UNIVERSITY OF SOUTHAMPTON

Faculty of Engineering and Physical Sciences
School of Electronics and Computer Science

Exploring Identity Assurance as a Complex System

by

Vincent Joseph Marmion

A thesis for the degree of
Doctor of Philosophy

September 2021

University of Southampton

Abstract

Faculty of Engineering and Physical Sciences
School of Electronics and Computer Science

Doctor of Philosophy

Exploring Identity Assurance as a Complex System

by Vincent Joseph Marmion

Personally identifying information (PII) are complex resources. Each item of PII, e.g., a fingerprint, holds a confidence-based utility that fuels identity assurance, i.e., processing fingerprints towards a desired confidence that a person is whom they claim. Each time we use an item of PII however, for identity assurance or otherwise, we inadvertently expose it to misuse. Exposure thus accumulates to deplete the confidence that may be extracted for subsequent identity assurance uses. Therefore, in terms of identity assurance, PII exhibit some of the properties of a commons, wherein resources are accessible to all, and whereby individual actions can affect the group. In this depiction of identity assurance, there is an underlying usage dilemma surrounding PII. This dilemma arises because coaxed by the affordance of the modern Web, PII of increasing veracity is being digitally exchanged, processed, and stored in ever-increasing volumes and varieties.

Towards a novel sense of identity assurance as a commons-esque system, this work combines empirical and agent-based simulation methods to investigate PII exchange between individuals and organisations. First, by repurposing Elo's (1979) ranking algorithm, I produce a unique user-centric measure of PII's personal utility by ranking identifiers based on the quantification of (N =125) users' willingness to disclose. These results also incorporate inter-contextual differences with a design spanning social, commercial and state-based contexts. Second, I qualitatively analyse 23 one-to-one semi-structured interviews regarding disclosure decisions. From this, I identify six super-ordinate classes of heuristics that users rely upon during disclosures: prominence, network, reliability, accordance, narrative, and modality, along with a seventh non-heuristics class; trade. Third, I combine my empirical results with theory to produce a dual-system decision model of users exchanging PII with organisations. Finally, I explore the dynamics of PII exchange via an agent-based simulation of my model that serves to illustrate the potential effect of interventions such as educating users or increasing competition. I show that our onus on disclosure self-management threatens the future efficacy of identity assurance methods.

Contents

List of Figures

List of Tables

Declaration of Authorship

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as: Marmion, V., Millard, D. E., Gerding, E. H., and Stevenage, S. V. (2019). The willingness of crowds: Cohort disclosure preferences for personally identifying information. In Proceedings of the International AAAI Conference on Web and Social Media, volume 13, pages 358–368
   Marmion, V., Millard, D. E., Gerding, E. H., and Stevenage, S. V. (2017b). The tragedy of the identity assurance commons. In Proceedings of the 2017 ACM on Web Science Conference, WebSci '17, pages 397–398, New York, NY, USA. ACM
   Marmion, V., Bishop, F., Millard, D. E., and Stevenage, S. V. (2017a). The Cognitive Heuristics Behind Disclosure Decisions, pages 591–607. Springer International Publishing

Signed:... Date:..................

Acknowledgements

My appreciation begins with my supervisors. To Dr Dave Millard, for invaluable patience, kindness, and for steering my focus, Dr Enrico Gerding for the astute tutelage, Dr Felicity Bishop for methodological support, and Prof. Stevenage for the opportunity.

My gratefulness also goes out to my family and friends for encouragement and support. In particular, Nicki Lewin, for her treasured guidance, Dr Maeve Marmion, the PhD trailblazer in the family, Dr Harvey, Dr Grey, and Dr Owsoski for their friendship and mentorship.

My gratitude extends to the Defence Science and Technology Laboratory (DSTL) for funding this project.

To all those facing a veil of fallacious fear, may you have the faith and the firmitude to venture forth, forever.

Definitions and Abbreviations

PII	Personal Identifying Information
CPR	Common Pool Resource
2*FA*	Two Factor Authentication
CIFAS	Credit Industry Fraud Avoidance System
GDPR	General Data Protection Regulation
ICO	Information Commissioner Office
K-Means	K-means clustering is a method of vector quantization
R	The R Project for Statistical Computing
NVivo	Qualitative Data Analysis Software
RQ	Research Question
PIN	Persona Identification Number
Phishing	Obtaining sensitive information using a false persona/entity
UK	United Kingdom
EU	European Union

Chapter 1

Introduction

Identity assurance is a process integral to high-stakes activities such as the police matching fingerprints or government services managing migration. Identity assurance is also an unavoidable aspect of living; we all frequently do it, and it frequently is done to us all. For example, a wine merchant assessing Alice's age from an ID card or the cat flap scanning a microchip when Bob enters his house. Likewise, greeting neighbours, boarding planes, sending an email, paying invoices, or going online each involve identity assurance principles.

Identity assurance is reaching a satisfying confidence level that a claimant is as they claim (Beres et al., 2007). In reaching this level of confidence, a relying party examines, formally or informally, a set of identifying credentials provided by the claimant. For the purposes here, a claimant is a natural person and, therefore, the identifying credentials are embodiments or artefacts representing the person. Moreover, the relying party is a digital system or intermediary, wherein confidence levels are critical to the process outcome that aims for unambiguous links to a single entity (Tsakalakis et al., 2017).

There is definitional variation in the legal treatment of identifying credentials. For instance, the US-centric legislation opts for ′personally identifiable information′ (PII) to include gender or biometric data items. In contrast, the European Data Protection Regulation (2016/678) (GDPR) has a broader encapsulation of data items, 'personal data', including data such as IP addresses. Essentially the GDPR recognises that with the right technology, process, and or supporting data, abstract identifiers can become linked or linkable data to a single identity, now or in the future.

In this work, the favoured term is PII, yet the broader reach of personal data remains. Essentially, there is value PII as a term. Chapter 2 examines each of PII components, personal, identifying, and information, each of which is a subtly different property. Therefore, when reading PII in this work, it refers to currently accepted items according to current legislation but tends towards a more comprehensive definition.

Despite the broad and often significant application of identity assurance, or indeed the significant ramifications of inadequate identity assurance, the state of identity assurance remains flawed. A persistence of individual misjudgements (Dhamija et al., 2006), imposter innovations (Tibbs, 2013), and organisation exploitations (Knight and Saxby, 2014), makes obtaining 100% confidence in an identity an unachievable goal (Uludag et al., 2004). However, using ever-innovative processes of extraction, i.e., biometric scanners, or a combination of methods, it is possible to get ever-closer to a person's true identity (Black et al., 2012). Nevertheless, this innovation occurs rarely without compromise.

Identity assurance is also more than a process; it is a socio-technical system, comprised of competing security and privacy desires, amidst commercial and social incentives, a system governed by national and global regulations (Solove, 2007; Nissenbaum, 2004; Cavoukian, 2008). Identity assurance is also an unstable system as each process innovation has potential consequences that take time to resolve (Vila et al., 2003). For instance, evolving from passwords to biometrics may increase assurance. However, it also provokes legal, social, and ethical considerations regarding privacy as it enables individuals to be covertly tracked across many systems (Jain and Nandakumar, 2012). Likewise, innovations such as those surrounding identity assurance can introduce a knowledge gap that necessitates users to play catch-up in understanding the technologies. This is a key user vulnerability as users are found to over-disclose or indeed mistakenly disclose to the (occasionally malicious) organisations with which they engage (Preibusch et al., 2013). Perhaps the most notable example of this is with the concept of digital technology and individual privacy, and for this reason, many of the lessons that we can learn in the context of privacy can be applied to identity assurance.

Vila et al. (2003) use a game-theoretic model to explore privacy as a social system based on the work of (Akerlof, 1970). In their model, they express privacy akin to a second-hand car market where users pay to check for faulty cars. They argue that the current consent model for PII disclosure costs the user in *time, effort or money* to discover whether an organisation is faulty, which in this context could be whether the organisation extracts more PII than required to sell for profit. These 'faulty' organisations can capitalise on the asymmetric information between user and organisation if it is deemed that users are unwilling to pay to discover any faults. Classifying their results as a free-riding problem, they describe an oscillation between users depending on others to discover a fault, yet many individuals with this same conclusion combine to enough ignorance that then tempts organisations to exploit the free-riding. In turn, this situation leads to more users inclined to pay to discover as there may be something warranting discovery, i.e., there is reduced trust in the market

(Figure 1.1)[1]. From their model, Vila and colleagues' results suggest the theoretic possibility of a long-term, mixed-strategy Nash equilibrium comprising some attending users, others ignoring, and some respecting organisations, others exploiting. Yet, despite the equilibrium, the authors maintain that this equilibrium is highly sensitive as innovations cause flux, meaning that exploitation persists.

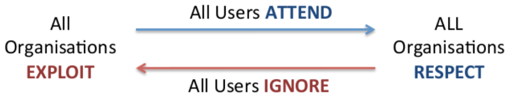


FIGURE 1.1: A Privacy Free Riding Problem (Akerlof, 1970; Vila et al., 2003).

While the game-model described by Vila et al. provides an insight into the persistence of data exploitation, the implication of any long-term sustainability is not addressed. That is, the equilibrium is sensitive to oscillation due to environmental changes repeatedly introduced by an ever-innovative approach to identity assurance, which in turn are simply outcomes of response to novel threats or new capabilities. This equates to an 'arms race' requiring ever-more identifier disclosure (Hong, 2012). During these oscillations the asymmetry of information between user and organisation is heightened, leading to a higher potential of exploiting organisations, contributing to an accumulating exposure of PII over time. This is an unwieldy cycle necessitating innovation which escalates the extraction of higher veracity PII.

In this thesis I build upon this idea of persistent exploitation and escalation, to argue that, along with some subtle properties of PII, the system of identity assurance can be considered as a system similar to a common pool resource (CPR). Moreover, like with many CPR scenarios, I argue that within this identity-based CPR there exists a sustainability concern. From this point, any reference to the identity assurance commons is for simplicity and analogy rather than a firm attempt to categorise identity assurance as being a CPR.

A CPR is any resource that is open to a group to exploit as individuals, but, it is also one that individual actions can have an impact on the group, e.g., our climate (Ostrom, 2008). In CPR scenarios, the general message involves a simple usage dilemma. On the one hand, there is *'the tragedy of the commons'* (Hardin, 1968), when resource depletion happens from over-use, wherein over-use arises from simple, seemingly harmless, yet self-interested, short-term decisions, within an environment lacking adequate regulations. On the other, there is *'the tragedy of the anti-commons'* (Heller, 1998), when wasteful resource management happens from under-use, wherein under-use arises from over-cautious decisions and protectionism, within an environment laden with regulations. With CPRs, regulation is a must, but, it must fit,

[1] A simulation of this system of oscillating behaviours is summarised in Chapter 5 as a prelude to the final simulation work.

but finding where along a continuum of governing decision is not trivial (Brede and Boschetti, 2009; Adjerid et al., 2013).

As with other CPRs there exists a PII usage dilemma in the identity assurance commons. In the identity assurance commons, items of PII are the collection of resources from which an identifying utility can be extracted. However, over-using PII can deplete this identifying utility in future uses. This depletion is perhaps more prominent in the digital age, wherein usage de-facto creates copies, and each copy in circulation adds doubt as to its legitimate user at any given time, e.g., using the same password on multiple systems (Ratha et al., 2007). Alternatively, under-using PII can waste this identifying utility in the present, which can enable illegitimate behaviour under a false identity, e.g., services using password systems for high-value systems when biometrics may be prudent. This usage dilemma is further complicated by the other properties of PII. It is easily overlooked that items of PII comprise of three properties; personal, identifying, informative. The identifying property has been discussed. Yet, items of PII are also information, and therefore, can be informative rather than identifying, e.g., a home address reveals probabilistic information about such things as wealth, education, and health. The third property is that PII are personal, and this property brings with it many subjective features regarding privacy, self-expression, and self-awareness. This means that PII is not exclusively used for identifying purposes, rather PII are being extracted and disclosed for a host of purposes, e,g, academic research. Consequently, PII of increasing veracity are being digitally exchanged, processed, and stored, with ever-increasing volume and variety, using ever-innovative methods of extraction, coupled with increasingly prevalent means of self-disclosure. This all amounts to a sustainability threat in the identity assurance commons.

1.1 Summary of Aims

My perspective is to consider identity assurance as a social system sharing properties and outcomes related to dilemmas depicted by the tragedy of the commons. One such property is that the actions of others can affect the group. Through this portal, our current, self-managed model of personal data regulation is a concern. At best, this model is insufficient as populations don't have the required tools or information to manage personal data disclosure adequately. At worst, this approach is inapt at preventing escalations in PII exploitation as people may not have the capacity to address personal data disclosure adequately.

Treating PII as a unique common pool resource for identity assurance, I aim to demonstrate how depletion of PII's identification value can emerge from straightforward, seemingly harmless, yet self-interested, short-term, individual decisions within a system of competing incentives. One objective is to understand better how individuals value PII within varying context. A second is to expand our understanding of how disclosure decisions unfold in the individual's mind. A third objective is to examine how individuals and organisations interact and exchange PII from a cost-utility outlook. These broad aims and objectives require interdisciplinary work a are limited as a result of high abstraction. The following questions aid these objectives,

RQ1 How are the properties of PII measured?

RQ2 How are users making disclosure decisions?

RQ3 How does context affect PII disclosure?

RQ4 How do RQ 1-3 combine within a system's view of identity assurance?

1.2 Research Approach and Contribution

The overarching method is to abstract a model of the identity assurance commons and uses simulation to explore the model. Figure 1.2 is a conceptual flow of how the research methods and designs fit together. Note that the intention of a model and simulation came before the empirical work. During the research, development and drafting of the pilot for Chapter 5, it became apparent that abstracting a realistic and insightful model of identity assurance would require specific empirical work. Therefore, these works are foundational to the development of the model, without which it would have required substantial assumption in their stead. The two foundation requirements fell into two distinct camps.

First, there exists no conclusive way to predict and or differentiate the intrinsic value a person puts on an item of PII from one instance of PII exchange to another. To, therefore, improve the realism of the PII exchange within the model, I produced a ranking of a set of PII according to a cohort of users' willingness for PII disclosure [RQ1]. This was achieved by repurposing Elo's (1978) aggregated ranking algorithm, within a three context design; disclosing to a social network (N=35), to a commercial bank (N=40), and to a government passport office (N=40) [RQ3]. This algorithm not only determines a ranking, but it also provides an underlying score that indicates a relative distance between the items of PII. This distance provides insight into how items of PII tend to cluster into distinct groups, while others stand apart.

My quantitative results add to the growing consensus that PII is a unique legal concept, requiring unique protections (Weitzner et al., 2006; Sullivan, 2013). The results contribute an alternative way to describe PII disclosure that goes beyond

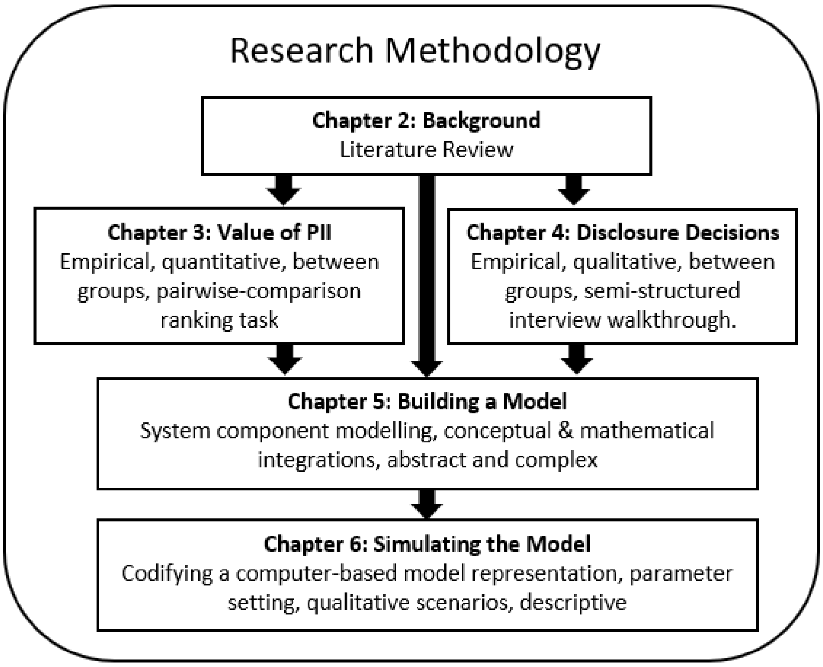

Research Methodology
Chapter 2: Background
Literature Review
Chapter 3: Value of PII
Empirical, quantitative, between groups, pairwise-comparison ranking task
Chapter 4: Disclosure Decisions
Empirical, qualitative, between groups, semi-structured interview walkthrough.
Chapter 5: Building a Model
System component modelling, conceptual & mathematical integrations, abstract and complex
Chapter 6: Simulating the Model
Codifying a computer-based model representation, parameter setting, qualitative scenarios, descriptive

FIGURE 1.2: Research Overview.

variety or frequency. They are a novel measurement of PII's personal utility and include insight into how context can affect users during disclosures. These findings also cast doubt on the ethics of practices that transfer PII across contextual boundaries, such as social networks sharing with governments, or governments sharing with commercial entities.

Second, the reliance on calculative rational agents within the free-riding model by Vila et al. (2003) fell short of contemporary thinking on how people actually approach disclosure decisions. So, from 23 semi-structured interviews, I provide a qualitative account of how users make disclosure decisions [RQ2]. These interviews cover two contexts [RQ3], reflecting two possible self-regulatory mindsets; entertainment and responsibility (Higgins, 1998). Adopting a framework set out by Metzger and Flanagin (2013) for credibility judgements, I find that users tend to rely on simple heuristics when making disclosure decisions rather than completing calculative measures.

My qualitative results contribute a novel framework of simple heuristics that users rely on during disclosure decisions. These results include the identification of six super-ordinate classes of disclosure-based heuristics: PROMINENCE, NETWORK, RELIABILITY, ACCORDANCE, NARRATIVE, MODALITY, and a seventh non-heuristics TRADE class. Moreover, they suggest that simplicity in disclosure decisions leads to herding type behaviours and misjudgements regarding risk in data extractions. This strongly suggests that regulatory efforts seeking to increase the autonomy of the informed user are inapt. Overall, this work adds evidence to a growing sense of the inadequacies of our current self-management model of identifier disclosure (Solove, 2012; Fairfield and Engel, 2015). This new framework of disclosure heuristics is used to incorporate more naturalistic user agents into my final model and simulation.

Finally, returning to the model. My empirical findings are combined [RQ4] within a more thorough game-theoretic model of data exploitation based on the model set out in Vila et al. (2003). The biggest challenge of this work stems from the *ideally* interdisciplinary nature of social simulation (Cioffi-Revilla, 2014). To this end, I work with a multi-disciplinary supervisory team that includes psychology and web science to combine the results from qualitative interview analysis, experimental quantitative analysis, game theory abstraction, towards an agent-based social simulation. A further challenge exists in balancing realism, generality, precision and tractability of the model. The empirical chapters aid these efforts. The quantitative work, revealing the personal utility of PII enables additional precision in accounting for the disclosure behaviours of users. Then the qualitatively produced heuristic schema enables additional realism in my user agents to move beyond the decision rationality prevalent in other similar models. The overarching output is to contributing towards a call to rethink how we regulate identity assurance, and specifically, how we regulate for the sustainability of PII. By setting out the complex issue of identity assurance within the framework of a commons, it is also a pragmatic step towards using well-established models to investigate what may drive or temper key dynamics within the commons.

Despite that social agent-based model as a methodology are gaining prominence for informing national policy (Edmonds and Gershenson, 2015), it is important to over-reach with the results, or in other words, it is important to mind the 'lure of artificial worlds' as the outcome is not empirical data (Bullock, 2014). That is, regardless of the quantitative feel of simulation, the result of this method for social systems is most powerful in a qualitative narrative rather than being used for predictive accuracy (Squazzoni, 2014). Therefore, conclusions are drawn from any of the simulation aspects of this work, also should be taken for their insight qualities rather than for any predictive qualities. Furthermore, insights garnered here should be taken from a UK perspective as there is likely some interpretation a bias from my own experiences, but also, as participant recruitment was exclusively within the bounds of the UK. The results will have their own generation influences, and this should be considered through any recommendations, and or any future extensions and replications to this work. I have little doubt that conflicting cultures, internationality or intergenerationally, would introduce many fascinating nuances. Nonetheless, the core principles of my findings regarding the manner by which disclosure decision are made, the clustering of PII values, and the notion of PII as a collection of depleting resources would likely remain agnostic of culture.

1.3 Structure

To recap and clarify, the structure of this work progresses towards the components comprising an agent-based social simulation. These are the environment, the agents, their interactions and adaptation, measurement and validation, each suggested in Bonabeau (2002); Cioffi-Revilla (2010), and Macal and North (2010). In Chapter 2, the background literature focuses on identity assurance as a problem environment, the unique properties of PII, the extraction behaviour of organisation agents, the disclosure behaviours of the user agents, and, more abstractly, the methods, and the perspectives of the validity of conducting a social simulation. Subsequently, in Chapter 3, a quantitative investigation provides a user-centric measurement of PII's personal utility. Then, based on interviews with users, Chapter 4 presents a set of heuristics used when people are making disclosure decisions. Chapter 5 presents model and simulation incorporating the previous findings. Chapter 7 summarises and details my future work. Figure 1.3 provides the chapter outline that maps onto the methodology overview in Figure 1.2.

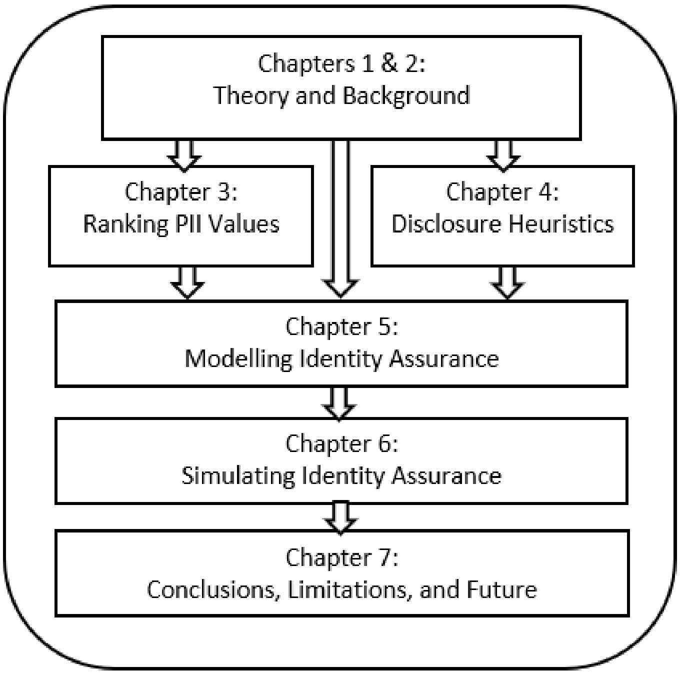


FIGURE 1.3: Flow chart of how the chapters are structured.

1.4 Publications

Each section of my work leads towards a single goal, yet they tackle diverse questions, this provides clear demarcation for publication. To date, sections of this work have been selected for publication on three occasions, with potential for a fourth. The ideas set out in this introduction, along with a comprehensive outline of the pilot model (Appendix C), were published as an extended abstract at 2017, Web Science Conference (WebSci) (Marmion et al., 2017b). My qualitative work in Chapter 4 was presented as a full paper at the 2017, Social Informatics conference (SocInfo) (Marmion et al., 2017a). The work outlined in Chapter 3 was presented as a full paper to the 2019, Web and Social Media conference (ICWSM) (Marmion et al., 2019).

Chapter 2

Background and Related Work

This chapter is a review of literature, concepts and models that I use, it comprises five sections. The first section introduces complexity as a theory and how it guides the world view of this work. The second discusses identity assurance as a complex problem environment, setting the bounds of identity assurance as a process but also as socio-technical system of PII disclosures and extractions, exposures and retractions. The third section focuses specifically on the unique properties of PII that make it a resource conflicted in its utilities, and how interactions over this resource concern the sustainability of the identity assurance commons. The fourth section then reviews the extraction of PII from the organisation agents perspective, and how interactions of this nature have escalated of the recent decades. Then the fifth section discusses the user agents in the system, and how they approach PII disclosure decisions.

2.1 Complexity Theory

Complexity theory is a catch-all for the scientific study of complex systems. A complex system is a network of interacting components that are independent of central control that gives rise to non-linear behaviours, and that adapt to feedback (Walby, 2007; Mitchell, 2009). A complex system takes other systems, that are often complex in their own right, as their environment (Walby, 2007). However, complexity theory can be considered as more than these abstractions, for it somewhat resists reductionist definitions (Morin, 1992). Turner and Baker (2019) provide an extensive overview of the current definition variation within complexity theory. Not only in definition, but also in the characteristics that one may assume to be evident within a complex system. Complexity theory is also a philosophy of science, aiming to divert one's perspective from a 'centralised mindset' that conceives systems as being of reduced or neatly nested parts, towards a complex systems view of overlapping, co-evolving, systems within systems (Resnick, 1997).

Complexity theory lacks a unified approach to align with the growing number of tools available to those adhering to the complexity paradigm (Walby, 2007). Nevertheless, as the perceived problems within the system of identity assurance typify a classic resource exploitation problem, formal analysis can be an aid (Brede and Boschetti, 2009; Vila et al., 2003). Many formal attempts to understand the confines of privacy and identity disclosure have relied on the principles of economics (Acquisti et al., 2012; Laudon, 1996; Preibusch et al., 2013; Varian, 1996). Such formalisation tends to look at the equilibrium solution of the system (Rust et al., 2002). Similarly, game-theoretic models offer an alternative method of understanding competing incentives and ultimately, any possible equilibriums. One particular merit in finding such equilibriums is in recommender systems, whereby the individual would 'play' against a service to optimise their identity exposure with security. If enough individuals interact this way, then the systems can optimise their strategy in real-time (Manshaei, 2011; Zhu and Zhu, 2009).

This work develops towards the method of agent-based social simulation to explore identity assurance as a complex adaptive system. The intention is to combine game-theory with simulation to harness the best of both scenarios (Bonabeau, 2002). That is the agent-based system that has a capacity for agents to adapt by learning from social and environmental stimuli, built on the underpinning of a static game-theoretic model of privacy economics (Vila et al., 2003). This approach encourages the bottom-up representation of a system and from the collation and interaction of simple individual rules (Axelrod, 1997). Such endeavours can subsequently support those of the remit to do so, e.g. policymakers to affect the system in a targeted manner (Edmonds and Gershenson, 2015) positively.

Simulation, and specifically, social simulation, is not without debate. In three generations of advancements, complexity and simulation science has encountered substantial criticism, doubt and definitional inconsistency (Alhadeff Jones, 2008). Principally there remains doubt as to the extent simulation science warrants a new philosophy or a new methodology, with added discussion as to whether it is a pseudo-science (Frigg and Reiss, 2009; Humphreys, 2009). While my decision to use simulation indicates to where I stand on many of these criticisms, combining simulation with empirical work is a prudent approach to mitigate any doubts and add rigour. The majority of the debate encapsulated by Frigg and Reiss (2009) centres around the following,

- **Metaphysical**: Simulations create some kind of parallel world in which experiments can be conducted under more favourable conditions than in the real world
- **Epistemic**: Simulations demand a new epistemology

- **Semantic**: Simulations demand a new analysis of how models/theories relate to concrete phenomena.
- **Methodological**: Simulations are a *sui generis* activity that lies in between theorising and experimentation

Frigg and Reiss set out to counter-argue each of these claims. They argue that while classifying simulation as a new type of science, that is the end of the novelty. Instead, Frigg and Reiss (2009, p. 4) perceive the value in debating the philosophy of simulation as contributing to the understanding of existing issues around, amongst others, scientific modelling, idealisation or external validity, rather than as exploring completely new and uncharted territory. On the other side of the debate, Humphreys (2009) directly challenges these claims and firmly sides with the idea that simulation science does raise new philosophical issues. Proponents of this standpoint do deem that computational, and simulation science introduces novel issues, but beyond considerations regarding aspects such as programming language choice, computational confines, simulation method, the debate regarding the validity of creating artificial worlds is firmly a valid one (Axelrod, 1997; Bullock, 2014; Cioffi-Revilla, 2014).

Kallemeyn et al. (2020) suggest that alongside this reductionist rejection, perhaps the emphasis on methodological pluralism is a core contribution of complexity theory. For this work, one that situates distinct research projects into a complex whole, it is not only pragmatic to adopt his complexity mindset but perhaps imperative to resist a reductionist framework. An open systems mindset to conduct multiple projects to understand better the same open and evolving system (Turner and Baker, 2019). The following sections outline the significant components of identity assurance as an open system.

2.2 Identity Assurance as a Problem Environment

2.2.1 Identity Assurance as a Process

Identity assurance is the confidence that arises from a process that determines if a person is who they claim to be, or that a person is who they are believed to be (Beres et al., 2007). We may distinguish such processes to include instances of an individual asserting their own identity, such as verification, i.e., you are who you say you are, or authorisation, i.e., you are verified and approved to proceed. Alternatively, an identity may be asserted onto an individual, such as with profiling, i.e. you are a person of type x, or identification, i.e., you are person x.

Predominately, these processes have been intuitive, such as visually recognising a familiar face, or perhaps judging a friendly one. For instance, our ability to recognise familiar faces remains a rudimentary but essential form of identity assurance that can suffice for transactions in small communities (Sullivan, 2013). In this latter situation, the face is the personally identifying information (PII) being processed by a recogniser to extract its identifying value (or utility) for some purpose such as personal comfort and trust.

Over the years, advances in technology have provided the ability to represent PII (or identifiers) as artefacts or 'tokens', to introduce new identity related affordance (Boult et al., 2007). For instance, authentication is the process of determining the provenance and truth of an artefact (such as a passport) from which a user implicitly asks the system to accept or reject a proposition, "am I, who I claim I am?" (Ratha et al., 2000). Such tokens are vital for modern transactions such as banking or international border control (Sullivan, 2013). A token could also be physical like a tattoo or scar (Schildkrout, 2004), or perhaps something more behavioural like a signature or a 'secret' handshake (Yampolskiy and Govindaraju, 2008). Such tokens in their own right provide additional confidence to bare intuition, Regardless of context and criticality, each identifier assessed adds to the confidence, for example, a signature, handwriting, or a post-mark each provide letter recipients an added degree of confidence as to the sender's legitimacy within every day correspondence, in the same way one may be asked for two forms of photographic identification to open a bank account.

However, with tokenisation comes compromise. Markings on the skin may act as a community identifier, but they also can profile the wearer as a target, an outsider, or an outcast. Writing enables remote, asynchronous communication, but this means of communication also enables interceptions. The development of print media meant that identifiers could be printed, distributed on mass, and stored indefinitely; attributes that have long been debated for their privacy implications (Warren and Brandeis, 1890). The crux of the matter is that tokenising identifiers enables impersonation; the antithesis of identity assurance. For instance, fake tribal markings to bypass suspicion, or forged letters to spread misinformation. Then, as such instances of impersonation are discovered, a response can be that more identifiers are sought, distinguishing marks become more intricate, letters carry wax insignias, and so forth.

Just as the advent of print media marked a sea change in the distribution of information, the development of digital media has been a pivotal time for identity related issues (Nissenbaum, 2004). Digitally tokenising PII has enabled remote identity assurance, which enabled the World Wide Web (or just the Web) to transition from the information repository of the 1990s to the interactive, social and commercial hub that exists today. However, the compromise is that digitally tokenising, then transmitting, processing and storing PII on the Web, reduces any natural protections afforded by 'practical obscurity' (O'Hara, 2010). Data are no longer dispersed across the globe in filling cabinets or disc storage, instead data sets are increasingly being stored on remote servers using cloud computing (Cavoukian, 2008). Moreover, this new practice coincides with the 'era of big data', wherein we mine PII, scrape it, process it, and consume it, in great quantity, for a multitude of reasons (Brown et al., 2011). Yet, the scale and efficiency of all this far outpaces the consciousness of most of the individuals engaging with these new technologies (Beer, 2009; Balebako et al., 2013).

2.2.2 The Opaque Disclosure and Extraction of PII

Directly alongside the exchange of identifiers for identity assurance, the exchange of personal data can assist innovations and efficiencies in a growing economy by focusing marketing efforts, personalising consumer experiences and informing research (Knijnenburg and Kobsa, 2013a; Tene and Polonetsky, 2012). In fact, organisations thrive on personal data and explore novel ways to exploit its value like any other commodity (Laudon, 1996; Rust et al., 2002). From a user's point of view, Van Zoonen and Turner (2014) classify identity disclosure in terms of a person's desire for expression (social interaction), for a transaction (e-commerce), or during instances of submission (state interaction). From an organisation's perspective each extraction of personal data can simply be viewed as part of an overall transaction (Sullivan, 2013).

Beyond one-to one or face-to-face transactions, when these transactions occur digitally an additional transaction is introduced through an intermediary which complicates the transparency of the transaction. Moreover, when conducted on the Web, the line is blurred further between the identifier extraction required for the purposes of identity assurance, and the identifier extraction for the purposes of profiling users. This is the case because whatever a person does via a Web service it involves a de facto transaction with one or, increasingly the case, multiple third party services (Balebako et al., 2013). For interaction with the state this transaction is quite straightforward, hence the notion of submission, for there is little room for negotiation when it comes to disclosure for the attainment of a passport or driving license. Whereas, in the pure transactional and the expression sense, a person has options regarding the medium chosen, competitor offerings or payment method. For instance, for self-expression which involves relating to others, the ability to authenticate an individual's identity aids relations and reciprocity whether online or not (Cozby, 1972; Millen and Patterson, 2003). Yet, while a user's principle aim may be to express and relate, almost invariably, a third party provides the service and thus acts as a data guardian.

More often than not, this transaction means a service may come at no (or a reduced) economic cost to the user who accepts some form of intangible cost such as displayed advertisements (Ur et al., 2012). As a result of these additional transactions, some organisations may harvest excessive data from their users in order to tailor their marketing efforts e.g. behavioural advertisements (Toubiana et al., 2010) or to supplement revenue by trading data (Balebako et al., 2013; Zhang et al., 2012). Importantly, many of these practices are legitimate, and organisations reveal them in a privacy policy (Furnell and Phippen, 2012). Nevertheless 'externalities' exist because the advantage of the trade is to the seller, yet there is a cost to the individual (Varian, 1996).

One line of thought is that individuals should be compensated for the trade of their data (Laudon, 1996). Such trades however, are often intricate and opaque. Figure 2.1, is an example illustration of an intricate ecosystem of data exchanges ranging from voluntary, observed or inferred data creation to final consumption by industries and public organisations (World Economic Forum, 2011). Users disclose personal data implicitly through everyday digital interactions, leaving a trail of metadata such as transaction audit logs that an organisation can observe and store (Clarke, 1988; Yampolskiy and Govindaraju, 2008). Enrichment models enable data aggregators to infer new personal information from volunteered or observed data sets, meaning the original volunteered data is often a gateway to richer data sets (Black et al., 2012; Bonneau et al., 2009). Ten years on, this high-level overview remains a succinct means to grasp the PII industry, yet, increasingly the only controllable variable is the volunteered personal data disclosed. What is perhaps misleading is that the linear order suggests a tangible end goal, another way to see this figure is a wraparound where consumption empower the services that promote more disclosure. It is circular, quick to innovate, and ultimately, hard to avoid. Perhaps fortuitously, PII being caught in the *nets* of big data is gaining public awareness (Debatin et al., 2009), and this awareness is disrupting individuals' sense of privacy and security (O'Hara, 2010). Yet, as Debatin and collogues find, it is only really those that directly experience privacy invasions that actually change behaviours. In other words, awareness is not enough, and waiting for experience is rather unsatisfactory.

2.2.3 Theft and Fraud

Digitally tokenising PII for use in remote identity assurance on the Web radically changes the manner of identity related crime (Anderson et al., 2014). The Web means that malicious entities distributed across the globe, are now effectively co-located within a single community, and able to remotely target an individual or service (Anderson et al., 2014; Kahn and Roberds, 2008). Subsequently, in the UK, the first quarter of 2015 saw 80% of reported identity fraud perpetrated online (CIFAS, 2017). In response, the UK Government announced a National Cyber Security Strategy which invested £1.9 billion in cyber security over a five year period

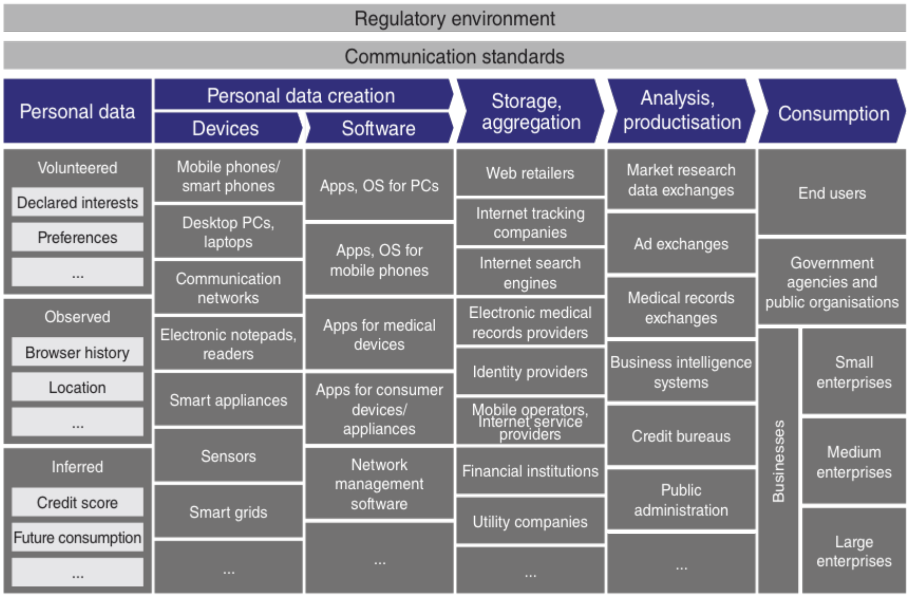


FIGURE 2.1: A representation of the intricate personal data ecosystem that details the movement of personal data from its creation to consumption (World Economic Forum, 2011)

(UK-Cabinet-Office, 2016). Despite such efforts, and that fact that some types of fraud decreasing, such as application fraud, the overall trend is that fraud is increasing year on year (Cifas, 2019). In fact, in the interim progress report for the National Cyber Security Strategy, there is as much regarding education and infrastructure rather than language of winning, or turning the tide of fraud, fully suggesting that greater challenges lie ahead (Office, 2019).

In the meantime, to maintain the integrity of online systems, the Web has become host to an increasing array of authentication methods each designed to increase assurance (or reduce uncertainty), and do so by requesting higher veracity personal identifiers in increasing volume and variety. Subsequently, by symmetry, users are faced with disclosure decisions for personal identifiers of increasing volume and variety. Consequently, privacy and identity preserving technologies are a growing commercial venture, yet in as much as they help, they also obfuscate the flow and reliability of the information available (Conger et al., 2013). Therefore, without adequate protection, over exposing PII is risking a principle facilitator of trusted interactions between individuals, organisations and nations across the globe (Sullivan, 2013).

For simplicity, I refer to any entity committing an attack on an identity assurance system as an imposter, when in reality this entity could be an individual, a group, or even autonomous computer programs. There are broadly two attack vectors at an identity assurance system, front-end and back-end, I focus here on front-end attacks. The difference is that, in back-end attacks imposters access a system by changing or circumventing the process, with methods such as Trojans, replay attacks or insider fraud (Lodge, 2007). Whereas, in front-end attacks imposters access systems by deceiving the process with methods such as stolen or spoofed identifiers (Jain and Nandakumar, 2012). The distinction is that whilst back-end attacks do not require seemingly legitimate identifiers, a front-end attack does.

Generally, imposters perform three distinct phases when circumventing security measures using identifiers (Newman and Mcnally, 2005). The first phase involves the theft of personal information. The second phase either involves the imposter redirecting the raw phase one data to directly access other accounts that may have used the same credentials, or breeding new data through some form of inference. The third phase involves the imposter fraudulently using the identifier for enriching phase two data, i.e., applying for new credentials or for some specific gain.

It is the third phase that typically brings hardship to the victim of the the fraud. For instance, new accounts, e.g., tenancy agreements or credit cards, can be opened with impersonation, reducing the credibility of the victims record (Anderson, 2007). An imposter can use an existing identity concurrent with the legitimate account holder e.g. for health insurance, or during an arrest (Taitsman et al., 2013). Alternatively, an imposter may aim to takeover a facility by changing the passkey, and then seek to maintain control or seek ransom, a type of attack growing inline with what is an accumulating array of digital assets (Aïmeur and Schonfeld, 2011). In addition, new account fraud does not need to directly concern an individual (although can also lead to tariff increases), instead, a fictitious or synthetic identity is used, perhaps to apply for official documents, that once in possession of the 'new' identity can be enriched with further documentation (Aïmeur and Schonfeld, 2011).

For an individual or corporate, the disruption can be considerable, often they must invest significant time and financial resources such as legal fees to rectify the situation (Aïmeur and Schonfeld, 2011; Newman and Mcnally, 2005). Moreover, despite insurances promoting a common assumption that costs can be recovered, any financial assurances fulfilled by an organisation are then recovered from the community of consumers through higher tariffs or prices (Anderson et al., 2014; Newman and Mcnally, 2005). Not least, the psychological impact can be all encompassing, causing severe emotional distress, and feeling 'betrayal', 'defiled', 'powerless' etc., including the eventual path towards suicide (Foley et al., 2009; Walters and Betz, 2012). We may summarise by siding with the notion and thus continuing this work under the assumption that while cases of identity based fraud continue to rise year on year, and the risk to corporate or living entities is real and potentially catastrophic.

2.2.4 PII Exposure and Retraction

Privacy in the context of this work is the reasonable expectation of information empowerment (Steeves and Pinero, 2008), including the ability to self-determine the processing of PII (Cavoukian, 2008; Fischer-Hübner et al., 2011; Westin, 2003). For clarity, the extreme end of the privacy spectrum is anonymity which is the ability to place absolute boundaries around an identity. Accordingly, privacy intrusion occurs via the circumvention of boundaries that include natural boundaries such as clothes and walls, social boundaries such as medical and legal, confidantes or temporal boundaries such as a phase of life or aspect of living (Bohn et al., 2005). Hence, a modern privacy concern, relates to the unexpected or unwitting transfer of personal information across perceived boundaries.

Regardless of whether or not a privacy intrusion occurs, innovative data collection methods can introduce what feel like deceptive methods (Balebako et al., 2013; Schwab et al., 2011). Even a perceived privacy breach can "impose a prominent affect on the way consumers view a product and its brand" (Afroz et al., 2013, p. 17). Therefore, current commercial practices that rely on a steady flow of user-generated content, yet fail to operate within users expectation of privacy could actually cause users to retract personal data (Balebako et al., 2011). This retraction could be through disengagement (Krasnova and Günther, 2009; Staddon et al., 2012) and self-censorship (Schauer, 1978) or perhaps through disruptive user behaviours such as falsification of data (Wirtz and Lwin, 2009) and obfuscation (Brunton and Nissenbaum, 2015).

Arriving full circle, a boundary for privacy purposes is in another sense a boundary for concealment that can enable an individual to misrepresent themselves to the disadvantage of others (LoPucki, 2002). In an expression sense, misrepresentation could be in the form of cat-phishing, essentially the act of fraudulently bypassing a person's internal authentication. In a transactional sense, concealment is the essence behind identity fraud. In a submission sense, concealment could represent an evasion of authorities.

The UK Information Commissioners Office has previously declared that organisations passing or selling personal details onto other organisations represents a principal privacy concern (Office, 2010). For instance, in one study, 88% of participants acknowledged privacy concerns stemming from online industries, with many reducing information disclosure accordingly (Krasnova and Günther, 2009). This study finds that when perceiving corporate threats, users tend to give less data, and when it is social threat, they give modified data. For identification processes, the extent to which users are willing to disclose personal information online determines the extent to which this identifying process is profitable for commercial dealings (World Economic Forum, 2011). What is ever clearer however, is that data policies are fraught with complexities, and that new frameworks are required, from which future polices might arrive that balance the modern thirst for data, and the innate sense of valuing privacy (Anne Toth and Lin, 2018). Missing from the World Economic Forum's assessment of the future of data policies, and the need to consider the characteristics of personal data, is the theory set out here, in that personal data in an identifying context has a depleting dynamic, and this should be incorporated into the policy design of the future.

In summary, disclosure and extraction of personal identifiers enables both legitimate and illegitimate endeavours. Users get access to secured services, buildings and borders, organisations can protect assets and or harness data for profit, and imposters can harvest and use personal data for nefarious gains. Identity assurance technologies aim to permit the former two and resist the latter. The next section takes a further look at how this situation is complicated when identity assurance is perceived beyond being a process, and instead as a system.

2.3 Identity Assurance as a Commons

2.3.1 PII as a Unique Resource

PII are the resource that enable identity assurance, yet each item of PII holds three properties that make it a conflicted resource in the world. Teasing these apart aids this work by providing a language to explore how these properties interconnect.

First, PII are information, and in any guise, i.e., knowledge or wisdom, information is a resource (Buckland, 1991). As a resource, PII's informative utility enables ventures such as marketing and political campaigns or social and medical research, consequently, the economic market for the processing and exchanging of PII is vast, and growing fast (Laudon, 1996; Schwab et al., 2011). This informative utility does not deplete in the same manner as resources such as oil or fish, instead it increases in exchange as the information is replicated (Eaton and Bawden, 1991).

Second, PII are identifying, hence they are the resource that enable identity assurance. Its utility is in providing a degree of confidence towards an entity's true identity. But, unlike the informative utility of PII, this identifying utility can deplete. Because, using an identifier can reduce its reliability for identity assurance, as the more copies the more likely it will be misused by an imposter (Jain et al., 1999). The effectiveness of identity assurance is dependant on finding identifiers that can reliably identify an individual time after time, yet due to depletion often this involves finding new *untapped* identifiers (Jain and Nandakumar, 2012; Guest et al., 2014). This is perhaps the quirk of identity assurance as a commons. In one sense, each item of PII can deplete, and therefore as a collective one may refer to these PII as a commons, but this would mean that identity assurance could actually be considered as a set of commons, with each type of PII potentially independent of the others depending on the systems employed.

Finally, PII are personal, and therefore, they have a personal utility. Whilst this utility aligns with PII's other utilities, it is not a clear alignment due to the concept of privacy, as in the ability to place boundaries around information (Bohn et al., 2005). That is, a person may disclose PII to exploit its informative utility. For instance, to peers for self-expression (Leary and Allen, 2011), to a business for some promotion (Grossklags et al., 2007), or maybe to a charity for altruistic purposes (Krasnova et al., 2010). Alternatively, they may disclose for its identifying utility. For instance, when forming an identity on social media (Knijnenburg, 2013), when transacting with a business (Sullivan, 2013), or when submitting to legal requirements (Van Zoonen and Turner, 2014). Yet evidence, under the banner of privacy concerns, repeatedly shows that this personal utility not only varies between individuals (Westin, 2003), but also within individuals as an effect of context (Martin and Nissenbaum, 2017) , or as an effect of change over time (Bohn et al., 2005).

Therefore, in as much as some individuals readily disclose, others invest time and money towards obfuscating their data (Brunton and Nissenbaum, 2015). Equally, as PII's informative utility does not deplete, some organisations seek novel ways to open it up to exploitation for public good, while others invest in privatising and enclosing previously 'free' information for private interests (Yakowitz, 2011). Whereas, regulation attempts are in the middle trying to strike a balance between transparency of data for individual and public good, with protecting the rights of individuals and organisations for private gain (Hess and Ostrom, 2003). One design feature of the UK Verify infrastructure is to overcome disclosure reluctance through a federation of online ID providers (Tsakalakis et al., 2017). In free-market theory, such a marketplace can provide organisation agility to respond to user disclosure preferences and resist unwarranted data aggregation.

2.3.2 An Identity Assurance Dilemma

The dilemma of identity assurance is in how to *maximally* exploit the identifying utility of PII, avoiding its depletion, amidst competing incentives to exploit its informative utility, all while adhering to its personal utility.

One perspective is that we under-use PII, and are therefore wasting potential in the reliability of identity assurance, this under-use, and therefore vulnerability in the system is somewhat indicated by 'reported' identity fraud in the UK continuing to rise year on year (Cifas, 2019). These crimes affect over 150,000 individuals at a cost of around £1.2 billion per year (Dickenson, 2015). Consequently, the solution for many organisations is to extract identifiers in ever-increasing volume, variety and veracity: More volume as services integrate further into our lives (Vidalis and Olga, 2014); more variety as multi-factor authentication protocols become an agreeable approach to modern security (Bhargav-Spantzel et al., 2006); and, more veracity as security systems innovate to gain access to highly identifying biometric identifiers (Jain and Nandakumar, 2012).

Another perspective, is that we over-use PII, which not only depletes the reliability of PII, leading to ever-more extractions, but likewise, over-use has sparked concerns that privacy is also depleting over time (Rust et al., 2002). Many of these concerns actually stem from the extraction methods, i.e., biometric systems, that are designed to improve identity assurance (Simoens et al., 2009). Consequently, a solution for a small, yet rising, selection of users, is to give less: Less volume as these concerns inhibit engagement (Krasnova and Günther, 2009; Staddon et al., 2012); less variety as new methods are met with suspicion (Nandakumar et al., 2008); and less veracity, as users obfuscate their data through falsification or anonymity methods (Brunton and Nissenbaum, 2015).

A further perspective is that this is a circular situation, in that organisations seeking more PII may inadvertently increase its exposure, risking privacy and theft, resulting in less disclosure, in turn, leading to organisations seeking more. This cycle, of personal identifiers being used for identity assurance, and becoming less secure and less useful for future identity assurance means 'that society may ultimately have to decide on a rate of identity theft that balances its preference for privacy with its tolerance for transaction fraud' (Kahn and Roberds, 2008, p. 31). Accordingly, society needs ways to examine this situation, not in isolation, but as an interdependent complex socio-technical system. Whilst this third circular perspective aligns closely to this work, there is slight distinction. That is, instead of top-down, in '**that society** may have to decide', rather what warrants further scrutiny is **how society's** decision is formed.

2.3.3 Identity Assurance as a Private and Public Good

Treating PII as a collection of resources for identity assurance leads to the potential of identity assurance being a good, i.e., a commodity. First, take privacy as an illuminating example. Privacy as a private good, or as a 'right' is long-rehearsed (Warren and Brandeis, 1890). That is, an individual has the right to self-determine the use of personal information, and that the benefits of privacy accrue to the individual. Whereas, privacy being a public good is also a compelling contemporary view brought about in the context of the digital networks (O'Hara, 2010). That is, the affect that a 'careless' individual disclosure has on that privacy of the individual, is often also born in some way by a community (Fairfield and Engel, 2015). One individual disclosed data often leaks data about other individuals. Imagine being denied a loan because your social network are not fiscally responsible, or simply because the people in your inner circle are inherent risk takers, this is a near reality and patents for such information technologies are being sought today (MacCarthy, 2010).

Whether privacy is a private or public good, places similar debates on identity assurance, this is because both relate to the exposure of personal information. However, there is an important distinction between PII and other personal data, owing to the effects of their disclosure. For instance, consider fingerprints as an item of PII, and religious belief as an item of personal information, there can be good reasons for keeping such information private, likewise there can be good reasons to disclose. However, the disclosure of religious belief has no affect on the strength of the belief (despite possible social repercussions), whereas the disclosure of a fingerprint can compromise its future use.

Consider identity assurance is a private good, in that the control, and therefore, the decision to exchange PII is a private decision accruing private consequences (benefits or otherwise). When many individuals come to the same conclusion, for instance exploiting their fingerprint for benefits of security and access, yet do so to a point whereby these organisations question the veracity of fingerprints, then one may begin to consider identity assurance as a public good. In that, the accumulation of these private decisions is of public consequence (benefits or otherwise). Because, if these organisations change, it can affect all users, not only those that were over exposing their PII. This is perhaps most apparent at the state level, wherein the need for the highest assurance can escalate from weaknesses in a method brought about by over-use (see Section 2.3). Considering identity assurance as a private or public good is not straight forward, and much work would be needed to expand on this, nonetheless the parallels with the privacy debates are there, and for this work, they provide a common language from which to describe cause and effects.

As a subset of all personal information, one may look to privacy regulations as a safeguard against over-exposure of PII. Privacy affords us the ability to place boundaries around personal data and personal actions (Bohn et al., 2005). This definition of privacy makes privacy inseparable from identity assurance. Without some boundary, the 'private' exchange of PII between individuals and services for identity assurance would be a misnomer (Jain and Nandakumar, 2012). In this regard, a tendency to treat privacy and identity assurance as a zero-sum situation, i.e., to increase one at the direct expense of the other, is giving way to a more co-dependant mindset (Cavoukian, 2008). That is, that dichotomies of winners and losers, are considered, at best limited, perhaps wasteful, and at worst misdirecting (Halperin and Backhouse, 2008; Cavoukian, 2011; Strahilevitz, 2013). The concern is that such thinking promotes what Resnick (1996) would term 'a centralised mindset', within which, centralised solutions are compelling, more privacy, more security or perhaps more openness.

Yet we know that PII is a special subset of personal information, broadly forming our 'transactional identity' (Sullivan, 2010), providing a means for *transacting* social, commercial or state activities on the Web. So, simultaneously one can discuss the incentives to promote and to restrict the use of PII on the Web. It is argued here that these conflicting incentives render privacy regulations as inadequate, and that a *more nuanced* regulatory approach is required rather than simply *a more* approach (Katos et al., 2010; Anne Toth and Lin, 2018).

2.3.4 The Tragedy of PII

Dilemmas such as in those in contemporary privacy debates discussed by Fairfield and Engel (2015) and O'Hara (2010), and likewise those in this work on identity assurance seem to parallel scenarios involving 'common pool resources' (Ostrom, 2014), e.g., fisheries, forests or public grazing pastures. These CPR scenarios can be explored in terms of *'the tragedy of the commons'*, whereby, simple, seemingly harmless, yet self-interested, short-term decisions can aggregate towards the detriment of all (Hardin, 2009). For instance, a single self-interested individual deciding to over-fish, may have negligible affect on a fisheries sustainability, whereas, many individuals independently coming to the same decision may cause depletion.

Staving depletion in these 'classic' usage dilemmas generally requires governance from some form of self-organisation or third party regulation (Ostrom, 2014). On this note, it is also important to not simply advocate more regulation, as over protective regulation due to fear of exploitation and depletion, can inadvertently result in the value of the CPR remaining largely unrealised. An outcome that is referred to as *'the tragedy of the anti-commons'* (Heller, 1998). For instance, overly strict catch quotas despite an abundance of fish stock frustrates the potential of the fishing community.

These two 'tragedies', are the two extremes of a continuum (Brede and Boschetti, 2009). This continuum is traversed with two interdependent yet distinct dilemmas; an individual dilemma as to how much resource to use, and a regulation dilemma as to how protective to be (Ostrom, 2008). Similarly, one may also describe the 'tragedy of PII' as a continuum resulting from usage and regulation dilemmas.

At one end of the continuum, we imagine a world with too little PII exchange. In this world; the Web is a less commercial place as users remain reluctant to exchange financial details (Udo, 2001), it is a less social place, as the social networking model is dependent on a steady source of personal information (Knijnenburg and Kobsa, 2013b), and it is a less efficient place For instance, the UK Government's vision ('Verify') to digitise all central systems is reliant on users willing to exchange the requisite PII, without which such systems will remain in the snail mail era. Equally, it would be less social, less commercial, and less efficient, if adhering to strict regulations made the extraction, processing and storing of PII prohibitively cumbersome for organisations. In essence, without some level of digital identity assurance, then trust in social and commercial interactions online would be limited to something more akin to blind faith.

At the other end of the continuum, we can envisage a world with too much PII exchange. In this world, users are highly willing to digitally exchange PII, moreover, regulations are such that organisations can cheaply extract it, process it, and store it. Online social networking may boom, online markets would be vibrant, and central infrastructure systems would be integrated into our daily digitized lives (Yakowitz, 2011). But with too much exchange, come costs. There is a cost to individual privacy and with that individual security, as access to an accumulating set of personal information becomes a click away, and another cost to security, as imposters have many more opportunities to obtain these valuable sets of PII (Cavoukian, 2008; Jain and Nandakumar, 2012). Maintaining the desired level of identity assurance, would require the exchange of PII in ever-increasing volume and variety to stay ahead of the imposters. The end game however, may be zero anonymity and therefore, total and constant surveillance.

As with the 'classic' CPR dilemmas, staving the depletion of PII reliability requires some form of self-regulation or third party regulation. Formal, third party regulation does exist. Namely, the UK Data Protection Act (1998), more recently, the European Data Protection Regulation (2016/678) (GDPR). Moreover, while currently stable, the GDPR regulation and the interpretation of Identity Assurance could fragment as the UK's departure from the EU matures (Tsakalakis et al., 2017). These regulations detail how organisations must obtain consent for PII extraction, and also, the bounds to which they can process it. In addition, many forms of informal self-regulation exists. For instance, the non-disclosure of data, the retraction of data, using an alias, or private browsing (Brunton and Nissenbaum, 2015).

Current formal regulation is not overprotective rather it is inadequate, moreover, that informal regulation is unfair. In part this is because, unlike in 'classic' resource dilemmas largely involving financial gain, within the context of PII exchange there are interplaying social, commercial, and security incentives (Krasnova et al., 2010; Sullivan, 2010; Jain and Nandakumar, 2012). For instance, commercial incentives exist in terms of organisations extracting excessive identifiers (Laudon, 1996; Toubiana et al., 2010). Or perhaps usability incentives exist from the perspective that identity assurance is simply a means to an end, and therefore, frictionless and covert extraction of identifiers holds much appeal to both parties (Preuveneers and Joosen, 2015). Equally, the social incentives for individuals to disclose identifiers should not be overlooked (Krasnova et al., 2010).

Naturally, many more parts comprise this complex socio-technical system, but these are principle components that receive attention in their own right, and herein lies a considerable research challenge. How do these parts fit together, interact, adapt, and co-evolve, and how are the resulting dynamics subject to change? A more thorough sense of these questions is needed, and is hence approached in this work, this along with much future research in the field and legal effort will be intrinsically linked to the success of any future regulatory attempts.

2.4 Extraction Escalation in Identity Assurance

To achieve identity assurance, an organisation chooses a method to extract personal identifiers from a claimant (the user). These methods develop over time to reflect the technology available, and the efficacy of any threats, but they also evolve with the behaviours and engagement of users as the collective action of individuals feedback into the system (the commons) as a whole. While these developments have been gradual, I describe three broad divisions under the headings of knowing and having, presence and being, and living and doing. The aim is to provide an abstraction of the methods of PII exchange to illustrate an increasing veracity of PII extracted, coupled with an increased passiveness in the disclosure. Moreover, this section explains how these milestones occur in a manner that disadvantages the individual because the exchange contains much asymmetric information in favour of organisations (Acquisti, 2004b; Vila et al., 2003), resulting in methods that are increasingly outside the user's understanding (Beer, 2009; Preuveneers and Joosen, 2015), all while users continue to misplace confidence regarding the risk of over-exposure (Solove, 2007).

2.4.1 Knowing and Having for Assurance

Most early and many existing online authentication practices, rely on either what you have, i.e. tokens, what you know, i.e. passwords, or a combination of both. A system and user agree a passkey (password or physical token) that the user provides to verify legitimacy. The ideal secure approach is a one-to-one relationship between the passkey and the system, relying on the uniqueness, intricacy and privacy of the passkey.

In practice however, users manage their passwords on a few-to-many basis, opting to recycle passwords or use highly predictable words (Jain et al., 2004). For example, Gaw and Felten (2006), found that to reduce the cognitive burden the majority of undergraduate students only had three or fewer passwords, despite in some cases having eight times as many accounts. Consequently, a weak-password enables an imposter through brute force, dictionary attacks or phishing to complete phase one (Lin et al., 2001). Then, following Gaw and Feltons findings, the likelihood is that the same password has been recycled for several other systems. A successful breach of one password (phase one attack) can support the attack of multiple systems (phase two).

Subsequently, security tactics are deployed to minimise users relying on common or predictable passwords, such as mandatory composite passwords, e.g., My?a$$w0rd. Alternatively, a tactic may be to spread the risk of a compromised identifier by using multi-factor authentication, e.g., chip and pin. Whilst these tactics increase the strength or 'entropy' of a known or held identifier, there remain key vulnerabilities. First, the increased effort leads to users increasingly writing passwords down or recycling them, and less likely to update them at recommended intervals (Komanduri et al., 2011). Second, it increases the rate of forgotten passwords and as a result, significant organisational resources are required simply to manage password-reset requests (Inglesant and Sasse, 2010; Uludag et al., 2004). Third, knowing and having methods cannot distinguish between a legitimate user and a person simply in possession of the tokens or knowledge (Ratha et al., 2006). Therefore, despite their widespread use, token and knowledge-based systems are gradually giving way to biometric-based systems (Inglesant and Sasse, 2010; Wayman, 2008).

2.4.2 Presence and Being for Assurance

In a biometric system, cameras and sensors capture the physical attributes of a person such as a fingerprint or iris scan and use these captured attributes as a unique passkey. These systems eliminate vulnerabilities associated with weak passkey reuse or frequent passkey resets as they rely on an individual being present (Nandakumar et al., 2008; Uludag et al., 2004). Furthermore, a physical biometric is significantly more difficult to steal, transfer or replicate than knowledge or a token (Ratha et al., 2001).

Compared to token or knowledge-based systems, attacks on biometric systems are significantly more resource intensive for an imposter (Nandakumar et al., 2008; Roberts, 2007). An imposter would need the skill to spoof the biometric, the opportunity (or stomach) to steal it or coerce an individual. Yet criticism of early biometric identity assurance systems has impeded widespread adoption (Inglesant and Sasse, 2010; Uludag et al., 2004). A prominent criticism centres on the irrevocable quality of biometrics, which increase the potential consequences of a compromised metric (Ratha et al., 2006; Rathgeb and Uhl, 2011). By default, a biometric system requires users to reuse identifiers, i.e., one fingerprint for multiple systems. A successfully spoofed or obtained biometric (phase one attack) can greatly enhance a phase two attack (Gaw and Felten, 2006).

Subsequent advances in biometric technology have confronted these weaknesses, including liveliness or presence testing (Sandström, 2004), multi-factor biometrics, i.e. fingerprint and retina scans (Bhargav-Spantzel et al., 2006), and even consent biometrics that can detect such attributes as duress via a retina scan (Yang et al., 2013). However, the issue of irrevocable identifiers persists. Cryptography is key to a solution yet such solutions have a usability trade-off for the users. Cryptography techniques can essentially create a unique identifier by combining a pure biometric image with a private key from the system (Chikkerur, 2008; Ratha et al., 2007; Rathgeb and Uhl, 2011). Hence, a single biometric attribute can act as a unique cryptography key for numerous systems, limiting the potential for multi-system compromises occurring from a single biometric compromise (Ratha et al., 2007). Unfortunately, the advantage in terms of cross-system security and the subsequent boost to the user's privacy, introduces high false reject rates as legitimate identifier images fail to match due to small variations in the image capture process, i.e., rotations of the finger on the scanner (Ratha et al., 2007). This trade-off reduces a key usability advantage over knowledge systems. Consequently, obtaining adequate usability during field tests of encryption remains a challenge for successful integration into systems that simply store the raw image. Meaning, "lingering concerns [remain] about the security of biometric systems and potential breaches of privacy resulting from the unauthorised release of users' stored biometric data" (Jain and Nandakumar, 2012, p. 87).

2.4.3 Living and Doing for Assurance

The 'knowing and having' and the 'presence and being' authentication systems only resist an imposter at the point of entry using static one-shot authentication. Once past this initial check, there is little to no resistance on many systems. This fortress model of security whereby the hard outer is believed to protect the soft inner has proven over time to be ineffective against sophisticated threats (Simpson and Foltz, 2017). Hence, recently, a third wave of identity assurance methods have emerged, which perform dynamic and continuous authentication. Whilst still largely based on physical identifiers, this next generation of assurance techniques draws on features of user actions to form a behavioural metric e.g., keystroke records. The attraction of these methods is that they continually authenticate, therefore, thwarting an imposter that circumvented the one-shot authentication stage (Niinuma and Jain, 2010). This reliance on user agency is why I differentiate them as living and doing methods. To avoid the obvious usability overheads, these methods covertly observe the behaviours of users as they engage with a system (Yampolskiy and Govindaraju, 2008). In essence, these methods require a certain amount of user surveillance to initially establish a behavioural baseline then match this with future behaviours, flagging anomalous behaviour (Niinuma and Jain, 2010).

These methods signify a shift in identity assurance from volunteered information that is explicitly disclosed by the individual, to a complex mix of volunteered and observed data (Yampolskiy and Govindaraju, 2007). Instead of a single high value identifier such as a biometric trait, a collection of low-value, almost by-product, data items can create a high veracity behavioural biometric (Clarke and Wigan, 2011;

Eckersley, 2010). For instance, Eckersley found that the way a web browser is personalised, i.e. add-ons, cookies, history or bookmarks, could act as a unique identifier used to track individuals across online activities.

During this evolution of authentication, the problem of users reusing passwords or tokens being duplicated, has been addressed by the use of biometrics that explicitly link to a person. But, the dangers of losing a token, or secret, have been replaced by a danger of compromising an irrevocable biometric that can identify and cross-match an individual across different data sets. Black et al. (2012) suggests that a person's disparate and peripheral identities lead back to a core 'SuperIdentity'. This is the case, even when using what a user may deem to be innocuous, non-identifying information (Hildebrandt, 2009; Solove and Schwartz, 2011). Certainly this is a powerful tool to identify malign individuals who exploit pseudo identities. However, this capability can also reduce the privacy boundaries that individuals believe exist to manage everyday living. Moreover, the advent of linked data methods can only make this capability and concern more pressing (Bizer et al., 2009).

In essence, the move to behavioural biometrics is services attempting to overcome technical frailties by extracting more personal identifiers. The subsequent concern raised by privacy researchers is that this threatens an individual's information self-determinism, because users are decreasingly aware of what identifiers have been exchanged and that this exchange is in effect a direct trade-off between privacy and security (Beer, 2009; Cavoukian, 2008; Lash, 2007). Therefore, despite recognition of the enhanced security and usability, the boundaries that people rely on to maintain comfortable levels of privacy start to blur as authentication methods can trend towards methods that, at their heart, enable covert and continued surveillance. Table 2.1 aims to encapsulate this transition, and hence is a central component of the overall model development.

TABLE 2.1: Trade-off Summary in the Evolution of Identity Assurance

Knowing and Having	**Presence and Being**	**Living and Doing**
Token and Password *E.g. Chip & Pin*	*Physical Biometrics* *E.g. Retina Scan*	*Behavioural Biometrics* *E.g. Keystroke Records*
✓Easy to Implement ✓Ubiquitous ✓Impersonal ✓Revocable ✓Many to Many	✓Hard to Steal/Forge ✓Hard to Lose ✓Novel	✓Low Intrusion ✓Dynamic ✓Front to Back security
✗Stolen/Forged ✗Forgotten/lost ✗Transferable ✗Static (Front)	✗Can be Intrusive ✗Static (Front) ✗Irrevocable ✗One to Many ✗Cross-matching	✗Cross-matching ✗Irrevocable ✗Surveillance ✗Covert ✗One to All

Note: As usability and security issues are overcome, authentication methods tend to rely on passive surveillance methods, or the capture of high value identifiers.

2.4.4 Political and Legal Oversight to Extraction

Identity assurance is not conducted without oversight. In the UK, it is a legal requirement for an organisation to provide outlines of the operation and treatment of personal data (Data Protection Act, 1998), with the assertion that personal data will be;

- accurate
- kept safe and secure
- used fairly and lawfully
- used for limited, specifically stated purposes
- kept for no longer than is absolutely necessary
- handled according to peoples data protection rights
- used in a way that is adequate, relevant and not excessive
- not transferred outside the UK without adequate protection

The recent GDPR regulation extends these protections with such changes as: including online identifiers, location data, and genetic data; increasing fines; and providing right to erasures. Perhaps the most salient difference is that the GDPR regulation is fully underpinned by the mechanism for consent, with clear notices and opt-in controls. However, whilst legal doctrine continue to set the boundaries for what can and cannot be done with personal data, this process can take a considerable time for the general rules to propagate to the specific. So, there remains the need for immediate and evolving policy and technical guidelines for compliance, such as in Tsakalakis et al. (2017).

It is incumbent of governments and large organisations to set such guidelines within the legal framework. However as Solove (2007) points out, often the practices of such organisations are bolstered by increments in the capacity of data processing, and that this is seemingly outpacing the scope or reform rate of the laws and regulations. This has the potential to leave a legal void within which only a hazy and inconsistent understanding of what is right and wrong persists. For instance, in 2014, the UK Government released guidelines for identity proofing and verification (UK-Cabinet-Office, 2016). These guidelines state that a single piece of evidence cannot be proof of an identity for their central systems. Instead, an agreed combination of evidence can create a level of assurance. Furthermore, the guidelines included a mandatory requirement for the inclusion of biometric identifiers, including activity monitoring at the highest level (4 of 4) of identity assurance.

On the surface this seems sensible, however, what scenarios warrant the level 4 and therefore maximum data extraction are subjective to the organisation. Therefore, other services, less aligned with national security endeavours, can also set their own high assurance requirements and accordingly aim to extract identifiers of high value. In situations where identifiers are covertly extracted, such as those in the living and doing category of authentication, organisations may be acting within current legal frameworks, yet the full social effects of these modest departures are still to be understood. As the legal and social understanding matures, it may be that retrospectively organisations may find that many practices deemed normal, are considered socially and ethically unacceptable to users.

In 2014, In Spain, a case was argued on the basis that information relating to a Mr Gonzalezs past, specifically, legal proceedings for the recovery of social security debts in 1998, had expired in relevance, yet, were nevertheless widely available through Google Search (Google Inc. v Gonzlez, 2014). The judgement supported the argument ruling that the time lapsed (i.e. temporal boundary) since an event can render that event irrelevant, and subsequently, ordered Google Inc. to remove this information from its search results. Essentially, despite the ruling being upheld on Article 7 and 8 of the Charter of Fundamental Rights of the EU (2012/C 326/02), this was considered a test case on the right to be forgotten and to erasure as described in, and recently ratified in Article 17 of European Data Protection Regulation (2016/678).

This regulation essentially prevents the indefinite storage and exchange of electronic data (Tsesis, 2014). The ruling was controversial, because it did not remove the information source, but simply restricted the mechanism that brought expired information to prominence. This meant, that despite Google Inc. defending their actions that they only process what is already public information, the judgement described Google Inc. as a data controller and accountable to the new Data Regulations.

The practicalities of this ruling and 'the right to be forgotten and to erasure' remain highly contested, a common predicament as regulatory and legal attempts to protect and define privacy are fraught with contextual ambiguities (Nissenbaum, 2004). Nevertheless, this case helps to highlight three things; boundaries such as described by Bohn et al. (2005) remain valued, modern technology can unacceptably increase the availability of PII across these perceived boundaries, and there remains forums and desire to discuss and resist aspects of technological capability that are misaligned with social expectations. Martin and Nissenbaum (2017) add an empirical investigation to Nissenbaum's 2004 influential theory on privacy as contextual integrity. Essentially, that context of data collection impacts the integrity of the data use. This study focuses on the presumption that public records resist privacy claims by nature of being already public. The authors find that judgements about the processing of public records misalign with policy. Moreover, consistency in the findings points to an unease with the existence of data brokers and the role they play in reducing the practical obscurity of public records.

2.5 User Agency During PII Exchange

To this point, this review has focused on identity assurance as both a process and a commons type system, it has focused on the technology and the interaction with the technology, legitimate or otherwise. However, here the review peers deeper at the rationale behind user's interactions with identity assurance technologies, and, how this fits with the mechanisms regulating such interactions. Essentially, this is a look at the potential disconnect between the theory of PII disclosure, and its practice, with the overall intention of being able to abstract user agents with realism and not idealism.

2.5.1 Intentions and Behaviours

Research in the area of behavioural economics has studied the monetary utility of PII from a personal perspective (Acquisti, 2004b), wherein individuals often attribute a relatively low value to their PII in exchange for desired services (Grossklags et al., 2007). However, work regarding the privacy paradox repeatedly tells us that relying on actual behaviours, while initially seeming to be a sound approach, actually misrepresent the personal intentions and ideals of disclosure (Norberg et al., 2007). This paradox involves the distinction between a person's valuations of personal information in terms of their intended, versus actual behaviour. Consequently, users disclose more information than they themselves consider appropriate. This important observation has captured the attention of researchers from an extensive array of disciplines including law, social psychology and behavioural economics. Largely the attributions of these scholars fall into three camps.

In the first camp, there is a perceived tension between a person's desire for privacy and their desire for self-expression or gratification. For instance, some disclosures are seemingly the result of individuals weighing up the value of personal data against desirables such as usability, convenience, publicity or entertainment (Acquisti, 2004a; Acquisti and Grossklags, 2007; Krasnova et al., 2010). Because, in practice, disclosure ideals are outweighed by a combination of immediate, often small, gratifications (Acquisti, 2004b; Grossklags et al., 2007) and perceived security benefits (Udo, 2001), or social norms (Zafeiropoulou et al., 2013) and other such heuristic-based decisions (Sundar et al., 2020) (see Chapter 4).

In the second camp, a distinction is made between the risk-based perception that form a person's abstract intentions, and the trust-based perception that account for a person's specific disclosure decision (Norberg et al., 2007). In this manner, the discloser can self-assure behind sentiments of having nothing to hide (Solove, 2007), or it won't happen to me (Krasnova and Günther, 2009). Moreover, actual disclosures are complicated by information asymmetry because compared to the organisations that control data extraction, users poorly understand the extent of data extraction, the purpose of each extraction, the destination of the data, and the risks that exist (Vila et al., 2003; Preibusch et al., 2013; Acquisti, 2014). Finally, these decisions are also uncertain, as, once third parties are involved, the distribution paths can continually splinter as the technology evolves (Weitzner et al., 2006).

Lastly, in the third camp, the paradox is explained by the vast nuances of the disclosure in terms of contextual differences. Whereby, individuals may be happy to disclose some personal data to their peers, but not necessarily to organisations, or conversely, release another set, for example, to a hospital but not their peers (van der Velden and El Emam, 2013). Whether a user is in dealings with a government, in the process of self-expression or perhaps merely shopping, can affect their willingness to disclose (Van Zoonen and Turner, 2014), it can also alter their perceptions of risk (Higgins, 1998).

There is a clear temptation here to encapsulate these behaviours in a top-down conceptual model of disclosure. For instance, Van Zoonen and Turner (2014) forward a concept of identity that distinguishes the agency of disclosure in terms of transaction, submission and expression maps neatly to the three camps above. However, such conceptual models could cause false confidences as they fail to represent the internal structures aligned with agency, and that a person's perception of current norms shapes future norms (Zafeiropoulou et al., 2013). As a result one may think of many disclosures as a result of herding effects where others' decisions influences the decision rather than a true personal decision (Acquisti et al., 2012).

Despite the discord between user's intention, knowledge, and behaviours, due to the self-deterministic mechanism of consent, regulations place the consequences of disclosure with the discloser (Solove, 2012). Also, as regulations do not distinguish online contexts in the way that users do, organisations such as data brokers often transfer PII outside of the user's intended remit (Tsesis, 2014). With users readily disclosing, and retaining the risk, organisations are then emboldened in extracting and trading more and more PII (Vila et al., 2003; Tene and Polonetsky, 2013b). This may be for legitimate security purposes, such as Apple Inc. adopting fingerprint enabled access for its iPhone 5 in 2012. Alternatively, it may be for non-security purposes, such as Rovio Inc. extracting location data via the Angry Birds game. Either way, despite the iPhone 5 being hacked on day one (Goodman, 2015), fingerprint enabled access (and increasingly face recognition) is now a standard feature of modern phones, and despite unease at Rovio's tactics, it is one of a growing list of games and apps now extracting PII as a norm (Balebako et al., 2013).

Vast numbers of users are either engaging without consideration, without understanding, or with an eventual acclimatisation to PII requests (Tene and Polonetsky, 2013a). Yet, consciously, reluctantly or unwittingly, users engaging with one generation of technologies only serves to spur the next generation (Flanagin et al., 2010). This leaves unwilling users to face a diminishing choice between non-disclosure or non-engagement (Staddon et al., 2012). From a macro view, these actions combine towards situations whereby common systems introduce PII of increasing identifying utility, which in turn, normalises further everyday uses, a concerning and potentially unsustainable escalation. Subsequently, as behaviours can belie the intrinsic value of PII, using observations of actual disclosure to assess the privacy interests of users only favours those advantaged by disclosure. This situation motivates this investigation into the personal value of PII to individuals, not regarding what people disclose in a given circumstance, but on what they prefer to reveal. That is, relying on behaviours leads to escalation in PII disclosure, and the ultimate concern in this work is that this escalation is unsustainable.

2.5.2 Consent as a Contract for Identifier Exchange

Regardless as to whether a conscious decision is made or not, when a person discloses personal data to an online system, they are doing so within a self-managed consent model, meaning consent is deemed as an explicit decision (Solove, 2012). Consent is a mechanism to define disclosure boundaries and enable effective online interaction, communication and information exchange (Whitley, 2009). The consent process enables a user in determining the value of disclosed data and to where these data are disclosed. Accordingly, during an interaction with an online service, governments impose a legal requirement for a privacy policy and/or a set of terms and conditions that outline the general rules by which the user and the service are to abide.

Unfortunately, whilst these rules are extensive and wide reaching, there remains controversy regarding the efficacy of consent. Tene and Polonetsky (2012) find that, rather than being a helpful consumer document, privacy policies only equate to service liability disclaimers. For instance, consent may be sought so that that a service may, unless "otherwise informed, be free to use consumers data in any (presumably legal) way it sees fit ... placing the entire information protection burden on the consumer" (Donna L. Hoffman, Thomas P. Novak,, 1999, p. 82, parenthesis in original). It is therefore, the individual's responsibility to decide if they are satisfied with external privacy assurances or, if they wish to investigate further themselves.

Such concerns regarding consent are highlighted when placed in contrast with the medical consent model which has been essentially stable since the 1960s (Whitney and Mccullough, 2004). Consent in medical treatment classifies simple consent as that requiring a brief explanation of actions followed by the patient agreement or refusal, any discussion of risks and alternatives are included at discretion of the medical staff. Whereas, informed consent requires a detailed discussion of the procedure, purpose, risks and benefits of proposed interventions. In essence, the common factor as seen in Table 2.2 that determines whether informed consent is required, or not, is high risk, where in contrast, low risk procedures only require simple consent (Whitney and Mccullough, 2004). Following this basic framework it is possible to further understand the efficacy of the consent model for disclosure decisions, whereby it is incumbent on the individual to self-inform for consent.

Table 2.2: The Requirements of Informed Consent from a Medical Perspective

		Risk	
		Low	**High**
Certainty	**Low**	Simple Consent	Informed Consent
	High	Simple Consent	Informed Consent

Note: High risk of medical procedure dictates informed consent, whereas high certainty does not.

2.5.3 Calculating for Informed Disclosure Decisions

Proponents of the classical or rational perspective on decision-making hold a view that the human mind is capable of unconsciously conducting complex decision calculations (Lancaster and Lancaster, 1982) . They suppose that probability theory and human reasoning are two sides of the same coin (Gigerenzer, 1999). Edwards et al. (1954) reasoned that decision-making is an optimisation or maximisation process, conducted by a rational or economic man. A prime example is economic utility theory that suggests that, while a person is making a decision, they evaluate the expected utility gained or lost on each decision paths (MacCrimmon and Larsson, 1979). In this manner, the economic approach to disclosure would align with the self-managed model of consent. In that, an individual performs the prerequisite privacy calculus on the disclosures gains and losses, and therefore, proceeds (informed) when there is positive utility (Culnan and Armstrong, 1999).

In the main however, people have been found to lack the mental resources required to perform the 'privacy calculus' of rational disclosure decisions (Acquisti, 2004a; Krasnova et al., 2010). Subsequently, under the rules of economic or rational man, individuals systematically undervalue their personal disclosures (Grossklags et al., 2007; Solove, 2007). Moreover, even if the calculus is attempted, decisions regarding personal information disclosure are not deterministic. Often, beyond the initial exchange of personal data, there exists a latent incentive for secondary use, and or trading, which can result in a real, even if dormant, threat (Beer, 2009; Weitzner et al., 2008).

Simon (1955, p. 104) proposes that there is a "complete lack of evidence that, in actual human choice situations, of any complexity, that these computations [in game-theoretical and economic human decision models] can be, or are in fact, performed." Simon advanced the concept of 'bounded rationality' which balances the computational capabilities of the decision maker, and the decision environment. Like a halfway house, bounded rationality retained the essence of classical reasoning theory as a basis for the prescriptive evaluation of human decisions (Lipshitz et al., 2001). Yet, instead of a consuming optimisation task, decisions are quickly satisficed at some threshold. Concurring with Simon's view of decision-making, Tversky and Kahneman (1992) reason that the real world lacks the transparency within which rational theories thrive, instead the lack of transparency introduces uncertainty. Then, where there is uncertainty, rationality violations or reasoning errors can occur, or be induced by changes in the decision circumstances (Acquisti, 2004a; Acquisti and Grossklags, 2003).

This effectively means that revisiting the consent framework in Table 2.1, leaves an uncertainty of what may happen to the data, and subsequently uncertainty as to the risks involved. Such uncertainty does not lend itself well to a classical or normative view of human reasoning (Brandstätter et al., 2006). Therefore, the calculation between utility gained and utility lost from disclosure is highly asymmetric in favour of the environment controller (the organisation) (Vila et al., 2003). This may then help explain theories in camp one that account for the privacy paradox in terms of individuals acting seemingly in a trade like manner yet with incomplete data, and therefore, with hindsight there is often regret (Balebako et al., 2011). It also boosts the case for emerging research into smart contract data exchanges, such as described in Bünz et al. (2020), that help the individual maintain a level of privacy without losing utility. This form of zero-trust delegated or augmented version of PII exchange show much promise, but many technological hurdles remain, not least usability.

2.5.4 Fast and Frugal Rules for Simple Disclosures Decisions

Rather than starting from a fixed point of rationality and then accommodating for behavioural violations, naturalistic and heuristic decision models start from a descriptive account of decision-making. Such approaches are concerned with what people do, rather than what people ought to do (Gigerenzer and Gaissmaier, 2011; Lipshitz et al., 2001).

Advocates of heuristics believe that rationality violations or reasoning errors are a complete and systematic failure to adhere to rational decision-making, hence heuristic models should be viewed as separate enterprises (Tversky and Kahneman, 1986, p. 275). Instead of trying to calculate a rational choice, or trying to satisfice or approximate utility aspirations, heuristic models involve educated guesses based on experience, an external element of the environment or another person (Gigerenzer et al., 2008). Despite being described as incomplete ideas, heuristics are considered useful or even indispensable for analytically thinking about intractable problems, or where time is limited (Brighton, 2006; Gigerenzer and Gaissmaier, 2011). Nevertheless, heuristic based decisions are by nature, frugal, they ignore part of the information and are therefore inherently uncertain. Moreover, with frugality comes speed and with speed comes additional error (Kahneman, 2011).

In regards to speed, recent work has highlighted the role of impulsivity as a construct to explain privacy decisions, and that high impulsivity corresponds to risky decisions (Coventry et al., 2016). Similarly, Dhamija and Dusseault (2008) recognise that when a user faces a situation such as authentication, it is rarely a primary goal, and the users exhibit haste, and seek the path of least resistance. In this instance, least resistance means least mental effort and time, making heuristic decisions an ideal approach. For instance, when determining the credibility of online retailers, consumers favour quick and easy heuristics that rely on surface level cues over examination of in-depth policies (Sundar, 2008; Van Noort et al., 2008). Such behaviour incentives online services to incorporate intrusive privacy notices that obtain consent hastily yet explicitly, but serve only as a nuisance to the user's actual goal. Moreover, such bounded attention to security details has been linked with successful phishing attacks as users are quickly making judgements about credibility, and trust, and (perhaps understandably) not fully engaging with the process (Dhamija et al., 2006).

Returning to the notion of informed versus simple consent, disclosure under the banner of heuristic decision-making would undoubtedly be considered as simple consent at best. Furthermore, the second camp of theories that account for the privacy paradox are reconciled here, regardless of how abstract risks are determined, actual disclosure is based on fast and frugal trust evaluations (Metzger, 2006; Norberg et al., 2007).

2.5.5 Adapting for Simple and Informed Disclosures

The third camp of privacy paradox theorists tend to focus on the contextual nature of disclosure decisions, whereby, whilst one context may invoke a person's prime consideration, another context may pass with little or no consideration (Foddy and Finighan, 1980; Knijnenburg and Kobsa, 2013a; Van Noort et al., 2008; Wirtz and Lwin, 2009). There is a situation therefore, not unlike with the medical analogy, where informed consent is reserved for the high-risk situations, i.e. severity of consequence, otherwise simple consent suffices. In essence, this is an informal description of a two-system social behaviour model (Strack and Deutsch, 2004) or otherwise referred to as *thinking fast and slow* (Kahneman, 2011).

The concepts is that the two systems work in tandem as people navigate a highly complex social world, with one slow reflective system based on facts and knowledge, and a second fast, impulsive system based on quick associations and motivations. Therefore, an individual may decide to use simple heuristics to navigate decision situations at any given moment, yet at times of high motivation (perceived high risk), this same person may approach a decision more systematically.

Similarly, Higgins (1998) explains individual variations in terms of a self-regulatory focus. This theory distinguishes between promotion focused individuals that are 'eager' for positive outcomes, with prevention focused individuals that are 'vigilant' against negative outcomes (Hall et al., 1997). If the goal is to avoid losses, then a prevention focus is embodied, and a person in this state would seek to self-inform regarding safety and risk (Idson et al., 2000). In contrast, if the goal was some gain, then a promotion focus is embodied and a person may be more concerned with swiftly fulfilling their goal and less concerned with self-informing (Higgins, 1998; Kruglanski et al., 2000). For instance, Craciun (2018) reveal that prevention-focused individuals had less inclination to share data during opt-out and op-out default conditions. Although, this notion of goal orientation is linked to context dependency (Fridman and Higgins, 2017). Moreover, and perhaps significant to future work, an apparent consensus to disclose removes a previously prevention-focused intention to withhold. Johnson et al. (2017) also find that a leader can also invoke self-regulation to followers in a manner akin to the herding dynamics of heuristics. Essentially, the authors indicate, at least within a formal work relationship, that relationships and culture could have as much influence as the context of the goal when it comes to information disclosure.

Other evidence suggests that decision makers focus on the extent of benefit and harm rather than probabilities (Loewenstein et al., 2001). Moreover, that the lure of instant gratification can lead to inconsistent valuations (Acquisti, 2004a). Given the inherent uncertainty regarding the risks of disclosure, it is therefore possible that a promotion focused person actually discounts negative outcomes, whereas even prevention focused individuals may overvalue positive outcomes (Halamish et al., 2008). The manner in which people make decisions also seems to be open to manipulation or 'nudge' (Balebako et al., 2011). A nudge is a method of reinforcement of behaviour and indirect suggestion that do not prohibit bad decisions but instead steer people towards better ones (Hansen and Jespersen, 2013). A nudge could be used to elicit a more thoughtful response (Acquisti et al., 2012), or perhaps, the nudge could enable a prudent yet fast response (Choe et al., 2013). Alternatively, manipulation could be to promote an imprudent, yet fast response by exploiting reasoning errors (Preibusch et al., 2013).

In summary, due to uncertainty and asymmetry in the decision environment, coupled with the calculative limitations of the decision maker or their reliance of simple heuristics, any notion of informed consent seems unrealistic. A user may only evaluate the disclosure on the volunteered data such as email address and name, whilst failing to account for the observed or inferred data, recall Figure 2.1. However, currently the consent model of agreeing the boundaries of information disclosure and subsequent use covers the use of all types of volunteered, observed and inferred data. The implication is that a user inadvertently consents to data being used outside of their expectations (Balebako et al., 2011).

Consequently, users are disclosing personal data under the banner of informed consent, whilst the reality is perhaps more equitable to simple consent, and in the backdrop of this, fraud is rising as privacy declines. However, unlike in medical treatment where the burden of informed consent is on the service, in the context of privacy, the burden of informed consent is on the individual, and even then, much uncertainty remains. If however, the risks of identifier misuse in terms of privacy and or fraudulent use are deemed high, it may explain a growing perspective that an alternative to the current self-informing consent model is urgently required (Solove, 2012; Weitzner et al., 2008).

2.5.6 The Role of Credibility Heuristics in Disclosure

When faced with difficult, uncertain, or intractable problems, using heuristics can be rational (Gigerenzer and Todd, 1999) as they fit with observations of decision-making *in the wild*. Nonetheless, they are prone to misjudgement and bias (Tversky and Kahneman, 1975), and with that are prone to manipulation, or 'nudges'. While these nudges might be used to limit over-disclosure, they can also be used to encourage users to disclose more than they might otherwise be comfortable with (Acquisti, 2012; Hansen and Jespersen, 2013). This section will look at exactly what is meant by heuristics, how they have been linked to credibility, and how this can be used as a framework for exploring disclosure decisions.

Heuristics are used to reduce difficult decisions to solvable simple decisions (Gigerenzer and Gaissmaier, 2011). For example, whether to invest in company A or B is a difficult decision, which depends on many complex factors. Whereas an associated heuristic may be that size is related to success. Then the heuristic's decision variable, i.e., the cue, could be the number of service users or stock price. So whether $A(cue) > B(cue)$, substitutes for whether A is a better investment than B.

Early research around activities such as *phishing* and *fake news* showed the impact that online cues have on users' trust of web sites, and their judgements of the credibility and legitimacy of those sites (Fogg, 2003; Sundar, 2008). Fogg (2003) finds that rather than users seeking a particular cue, prominent cues affect users, and that with new digital interfaces comes changes in the prominent cues (e.g., the padlock shown on a search bar). Understanding how to manipulate these effects is an active research agenda for behavioural economists Balebako et al. (2011); Acquisti (2012); Adjerid et al. (2013).

The MAIN model (Sundar, 2008) structures ten years of psychological research into the cues that have been empirically shown to affect users. This model assembles the cues in terms of four technological affordances of digital media: Modality, Agency, Interactivity, and Navigability. The result is an extensive array of cues and associated heuristics providing a framework for more applied research. Drawing from this model, Metzger et al. (2010) conducted 11 focus groups, and found five prominent credibility heuristics: *Reputation, Endorsement, Consistency, Expectancy Violation and Persuasive Intent*; later adding *Self-confirmation* as a sixth (Metzger and Flanagin, 2013). For instance, determining the credibility of a website is a difficult decision. An associated heuristic may be the Expectancy Violation heuristic that illegitimate websites appear unprofessional. So, whether the website has spelling mistakes can substitute for whether the website is credible.

The value of Metzger et al. (2010) and Metzger and Flanagin (2013) is in their representing an expansive array of cues and decision variables, as is the case in the MAIN model, into a simple and coherent framework (as summarised in Table 2.3). In addition, this table contains the six additional heuristics that Metzger and Flanagin

TABLE 2.3: Heuristic Approaches to Credibility Evaluation Online

Heuristic	Description
Authority	An official or primary authority
Recognition	Familiarity, even in name only
Reputation*	A prestigious service would not knowingly be wrong
Endorsement*	Recommendation from known others
Bandwagon	Recommendations or perceived actions of unknown others
Consistency*	Agreement with another source or procedure
Consensus	Agreement between many sources or procedures
Self-Confirmation*	Alignment with a pre-existing belief
Coolness	New modalities of the technology
Novelty	New encounters with the technology
Expectancy Violation*	Inferior site design, errors, poor visual appearance
Persuasive Intent*	A feeling of bias or being pushed

Note: Twelve (*six prominent) credibility heuristics collated from Metzger and Flanagin (2013).

(2013) discussed as relevant, yet omitted for their purposes[1]. A concern is that the original six may be overly fitted to credibility judgements. Likewise, Gambino et al. (2016) used grounded theory on data from eight focus groups to reveal eight heuristics that they find underpin disclosure; four of which promote disclosure (Gatekeeping, Safety-net, Bubble and Ephemerality), and four that inhibit it (Fuzzy-boundary, Intrusiveness, Uncertainty, Mobility). Where the work of Metzger et al. provided the motivation for this work, the findings in Chapter 4 have been refactored in light of these eight heuristics.

2.6 Summary

This section has detailed how identity assurance is an essential feature of interaction, digital or otherwise. Yet also, that with more and more interactions conducted online, there is an abundance of personal identifying data stored online, and with it new threats. The common response being to increase the overall uniqueness of identifiers extracted during authentication, increasing the 'value' of personal data in circulation and at risk of misuse. This situation looks set to continue as system security evolves from password and token authentication, to physical biometric authentication, and to the use of behavioural biometrics. Also discussed is that this is not solely a technical issue. It is increasingly apparent that social, economic and security dynamics of PII exchange are at play, and are not easily segregated. The challenges surrounding effective identity assurance are an interwoven mix of social, economic and security

[1] Metzger et al., found six heuristics through a process of reduction, i.e., Recognition is subsumed under Reputation, as to perceive reputation involves a prior recognition.

dynamics surrounding identifier exchange. Therefore, the following two chapters are a complementary set of studies aiming to help understand these dynamics towards a more complete model of identity assurance, by addressing the first three research questions. Chapter 3 investigates the personal utility of PII as a plausible counterweight to the more intuitive informative and identifying utilities of PII. Chapter 4 explores the decisions users make to navigate the complexities surrounding the self-managed disclosure of PII. Then, Chapter 5 outlines ways to combine these findings within a model of the identity assurance commons. Chapter 6 simulates the model with dual-system agents to garner insights the efficacy and impact of a serious of interventions. Finally, Chapter 7 provides a summary of discussion, contributions an limitations, along with a of how these works fit together towards a novel insight into identity assurance systems and to provide pragmatic recommendations to regulation.

Chapter 3

Cohort Disclosure Preferences for Personally Identifying Information

3.1 Introduction

The remits of identity dependent technologies range from securing access to nations to delivering advertisements. These remits involve processing personal identifying information (PII) to reach the desired confidence that a person is as they claim or is whom they are believed to be (Beres et al., 2007). Ideally, these remits balance two requirements. On one side they need to exploit the identifying utility of PII to assure the secure access of different systems, and on the other they must respect the personal interests of users in terms of privacy (Williams, 2008). Unfortunately, efforts to meet these requirements are currently imbalanced. This is because exploiting PII's identifying utility is a clear and long-standing goal for both nation states and commercial organisations due to direct security and financial incentives (Clarke, 1988; Black et al., 2012). Whereas, attempts at respecting the privacy interests of users lack cohesion and maturity (Nissenbaum, 2004; Solove, 2007), and have mostly indirect, hard to measure benefits such as goodwill (Toubiana et al., 2010).

Despite this imbalance, there is growing interest in personal privacy, with organisations recognising the market potential in addressing user concerns in the expanding Internet of Things (IoT) (Böhme et al., 2007; Toubiana et al., 2010; Brunton and Nissenbaum, 2015). Likewise, the recent EU general data protection act (GDPR) has added timely impetus to this agenda. So there remains urgency in steering systems designers, researchers, and policy-makers towards identity-based technologies that can meet both requirements together (Cavoukian, 2011).

This work, the majority of which is published in (Marmion et al., 2019), contributes to understanding the personal sensitivities of users when it comes to disclosing their PII.

A focus is on the discord between individual action and intentions as discussed in Chapter 2. Perhaps counter-intuitively, the method here to focus on individual values is done this using a cohort approach that aggregates the partial preferences of individuals. A cohort approach makes this method scalable and adaptable in terms of how the preference indication work is divided. Then, concluding, the results return from that basis to show that this method can estimate individual preferences. To do this I explore a set of PII for indication of an underlying structure of personal utility, and compare the results across three contexts. Also, in the method section I present a novel application of Elo's ranking algorithm (1978), and later show that this method not only ranks subjective data types, but that also reveals relative distances between the ranks.

3.1.1 Summary of Aims

If actions are a poor representation of user values, then what remains are user's intentions and ideals. Hence, it is prudent to understand these intentions and ideals now, before the social norms of disclosure realign with what were originally reluctant actions. This work sets out a method to do this through an investigation into the personal value of PII to individuals, not in terms of what people disclose in a given situation, but on what they would prefer to disclose.

Out of the four main research questions set out in Chapter 1, the aims of this chapter contribute to two. The first aim is to develop a method of obtaining users' willingness to disclose, providing a novel insight into PII's personal utility [RQ1][1]. Aim two is to use this method over three contexts; submitting to government requests, transacting with commercial organisations, or self-expression on social media [RQ3] [2]. Then, the results in the form of a quantitative measure of PII's personal utility will enrich the final model with a nuanced set of PII, rather then an abstracted homogeneous set, towards RQ4[3]. Ultimately, these aims contribute to a better understanding of what drives the escalation of PII extraction, and how it may be tempered.

3.2 Method

The aim is in knowing how users perceive their PII in terms of personal disclosure preferences. To achieve this, the methodology is based around a participant experiment in the form of a pairwise comparison task. In each round participants are asked to choose between two different types of PII within a particular context. Once

[1] RQ1: How can we quantify the personal, identifying, and informative utilities of PII?

[2] RQ3: What are the contextual affects on PII disclosure?

[3] RQ4: How do RQ 1-3 combine within the identity assurance commons?

all the comparisons have been recorded, a ranking algorithm is used to combine the individual comparisons into an ordered list for the group. The following sections give a detailed description of the experimental design and the process with which data were collected and analysed.

3.2.1 Experimental Design

Pairwise Competition (The Match): The design involves using pairwise comparisons to mirror the process of a competitive match between two players. This feature is harnessed here, with opponents (or the players) being replaced by types of PII, and the win condition being a judgement from a participant. In each match, the participant (the judge) chooses from two items of PII, e.g., work email and personal email, that which they are most willing to disclose in a given situation. The winner is the chosen item, and the other is the loser, the format does not permit a draw. Then, if there exists a preference in the judge's willingness to disclose then, some items will win significantly more than others, and therefore, it is possible to produce a ranking on a set of PII. Then repeating the process with a cohort of judges makes it is possible to produce an aggregated ranking.

TABLE 3.1: Set of 33 Personal Data Identifiers

Knowledge and Token				**Biometric**	
Biographic	*Demographic*		*Knowledge*	*Physical*	*Behavioural*
First Name	First pet	Age	Password	Fingerprint	Signature
Surname	Postcode	Education	PIN code	DNA	Geo-location
DoB	First School	Nationality	Username	Iris Image	Swipe Dynamics
Work Email	Phone Number	Home Address	Favourite Colour	Face Image	Keystroke Dynamics
Personal Email	Ethnicity		Account Number	Hand Pattern	Voice Signal
Mother's Maiden Name	Sex			Ear Pattern	

Adopted from a Rethinking personal data report (Rose and Kalapesi, 2012), and adapted through colleague discussion. DNA and favourite colour provide sense test bounds to the set. One would expect these to be outliers in the results. The categorisation of these identifiers is for illustration only; the participants did not receive this information.

Set of PII (The Players): The algorithm requires players to compete. Here the 'players' are a set of 33 items of PII (Table 3.1). This set is a subset of those included in a report on personal data (Rose and Kalapesi, 2012). Not meant as an exhaustive set, rather a modest group of familiar identifiers that span everyday user interaction and span different abstract categories of PII; knowledge, token and biometric identifiers. The set also has a range of sub description diversity, biographic, physical, etc., that was sense testing the set's balance rather than to cue the participant. Participants did not receive this delimitation. In terms of familiarity, an informal run-through of the task with colleagues and verbal feedback surfaced with no issues with the set. Future work would benefit from extending this list. For now, this work focused on the method and what the results can tell us about the characteristics of PII's personal utility.

Experiment Scope: Assuming that the order of presentation is not important, i.e., Postcode vs Fingerprint = Fingerprint vs Postcode, the 33 items of PII produce 538 possible combinations. During an informal pilot, individuals made 50 pairwise comparisons, taking 5-8 minutes, without any apparent attention fatigue.

Initially, an even split of the 538 pairs was considered as it would involve less participants to cover the list, i.e., a cohort of 11 participants making 49 comparisons would cover the list with each item of PII playing 32 matches. However, changes to the list of PII, or in the participant pool would cause issues of inconsistency. For instance, would a single new participant be required to judge all 33 matches required for a new item? In such a case, this participant's experience would differ from the others, as they would judge each original item of PII once, each against the new item. Equally, if the decision was to add more participants, then would this need to be done in multiples of 11? This simple design is frugal but does not scale.

The decision was made to present the pairs randomly to participants as this simple design feature is frugal and scales consistently. The random presentation means that adding extra items of PII to be judged or adding participants to do the judging does not influence the experience of the participants. The point of distributing the work is so that the individual does not even need to see the entire list to contribute to the list in the same way a chess player need not play an opponent yet have a similar ranking.

Due to the work distribution, the number of judges is less critical than each item of PII having ample opportunity to play. For the algorithm detailed in the next section, a heuristic of fewer than 30 pairings implies a provisional player (Coulom, 2007). A mock simulation of this random process determined that each of the three cohorts of 40 participants each making 50 pairwise comparisons ensure over 75 pairing for each item (see Appendix A). The randomness does, however, impose variation in the pairings. Table 3.2 provides an overview of the actual outcome. For instance, Cohort 1, had fewer than 40 participants after removals meaning one item only played 73 matches, while another played 120.

Contextual Variety: To abide by evidence of the contextual nature of disclosure (Knijnenburg and Kobsa, 2013a), the ecological validity of this enquiry was helped by situating the task across three contexts, these are a social (networking) context to account for activities involving 'expression', a commercial context (banking) to

TABLE 3.2: PII Pair Presentations

	Social	Commercial	Government
Total Pairs Presented	1750	2000	2000
Min Presented Per Item	73	101	99
Max Presented Per Item	120	149	151
Range	47	48	52
SD	9.64	12.04	11.23

Note: This table shows the outcome frequency of the random pair presentations to the users.

account for 'transactional' disclosure, and a government (passport) context that aligns with 'submission' based disclosure (Van Zoonen and Turner, 2014). In a between subjects design, participants were assigned to one of three distinct cohorts that differed in the context of the task (see Table 3.3 for details of the scenarios presented to participants). Consequently, three separate personal identifier ranks will be generated, one for each cohort, representing the relative value of PII within that particular scenario of disclosure.

3.2.2 The Algorithm

Elo's (1978) ranking algorithm is the base mechanism to achieve the aim of quantifying willingness to disclose PII. The competition aspect of Elo's algorithm is mirrored in the process of pairwise comparisons (Sarma et al., 2010), and can be used more generally outside of sports to rank subjective-based entities, i.e., photographs (Coulom, 2007). In our instance, these entities, or 'players', are items of PII.

The algorithm produces an expected value for whether player A will win against player B, and uses that value to update both players' scores depending on the actual outcome. For example, players A and B are opponents, each with an associated pre-game score, A_i, B_i. If A wins, then $A_{i+1} = A_i + x$, and $B_{i+1} = B_i - x$, i.e. it is a point exchange of the value of x. This way, incorrectly scored players can rise or fall through the ranks, while correctly scored items remain relatively stable. When repeated with many players and many matches an overall ranking stabilises with the best players on top. This process is achieved with equation (3.1),

$$EA_{i+1} = \frac{1}{1 + 10^{(B_i - A_i/400)}} \tag{3.1a}$$

$$EB_{i+1} = 1 - EA_{i+1} \tag{3.1b}$$

$$A_{i+1} = A_i + x \tag{3.1c}$$

$$x = K(A_i - EA_{i+1}) \tag{3.1d}$$

TABLE 3.3: Task Group Conditions and Tasks

Cohort	**S. Social**	**C. Commercial**	**G. Government**
Context	New Social Network	Personal Bank	Passport Office
Scenario	Your best friends have been telling you about a new social networking site that is very popular and that you should join. You are keen and are determined to join. So you visit the website and begin the registration. However, this site is trying something different with their security.	Your bank calls to say they are upgrading their security methods, however, they are trying something different. You have been with the bank a long time and wish to comply.	You have had your passport for 10 years and it now needs renewing. You need a passport so you begin the registration process. However, the Government, are trying something different with their security.
Task	We take security very seriously and want to improve how we secure our service for your use. Instead of a single identifier, we intend to use a broader set of identifying information. However, we want to give you some control. Therefore, below is a pair of identifiers, what we want is for you to click on the identifier that is most appropriate for the security of this service. Think about this in terms of security and what you are comfortable giving to this service.		

Note: This is a breakdown of the conditions and the instructions given to the participants for the ranking task.

where, x is the number of points exchanged between A and B. The value of x is calculated using the relative magnitudes of A_i, B_i, to determine the expectancy of player A winning EA_{i+1}. Players with higher scores are expected to win, and an expected win yields a smaller x than an upset victory. Referred to as the K-factor, K is a constant that controls the maximum magnitude of x.

Algorithm Parameters: The parameters were set in line with those in Hvattum and Arntzen (2010). Each item of PII was allocated an initial score of 1500. Modifier values of 10 and 400, meaning that if $A_{(i)} - B_{(i)} = 400$, then the expected score of A winning is ten times that of B. These parameters have little impact other than to set the scale for the scores that underpin the eventual ranking. However, the K-factor has a more nuanced role. Typically this number ranges from 10 to 40 in general applications, with a higher number resulting in more sensitivity in score exchanges and also therefore, in rank positions (Coulom, 2007). This can be useful for situations where new players are entering stable ranks, for instance the K-factor can be tuned, i.e., higher for novices, this way a new player may affect the ranking more initially while they close in on a true rank, then after a certain number of games, they are assigned the same K-factor as others. Likewise, a lower K-factor for 'master' players means that they do not effect the scoring as much when playing lessor opponents. In this application, there is no concept of novice and master, so while choosing a constant 32 for the K-factor, may be considered as high, it is explained in the next section why this is a manageable issue.

Systematic Error: An unexpected feature of Elo's algorithm is that the match order along with an overly sensitive K-factor, can affect the final ranking. For example, using arbitrary results between five fictional players, and randomising the order in which the results are processed produces different scores, then from these scores differing ranks can form (Figure 3.1). Eliminating this systematic error involves using a combination of Monte Carlo simulation and rolling scores. Essentially, at T_0 all scores are equal (1500), then randomising the order of the participants and producing a score and rank at $T_{i=1}$. Then repeating the randomisation of participant order, and using the resulting scores from T_i as the base score for T_{i+1}, each item approaches its correct position (seen later in Figure 3.2). Each turn aims to produce less fluctuation until order sensitivity becomes insignificant. Adopting this technique also eliminates the potential hazard of setting the K-factor too high.

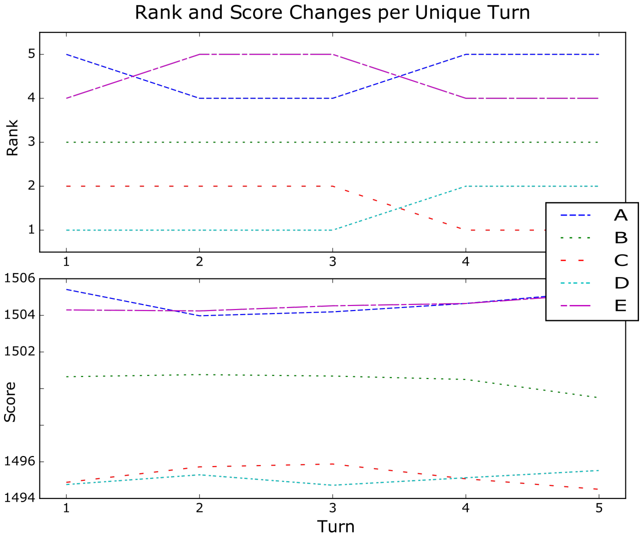


FIGURE 3.1: Using the ELO ranking algorithm on a random list of results between five mock players [A:E], produces slightly different rank [top] and scores [bottom] depending on the order that the results of passed through the algorithm. Each turn is independent.

Advantages of ELO: Using Elo's algorithm is advantageous due to the way the scoring occurs. For instance, in Condorcet voting, e.g., Copeland's method, it is the total votes that indicate rank position (Saari, 1999). Likewise, Likert type systems impose a rating scale on the individual. In Elo's method no scale is imposed and the number of 'votes' is not essential. Instead, it is who or what was in opposition that matters, the magnitude of a victory is primary. These magnitudes not only produce the ranking they also provide a relative value, meaning that it is possible to also interpret any score separations as more meaningful scale.

Comparing different pairwise ranking algorithms, (Goodspeed, 2017) shows that the Elo rating method provides a more discriminating outcome for subjective evaluations that elicit extreme responses than the simplistic win ratio and the Q-score that depends on the win ratio. This attribute of the Elo rating system suits the topic of privacy and disclosure decisions as the context and data sought can be driven by state and trait emotions that entangle hopes with risk environments (Fridman and Higgins, 2017). This method would be a worthwhile addition to other methodologies that aim to classify and order subjective data, especially those relying on human moderators rating all elements, such as examining the 'anonymity sensitivity' of social media posts (Correa et al., 2015). With our method, their participants would be able to cover a more massive repository of data without increasing workload.

Additionally, in using this ELO approach, not all items have to be paired an equal number of times or paired at all. Likewise, new items added to the original set will eventually find their place without the need to repeat the entire process. As PII is a continuously growing set (Black et al., 2012), the feature to add new items and or participant is a significant advantage. Later, discussed is the need for togetherness due to the interdependencies of PII disclosure, and it is from this perspective that one may perceive a snowball of engagement, with new information continuously added, and new participants joining the process, all without disrupting the core methodology. That sort of data set would be a powerful negotiating tool against a system that holds all the cards. An added advantage is that, given the differences in how people make disclosure decisions (Van Zoonen and Turner, 2014; Knijnenburg, 2013), the process can be repeated with diverse cohorts and aggregated over time, and or kept separate. This flexibility means that groups that share the same ideals and motivation could form forcible coalitions.

3.2.3 Procedure

Participants responded to an online advertisement. They followed a hyperlink to an online survey system and completed the required consent to continue. During the consent process, the participants were briefed on the procedure.

To provide additional insight into the psychological composition of the participants, prior to undertaking the pairwise comparison exercise, participants responded to three questionnaires. Regulatory Focus (11 questions) (Higgins, 1998), a two factor Regulatory Mode (30) (Kruglanski et al., 2000), and the four-factor UPPS impulsivity test (urgency, sensation seeking, lack of premeditation, and lack of perseverance) (45)

(Coventry et al., 2016; Whiteside et al., 2005). Questions were presented in a fixed random order.

The results of these questionnaires provide additional confidence that the participants were indeed behaving normally, not least as the distributions themselves represent normal distributions but also in terms of consistency with past studies (see Appendix **??**). The regulatory mode questionnaire included a 6-item lie index to assess the credibility of responses. This data will also be used in constructing the agents in the final model.

The three questionnaires took the participants approximately 15 minutes to complete. On completion participants were redirected to a bespoke webpage for the pairwise comparison task. The site presented one scenario to each participant, then the first of the 50 identifier pairs in the form of two hyperlinks, the link that was clicked counted as the winner, and the next pairing was displayed. The system randomly allocated the situational context from the three described in Table 3.3. Cohort (S) were invited to consider disclosure of information in order to sign up to a new social media site, cohort (C) to consider disclosure of information to their bank, and cohort (G) to consider their disclosure of information to a governmental site to support a passport renewal.

3.2.4 Participants

Differences could be introduced by variety in participant demographics, as described in (Knijnenburg, 2013). This work focuses on the influence of context, yet for alternative questions, the context could instead be factors such as age, education or gender. It would just be a case of controlling the recruitment for these factors. However, these differences can become overly granular whereas here PII and identity assurance are considered as a collective matter (Fairfield and Engel, 2015), so constraints may be of diminishing value. Participants (the judges) self-selected from within three UK Universities, based on an advert seeking those aged 18 to 26, of UK and Irish nationality. Table 3.4 summarises the recruitment and selection. Recruitment stopped when each of the three cohorts reached 40 participants, in total 125 participants took part (93f / 32m, mean age 20.21, SD 1.78). A part of three questionnaires (a two factor Regulatory Mode (Kruglanski et al., 2000)) is a 6-item lie index to assess the credibility of engagement. As a precaution, 10 participants were excluded from the results due to scoring over 1.5 standard deviations from the mean lie scale, or for leaving more than 10 answers blank over the three questionnaires.

TABLE 3.4: Participant Recruitment Overview

Cohort	Context	Male	Female	Mean Age	Age	N	N*
S	A Social Network	7	33	19.7	18 - 26	40	**35**
C	A Bank	10	32	20.5	18 - 25	42	**40**
G	The Passport Office	15	28	20.4	18 - 25	43	**40**

N* is the participant numbers after eliminations

3.3 Results

The results comprise three sections. The first section examines the results solely from Cohort S; the PII requests from a social network service. However, these results typify each cohort. The second section of results will incorporate Cohorts C & G to provide a three-way concordance analysis. The third then shows how it is possible to use these cohort preferences to bootstrap the efforts of the individual.

3.3.1 Ranking and Clusters

Within Cohort S, there were a total of 1750 matches from 35 participant judges each making 50 pairwise comparisons. Figure 3.2 [top] shows the ranking of the 33 items of PII, while Figure 3.2 [bottom] are the corresponding scores. Initial observations show that predictably (and consistent for each cohort), DNA sample and favourite colour occupy opposite ends of the scale in terms of willingness to disclose. Note that in practice these participants are potentially willing to disclose all of these items, this is a relative willingness. Higher scores indicate that an item was selected mostly as a winner, i.e., more willing to disclose. This expected coherence somewhat validates that participants cooperated by not randomly clicking through the pairwise comparisons.

As seen in Figure 3.1, Figure 3.2 [top] illustrates how ranking can fluctuate depending on the order by which the matches are processed. The closeness in the scores in Figure 3.2 [bottom] are what are causing these rank fluctuations. Lowering the K-Factor would produce less fluctuations, however using Monte Carlo simulations with scores at T_i used for T_{i+1}, these initial fluctuation dissipate, and although fluctuations persist, by T_{20} most positions have stabilised.

What is already apparent by T_{20} is that the scores have separated into clusters. Taking the results at T_{1000}, Figure 3.3 is an alternative view of these clusters as score box plots. Ordered by final ranking, this figure illustrates how these clusters form what appears to be step-levels of PII, with items such as age and sex in a cluster that is comprised of PII relatively more inclined to be disclosed by participants, whilst face and iris image are contained within another cluster of PII that the participants were collectively less willing to disclose. Due to the variation and overlap within these scores, it is prudent

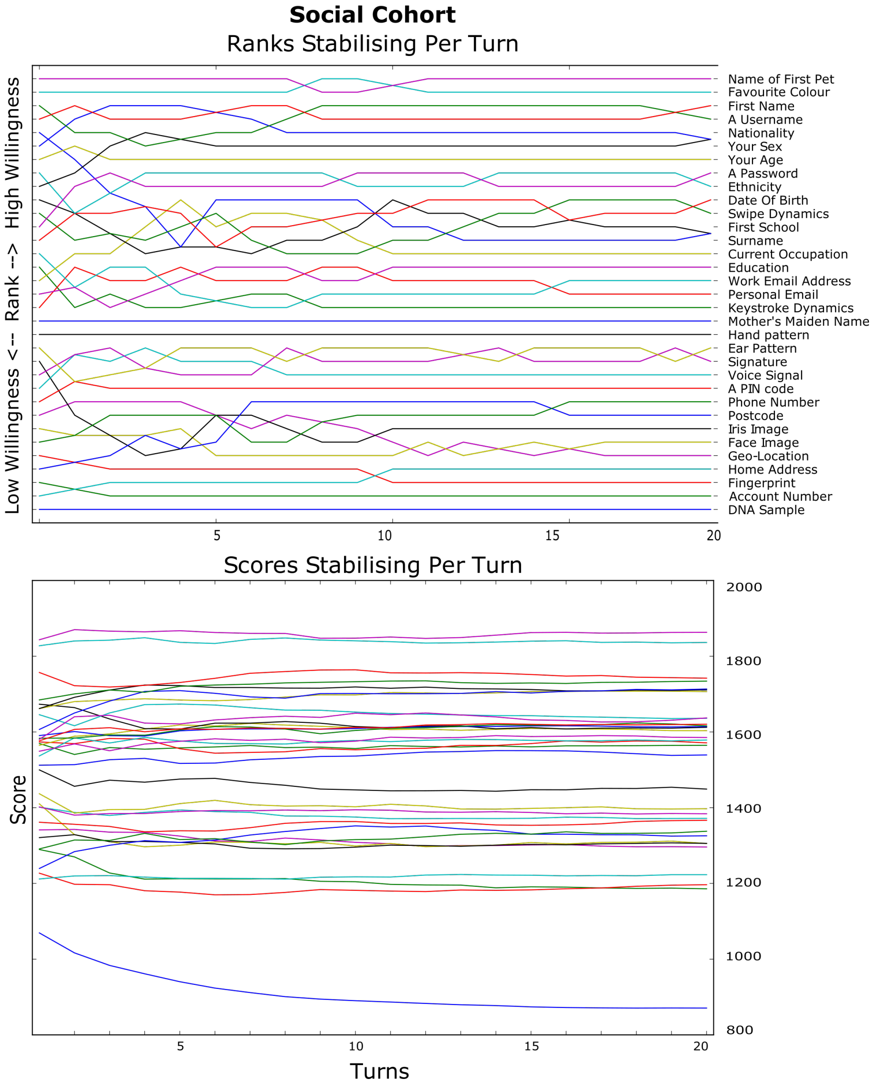


FIGURE 3.2: **Rank and Score Stabilising** [Top] The rank positions fluctuate at the beginning but visibly settle by T_{20}. [Bottom] The corresponding scores indicate a clustering of items of PII that are sensitive to small fluctuations. K-Factor 32 was used, a lower factor would produce less fluctuation. Scores at T_i are a base for T_{i+1}.

to consider rank and score together as rank alone does not tell the whole picture. For example, Ethnicity, Nationality, and Sex, rank 26, 27, and 28th respectively, yet the corresponding score are 1648, 1704 (+56), and 1716 (+12).

The next question then is how many clusters are there, and using R's K-means function as described in Murtagh and Legendre (2014) it is possible to determine four to five clusters from the scree plot in Figure 3.4 [left]. The corresponding dendrogram [right] shows the list divided into four, the length of the horizontal lines is a representation of the relative distance between the scores.

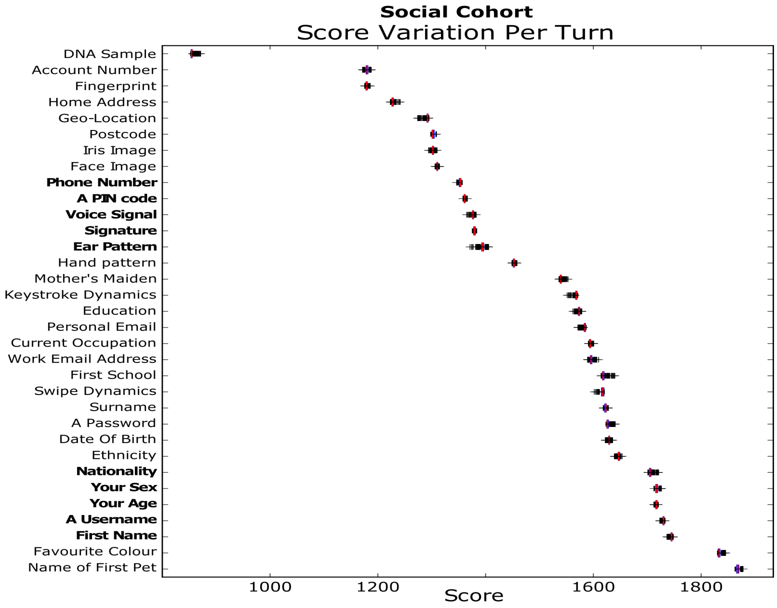


FIGURE 3.3: **Score Boxplots**. Although individual scores changed over T_{1000}, the spread of scores was narrow.

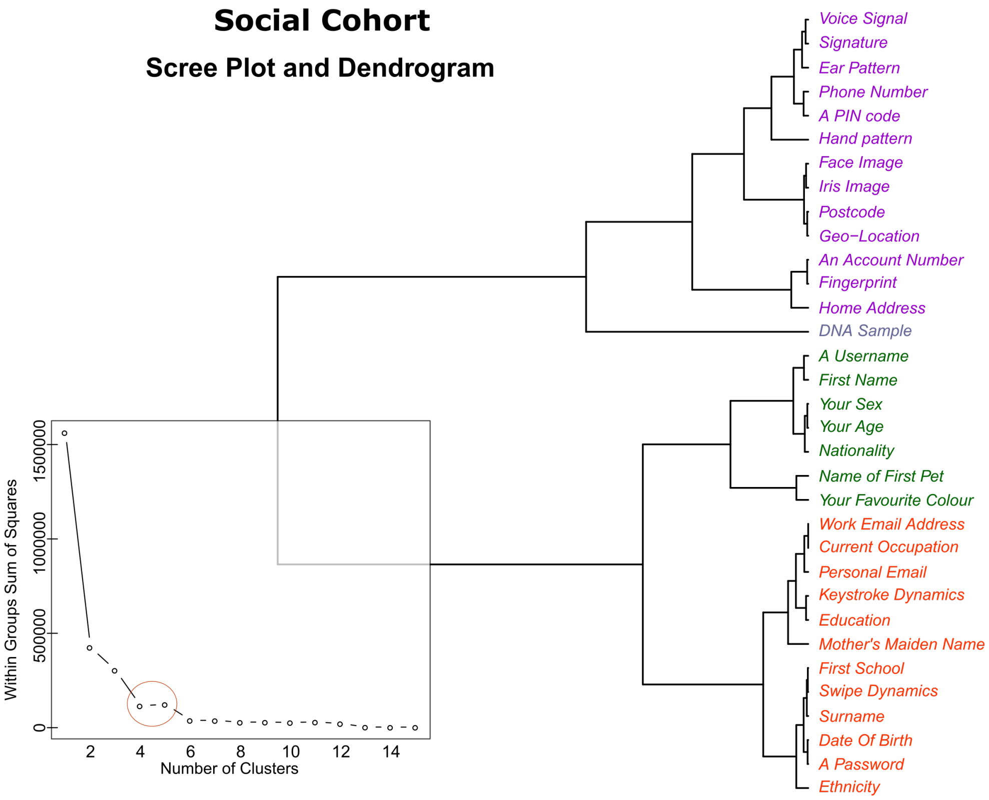


FIGURE 3.4: **Clusters**. The scree plot (bottom left) and dendrogram (right) for the social condition. Here 4-5 groups can be visually deduced from the scree type plot, at 6 there is no added benefit. The dendrogram signifies groupings due to mean distances, not in rank order. Technique: Hclust in Rs stats package using ward.D2 (Murtagh and Legendre, 2014)

3.3.2 Contextual Concordance

Table 3.5 contains the final ranks of each the cohorts after T_{1000}. The table has been sorted based on the average rank of an item. Also included is a σ score to indicate which of the items were either volatile or stable within the rankings, with a lower number indicating greater stability across contexts, i.e., phone number $\sigma = 9.64$, home address $\sigma = 1.52$.

TABLE 3.5: Ranks and Volatility

PII Label	S	C	G	σ
Username	30	33	29	2.08
Nationality	27	27	33	3.46
Surname	24	30	32	4.16
Favourite Colour	32	31	21	6.08
Password	23	32	27	4.50
First Name	31	29	22	4.72
Name of First Pet	33	25	24	4.93
Age	29	18	28	6.08
Sex	28	15	31	8.50
Mum's Maiden Name	15	28	30	8.14
Date of Birth	25	20	26	3.21
First School	21	22	23	1.00
Current Occupation	19	19	25	3.46
Work Email Address	20	26	13	6.50
Personal Email	18	21	18	1.73
Education	17	23	17	3.46
Swipe Dynamics	22	12	20	5.29
Ethnicity	26	11	16	7.63
Postcode	6	16	19	6.80
Signature	12	13	14	1.00
Phone Number	9	24	6	9.64
PIN code	10	17	11	3.78
Ear Pattern	13	6	15	4.72
Keystroke Dynamics	16	5	7	5.85
Voice Signal	11	8	8	1.73
Hand pattern	14	3	10	5.56
Fingerprint	3	9	12	4.58
Face Image	8	4	9	2.64
Account Number	2	14	4	6.42
Geo-Location	5	10	2	4.04
Home Address	4	7	5	1.52
Iris Image	7	2	3	2.64
DNA Sample	1	1	1	0.00

Ranks and Volatility. Sorted by the average rank across the three contexts (S)ocial, (C)ommercial, and (G)overnment. The standard deviation (σ) provides a reference to how volatile an item was across the cohorts. T_{1000}

Figure 3.5 depicts each of the three scenarios reaching stability. This stability resulted from comparing the ranking in Turn T_i to ranking in T_{i-1} using Kendalls Tau, as described by Legendre (2005). This figure illustrates a reduction of the systematic error within the algorithm by T_{300}, yet some persist beyond T_{700}. With some items the score grouping was such that some flux could not be entirely eradicated using the initial parameters and data set, e.g., in cohort S, sex (1716.15) $\approx$ age (1716.40) making the ranking susceptible to small movements in score even by T_{2000}.

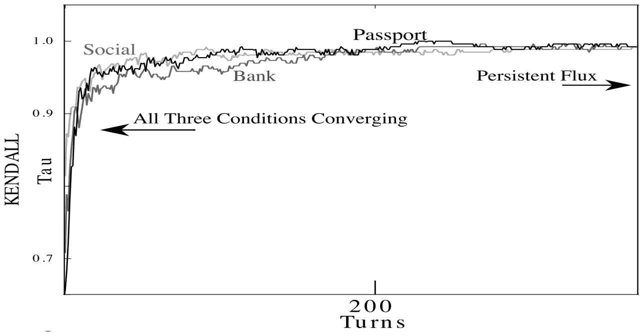


FIGURE 3.5: **Stabilisation**. Kendalls (τ) Rank Coefficient (Legendre, 2005), shows rank stabilisation up to T_{800}. The Bank Condition was the least stable. Also, a small flux persisted.

To investigate the affect context has on willingness to disclose, the final ranking from cohorts S, C and G, although comprised of 35+ judges each, are here considered as ranking from three independent judges. So essentially, the context becomes the judge, and one can compare the judges responses. Following Legendre (2005), Kendall's coefficient of concordance (W) measures the agreement between > 2 judges. The result $(W = 0.81, p < .0001)$ suggests an overarching similarity in rankings across the three contexts. Yet it is clear from items such as phone number or ethnicity that there is sufficient discord between the contexts. When examining the conditions in pairs, Table 3.6 indicates that this seemingly high value of W is due to the close similarity of two of the conditions; the passport and social cohorts $(\tau) = 0.62$).

TABLE 3.6: Rank Concordance Between Contexts

Kendalls Tau (τ)		
	Social	Bank
Bank	0.44	-
Passport	0.62	0.52

To this point the contextual differences have been the focus, however, the similarities between the contexts that resulted in $W = 0.81$, and illustrated in Figure 3.7, illustrates that distinct clusters exist across the contexts. Excluding the main outlier (DNA) the K-means clustering arrives at five clusters of PII that are separated mainly in terms of willingness to disclose (PC2). Moreover, Figure 3.6 illustrates the underlying consistencies between the groups, yet suggests that a) the bank cohort is less like the social and passport cohorts, and b) the social an passport cohorts differ from the bank group in similar ways. Finally, the ladder plot, Figure 3.8 serves as an overview the results so that it is possible to follow an item of PII across context to see differences, but also to see that similarities persist.

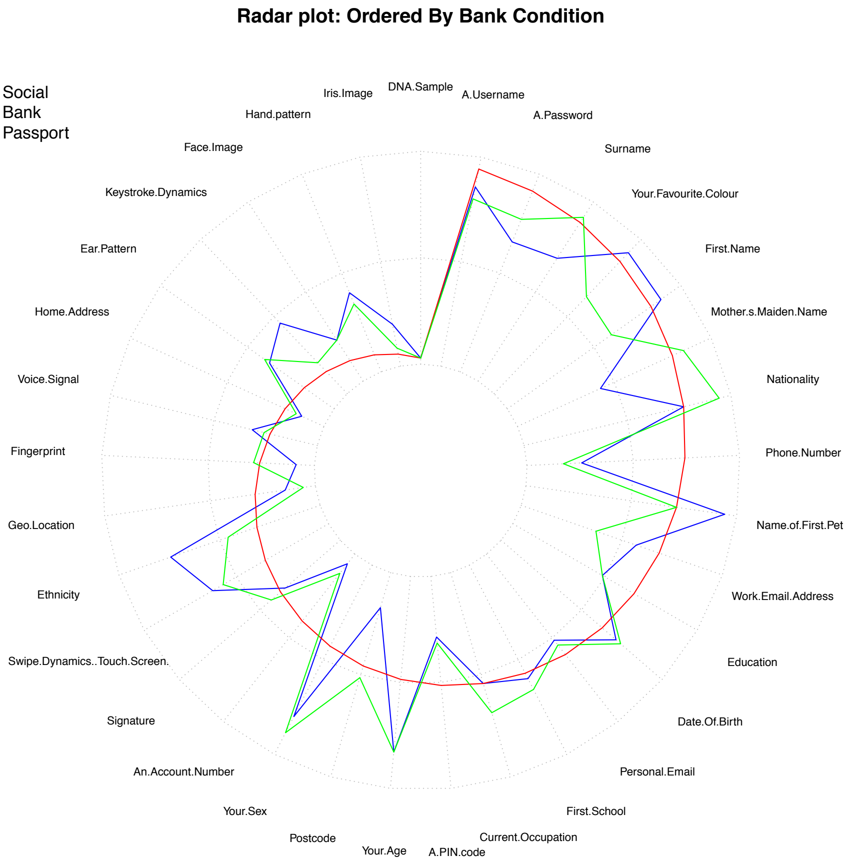


FIGURE 3.6: A radar plot ordered by the rank results in the bank cohort. This illustrates the underlying consistencies between the groups, yet suggests that a) the bank cohort is less like the social and passport cohorts, and b) the social an passport cohorts differ from the bank group in similar ways.

3.3.3 The Willingness of Crowds

It would take 538 pairings for an individual to rank 33 items using pairwise comparisons of each combination. However, in Table 3.7 it is shown that the partial sets of 50 individual pairwise comparisons, aggregated into cohort-based rankings, is a reasonable estimate of the individual's preferences.

Table 3.7 forms from comparing each individuals opening matches, i.e., fingerprint vs retina, to the final $N-1$ rankings. The percentages indicate that a new individual using the cohort ranking will have a good (77%, 67%, 64%) chance that is in the correct order, rising (90%, 87%, 88%) when considering close, i.e., (< 5) yet incorrectly ordered items. Then, with this bootstrapping, it would is possible to keep 'turning' the

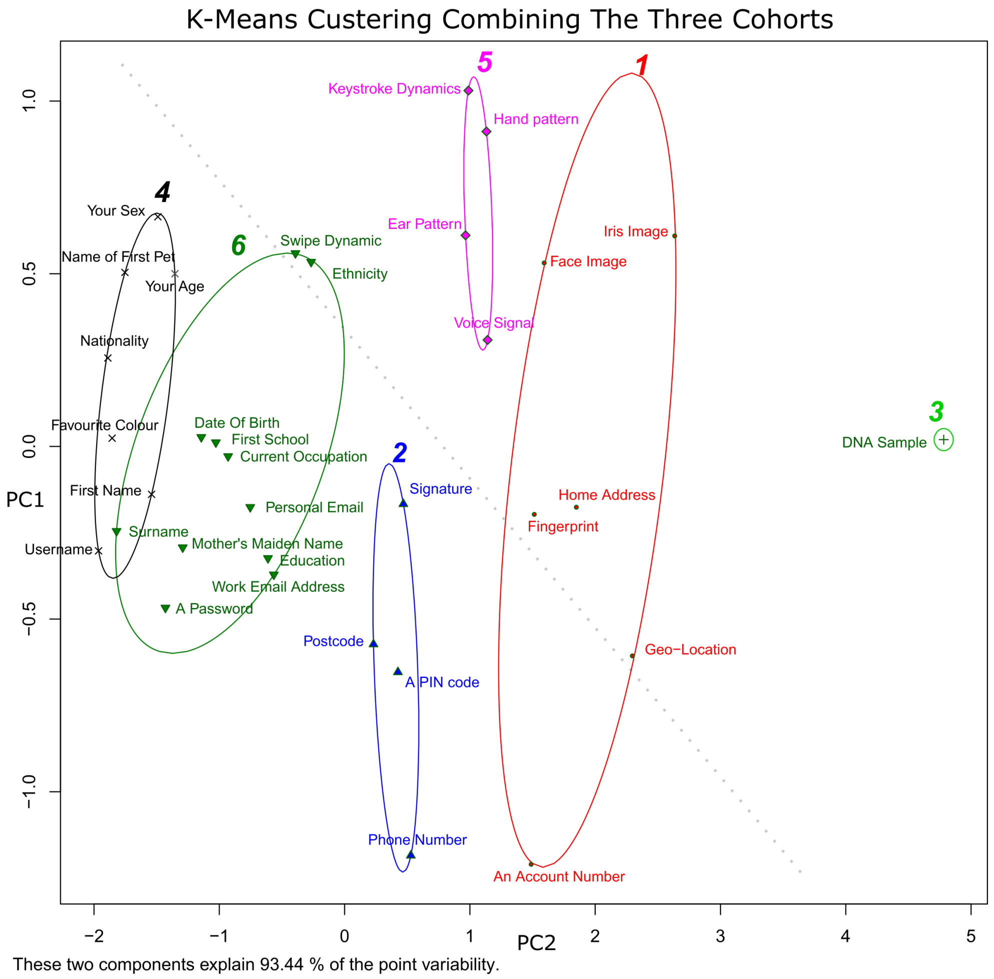


FIGURE 3.7: Cluster Plot. This plot uses the K-means clustering (max.iteration = 1000) function in the R program. The input was the scores of all three cohorts (T_{1000}). It illustrates that even with differences across context, distinct groups exist within the data. Principle Component 1 (PC1) closely associates with volatility between the cohorts, e.g., personal email ranked consistently and if found along the centre. PC2 associates with the general willingness to disclose. The dotted grey line was added after to illustrate a separation of token and knowledge types and biometric types

TABLE 3.7: Wisdom of Crowds: N-1 Individual Estimates

	S	**C**	**G**
Correct order, > 5	59%	52%	49%
Correct order, ≤ 5	18%	15%	15%
Incorrect order, < 5	13%	20%	24%
Incorrect order, > 5	10%	13%	12%

Percentage of correct and incorrect ranking (within and outside of 5 ranking places) when compared to the base pairwise comparison matches.

system using only the individual's 50 preferences or by providing the occasional manual override to bring raise these percentages further to rectify the list towards their preferences.

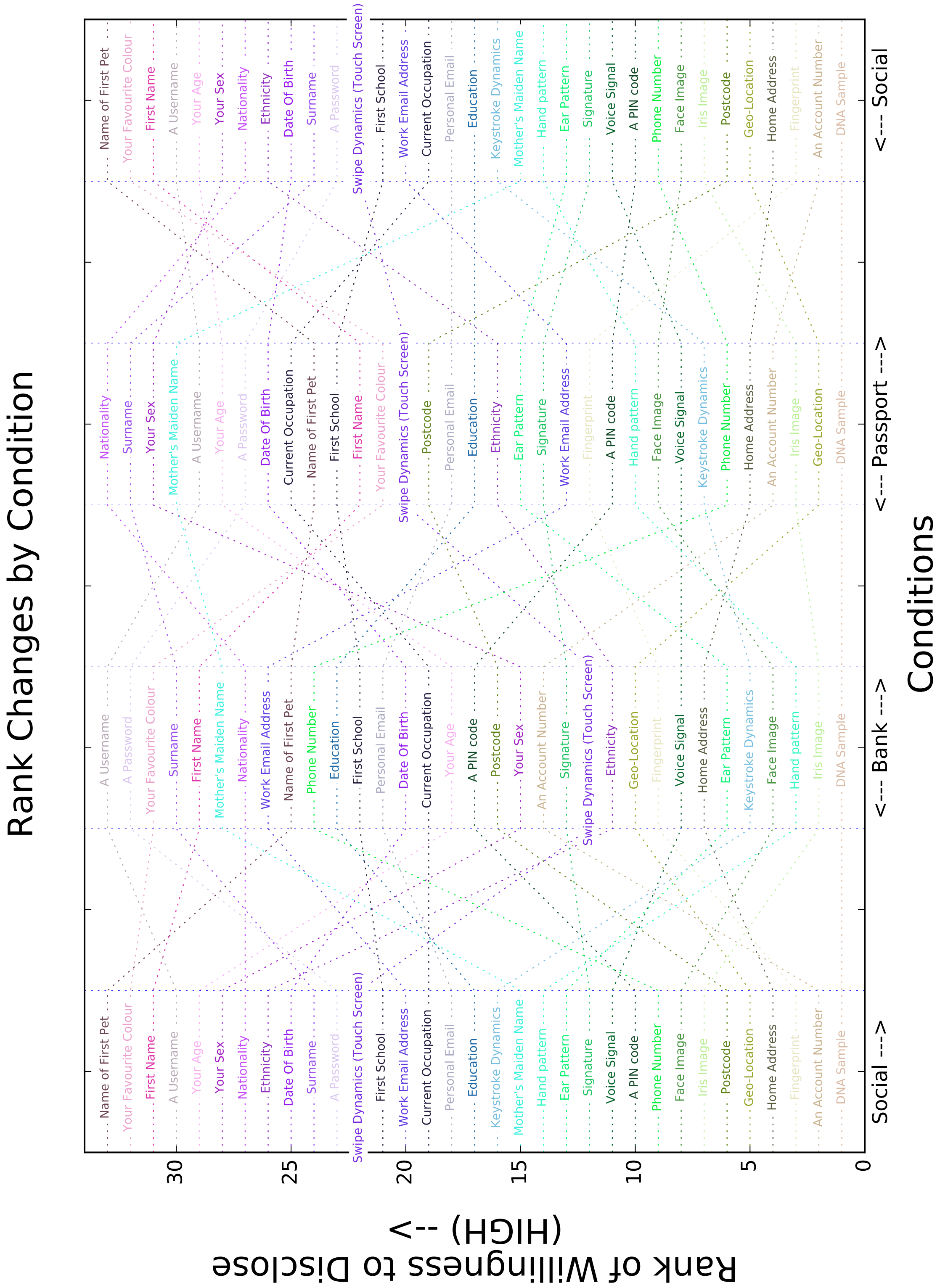


FIGURE 3.8: Overview of the rank results in cohorts 1, 2, and 3 Note: Column One duplicates in column four for easier cross-reference.

3.4 Discussion

Overall, the results suggest this to be a useful and economical method to work with subjective data. Using a modest number of participants for under 10 minutes each, it was possible to distinguish patterns within a list of 33 items of PII. The method is scalable in that increasing the list of PII requires more individuals, rather than increasing the load on an individual. In addition, the method is adaptable, as adding more individuals and/or more items can be done without discarding any original data.

It was unsurprising that items such as DNA faired differently from items such as email address, and therefore, the aim of producing a some order of ranking would be possible. The manner by which the items of PII seem to differ suggest that the aim of ranking a set was fulfilled. The results also show how the list forms clusters of similarly perceived items of PII. This means that along with a ranking, there is a relative distance between certain groups of PII that could be meaningful to identity-based systems that wish to collect PII yet also wish to avoid overreach in their PII requests, at least from the perception of users.

Repeating the process, yielded similar yet distinct results across the contexts, suggesting that this method is sensitive enough to reveal subtle inter-contextual differences. The results of the three contexts were mixed, in some instances context made little difference, e.g., personal email ranking at S:18, C:21, and G: 18) in others, for instance, phone number ranking at S:9, C:23, and G:6, suggests that context is significant for a sub set of PII. This echoes the suggestion that the online world should not be considered as one undifferentiated space for identity matters (Emanuel et al., 2014).

This ranking of PII from a personal utility perspective could provide system designers new insight into the data they may consider requesting. A designer of a social site may wish to have users' phone numbers, but not understand that there are potentially over 20 items of PII that users may find more agreeable. Perhaps, said designer could seek a combination of these lesser items, forgoing the phone number. In such a circumstance the user and the extractor may prosper, as combinatorial identification has proved to be powerful (Black et al., 2012), also the user does not face the dilemma of disclosing a highly valued item of PII or retracting engagement. This later dilemma has further importance because of the knowledge gap that exists, as designers and PII extractors may never know why lack of engagement has occurred, or whether users resent the disclosure.

Understanding such nuances could be advantageous for the efficacy of personal disclosure recommender systems (Balebako et al., 2011; Knijnenburg and Kobsa, 2013a) or automated disclosure negotiation agents (Baarslag et al., 2017). For instance, instead of a user training a system from scratch, all users can share the workload to arrive at a cohort ranking that has relies on the 'wisdom of crowds', or more apt, the willingness of crowds. Then, using the results, i.g., those in Table 3.5, to pre-populate a disclosure aid system, it would be relatively simple for a user to tune the results to their preference by conducting the occasional sense check on the ranking via a simple pairwise comparison. Likewise, new entries can reach their position by conducting pairwise comparisons in a simple binary search pattern. Future users may then avoid the quagmire of disclosure decisions required in the IoT and instead delegate the disclosure of PII to autonomous negotiating agents (Baarslag et al., 2017).

In Table 3.1, the set of PII is nominally separated into broad categories: knowledge and token, and biometric identifiers. The results here suggest that this conceptual model aligns with the psychological model of PII held by the cohorts. For instance, drawing a faint line in Figure 3.7 indicates a separation of the biometric identifiers (clusters 1 and 5) from most other biographical or demographical PII, and this separation was stable across each context. The implication of this is that it shows a discrepancy in how PII is treated in practice. Biometric PII is considered as being of greater identifying utility than biographical PII (Rathgeb and Uhl, 2011), but from these results that it also has a greater personal utility, yet when it comes to extracting, processing, storing and trading such information both sets are treated the same. Therefore, there is a cost imbalance to PII disclosures, as the benefit for the extractor is increased, but so is the 'cost' to the discloser. One way this might change would be to regulate PII by assigning special status to items of low disclosure willingness. Then, it would be plausible to rebalance disclosures by exerting a cost on the organisation using these 'special' items of PII. Whether it remains at the discretion of organisations to show restraint in their extraction of PII, or becomes a power of regulators to restrain the use of PII on behalf of the user, the results here inform these options in a practical manner.

3.4.1 Limitations and Potential Extensions

The demographics of the sample pose a limitation. For example, it is possible that results that place home address consistently low on the willingness to disclose rank regardless of context reflect that the participants comprise mostly students and there is an element of being displaced, and temporariness to student living arrangements. Unknown in these results are any latent variables within the data, that is, whether users envisaged an actual disclosure method and incorporated a willingness to engage with the usability or novelty of such a system. Equally, items of PII such as swipe dynamics would likely have been interpreted differently by individuals as some may not be even aware of such methods, so some explanations or demonstrations would enhance future implementations of this study, but only if done promptly and succinctly. However, these limitations lessen because social drivers and interdependence not only help promote and develop technologies in the extraction of PII (Fairfield and Engel, 2015), but they also provide social proof for individual and groups to establish disclosure norms (Das et al., 2015). That is because, correlations, inconsistencies, misunderstandings, and even maliciousness are all thrown into the mix with this method, just as they are in the real world. The strength of the process, however, is that outliers are not abandoned instead consensus merely counterweights them. Another point arising from the limitation posed by narrow demographics is the extent an individual may wish to exploit the aggregated cohort data from a self-reflecting homogeneous sample instead of seeking consensus from a diverse group. Either way, the final result can be trained to the individual, but how much training would be required could be a factor of such a decision, or perhaps the decision would be to take the consensus as is, thus delegating to a particular crowd. These questions require a qualitative investigation that falls beyond the scope of this work.

The Elo ranking algorithm, despite still used today, is foundational to other algorithms (e.g., Elo-Davidson) that exist to deal with imperfections in the Elo model, e.g., order of judgement or new items entering the data set or even veteran vs novice players. The PII set used, adopted from (Rose and Kalapesi, 2012), has novice and veteran items in some manner if we think of the terms by way of familiarity to the participant, i.e., biometric data remains novel. Future work could incorporate this notion of time to match the maturity of the technology.

Ranking models, pair-wise or otherwise, each have parameters that tune the outcome of the systems. When ranking subjective matters, the validation of the output often relies on coherence and other subjective means. In the case here, this limitation is valid. Altering the parameters can change the ranking. Nonetheless, the results becoming the parameters that inform the simulation in Chapter 6 are determined by the clustering Figure 3.7, not by the individual rank results. This compromise smoothes some of the noise and lessens any broad statement regarding one item of PII to another.

Two aspects of some pairwise comparison algorithms are absent from the ELO rating system used here. First, the lack of a draw condition can introduce error into the ranking by forcing participants to choose. Such behaviour may be uncomfortable for participants. Beyond an informal pilot, the design did not record any potential feelings towards the method from within the task. It could be of inconsequential. Nonetheless, (Szczecinski and Djebbi, 2019) propose a hybrid algorithm, k-Elo, to tackle the draw state. Further research design could warrant the inclusion of the draw condition and, therefore, the use of k-Elo that accounts for the frequency of draws, thus removing any doubt to the effect of the absent draw condition. Second, pairwise analysis methods exist that provide adaptive pair selection (Jamieson and Nowak, 2011) rather than equal distribution, or random as in the current design. Not only does adaption considerably lessen the number of pairs presentations required. This technique offers a capability for individual use learning that could empower any agent assistant attempts in the future.

Also outside the scope of this work, there are a few natural progressions to this research. The first is to consider expanding the list of PII to include other personal information such as credit score, sexuality, health records, and religious beliefs. While these are often disclosed or extracted through modern living already and can contribute to profiling methods (Kosinski et al., 2013), they are not included within this set of more formal identifiers. However, in the context of tomorrow, they may become valuable identifiers swept up along with metadata to provide additional 'behavioural signals' towards passive identification (Perez et al., 2018). Second, this method could track preferences over time, as new extraction methods become mainstream; for example, it would be fascinating to see how society's reluctance erodes, or becomes entrenched, regarding biometric identification. Third, to use this method across cultures and age groups, as the ideal and norms of the UK may differ vastly internationally.

3.5 Chapter Conclusion

As users repeatedly face undesirable disclosure decisions, a lack of strategy, regarding robustness and social acceptability, hampers the long-term effectiveness of identity-dependent technologies. This work premises that actual behaviour is often a watered down remnant of previous preferences, arriving at this view due to the sheer amount of disclosure requests faced each day as people try to get from permissions to the task. Moreover, whether these requests are active or passive, our actions consent to accept the uncertainty and risk in PII disclosures. So, if observing the actions of individuals in the wild is unreliable, then what remains are the users' intentions and ideals. It is prudent, therefore, to protect intentions and ideals now, before social norms adapt to what had initially been reluctant actions.

A contribution of this work is in developing a method enabling exploration of value in subjective data, in this case into PII's relative personal value. Using this method across three contexts, submitting to government requests, transacting with commercial organisations, or self-expression on social media has provided a novel insight into the intentions and ideals of users' willingness to disclose. These ideas, along with this new measure of the personal utility of PII will be brought forward to enrich the simulated model of identity assurance (Chapter 7) by providing a form of measurement for user agent's disclosures.

The results reflect a non-linear distribution within a set of 33 items of PII, indicating clusters of similarly valued PII, some that users seem relatively willing to disclose, but suggesting others that need more protection as they represent a high personal value. Also, it has been shown that for many items of PII, their personal utility changes depending on context. Together, these findings indicate an inadequacy within current PII regulations, which treat PII as a homogeneous set, whereas a more nuanced set of regulations would perhaps help avoid any over-use of high value items.

These findings indicate that different types of PII should warrant different levels of protection, where organisations perhaps should have to justify and bear some cost for the use of these higher value PII. Understanding these levels could promote regulation that transfers the value as a cost to those wishing to extract. Such control would be one way of tempering the escalation in PII extraction. Another would be to use autonomous consent-based agents that are less likely to succumb to the normalisation of disclosure or other cognitive bias that affect user disclosures.

The findings also raise questions regarding the transfer of PII across contextual boundaries, as the results suggest that it is possible to manipulate a user in such a way that disclosure sought in settings of high willingness to disclose get transferred to settings naturally of low willingness. Such an action would likely go against the ideals of the original discloser. Moreover, as the utility of the extracted PII increases for those now in possession, the original trade-off value, e.g., gratification, to the individual may not.

It is essential to protect PII for users across all societies, and not just from those at the head of the technological curve. For the reach of modern technologies means that regarding privacy and identity we need to think collectively, of PII as a public good, rather than individualistically (Fairfield and Engel, 2015). For if others are increasingly sharing their PII, however personal their own motivations, then the pressure increases on us to share our own PII. The goalposts move, and increasingly we will be expected to share PII that we consider of high value. A public good indicates a common problem.

Chapter 4

The Cognitive Heuristics Behind Disclosure Decisions

4.1 Introduction

Disclosing personal data is self-managed. It is the user that decides whether to consent to disclosure requests or whether to withhold their data. This consent-based model aligns appealingly with the ideals of information self-determinism (Westin, 2003). However, in practice these ideals are not being met. Although regulations such as the UK Data Protection Act (1998), and GDPR (2016/678) stipulate that organisations inform users of the operation of data processing, the explanation of risk is left to organisational discretion, making it incumbent on users to make the necessary risk calculation (Solove, 2012). Adopting the long-standing definition regarding consent in a medical context (Whitney and Mccullough, 2004), the reality of consent in disclosure is akin to simple rather than informed consent. Simple consent involves a brief explanation of operations, followed by a trust-based agreement or refusal; to elevate this to informed consent, a detailed discussion of risks is also required. Regardless of likelihood, high risk necessitates informed consent, and because the consequences of data misuse are increasingly high and decreasingly rare (Goodman, 2015), simple consent is unsatisfactory.

Unfortunately, increasing the autonomy of informed users may not be sufficient. This is because disclosure decisions are inherently uncertain, and when data is stored indefinitely there is no means of accounting for future uses or future capabilities, so the 'data controller' may also be ignorant of the risk (Weitzner et al., 2008). Even if organisations were able to explain the risk, users rarely read privacy policies (Vila et al., 2003), and when there is an attempt, they lack the time and/or the capacity for the required uncertainty calculus ('privacy calculus') to comprehend them (Furnell and Phippen, 2012; Kehr et al., 2013). Users are essentially left to trust that the data controller will behave in an expected and innocuous manner (Olivero and Lunt, 2004; Metzger, 2006) before making heuristic judgements (i.e. using 'rules of thumb') about whether to disclose (Tversky and Kahneman, 1975).

When thinking of our problem environment like a commons, recall the regulation continuum set out within Chapter 1. That is, there is an over-regulation theory that brings about an inefficient anti-commons tragedy (Heller, 1998), and a free-riding under-regulated, straight-up tragedy of the commons event (Hardin, 1968). Coupled with Chapter 3, wherein an individual's underlying value forms an overlapping hierarchical set of clusters. With biometric and location data holding relative strength over more mundane biographical PII, one would expect that self-regulation in an environment of government regulation would result in a veering towards anti-commons of low disclosure of the higher value PII, and low extraction or request of said PII. Instead, the reality is akin to a low value and high extraction as the digital world advances, taking more with faint resistance from the individual.

Sundar et al. (2013) and Gambino et al. (2016) consider cognitive heuristics as a key decision feature that aids our understanding of disclosure decisions. Moreover they find heuristics as key to illuminating the privacy paradox - the tendency of user's to disclose more in their actual behaviour than in their previously stated intention (Norberg et al., 2007). It is the often observed willingness to disclose that points strongly to a heuristic disclosure decision system at play. It is a trust over scrutiny exchange. Therefore an environment that aligns with the standard definition of a tragedy of commons, i.e., individuals trust that requests are valid and disclose freely. However, this behaviour only serves to incentivise greater reach in requests. In turn, this greater reach leads to a depletion in the value of PII being processed as a resource, thus depleting that particular PII common. A new commons of data is sought, repeating the cycle, and ergo, escalation in the requested volume of PII.

Credibility and disclosure decisions are both trust-based decisions (Metzger, 2006; Sundar et al., 2013), and heuristics are somewhat abstracted from the *weeds* of a problem Kahneman (2011). In fact, it is predicted that in some circumstances one simple heuristic can take an individual through a full cycle of disclosure through related yet independent decisions. That is, the user moves from a credibility judgement regarding the legitimacy of a service, through to a judgement to determine a service's trustworthiness as a data controller, and finally to whether the individual is willing to disclose a particular item of data.

4.1.1 Summary of Aims

This chapter, the majority of which is published in Marmion et al. (2017a), works towards addressing RQ2 (How are users making disclosure decisions?), by presenting a qualitative study into the heuristics that people use when making disclosure decisions. Established credibility heuristics have been used as a starting point to present an analysis of 23 semi-structured interviews, with the aim of exploring whether the heuristics related to credibility judgements are a general enough framework to also apply to disclosure decisions. To also address RQ3 (How does context affect disclosure?), interviewees were primed to think within one of two possible self-regulatory mindsets (Higgins, 1998); twelve were promotion primed around situations relating to social activities and gains, and eleven were prevention primed around situations relating to financial transactions. This work also seeks to develop superordinate classes of similarly themed heuristics, and to explore the importance and limitations of heuristics within those classes, and therefore within the disclosure decision process as whole. These six classes of heuristics feed into the next phase of the simulation work (RQ4: How do RQs 1, 2 and 3 combine within the identity assurance commons?).

4.2 Method

A series of semi-structured, one-to-one, face-to-face interviews were conducted. This approach was consistent with the qualitative nature of the focus groups in Metzger et al. (2010) and Gambino et al. (2016), but had an advantage of providing for a deeper focus on individual experience. The interview structure followed that of a cognitive walk-through. This was chosen as it is productively used for heuristic evaluations within human computer interaction (HCI) studies (Nielsen, 1994; Hollingsed and Novick, 2007). However, instead of a specific target system as in HCI studies, the interviewees were asked to recall an interaction with an online service. This meant that no system had to be contrived, and the focus on what had already occurred avoided talk of ideals that misalign with actual behaviour (i.e., the privacy paradox) (Norberg et al., 2007).

Furthermore, focusing on interviewee interpretation (the heuristic) means that the actual system, and the content of the cues was of less importance than the type of cues. For instance, one person may look for a celebrity seal of approval, whereas another looks for a Royal Warrant mark, either way they can both be using the Endorsement heuristic. This meant that the required sample were those who regularly engaged with online actives and services, and who could also reason and articulate about these engagements. With this in mind individuals were recruited from the Psychology department's participant pool, resulting in 23 Interviewees, all aged 18 to 25 years old.

The interviews lasted less than an hour (43 to 57 minutes), and were conducted on campus. Interviewees were briefed, and given the opportunity for questions before signing a consent form. The audio recording then commenced for the duration of the interview, and at the end they were debriefed. Participants were paid a small amount of compensation for their time. Finally, the audio files were transcribed verbatim by the interviewer, and imported into NVivo for analysis.

The interviews comprised three stages. Stage one (approx. 5 minutes), involved simple questions that established the interviewee's general digital engagement, whilst also easing them into the interview process.

In stage two (approx. 40 minutes), each interviewee was asked to recall a recent instance whereby they registered with an online service. The interviewees were then primed to think within one of two possible self-regulatory mindsets (Higgins, 1998); twelve were promotion primed to recall a system relating to social activities, entertainment or freebies (in the analysis these are denoted with S, for social, i.e., S1, S2,...S12), and eleven were prevention primed to recall a system relating to responsibilities or financial and commercial transactions (denoted with T, for transaction, i.e., T1, T2,...T11).

Once a relevant situation had been recalled they were asked to discuss the process from first considering the service, through to completing the registration or transaction. They were allowed to speak freely around the task, however, when recall became disjointed a prompt sheet of short questions that moved chronologically through the process was used for reference. While the interviewer steered the conversation around the contexts of the original primer, some contextual cross-over was unavoidable, and in many cases these instances were insightful.

Finally, stage three (approx. 5 minutes), included questions regarding the general concept of identity and privacy in the media. As well as a winding down exercise, this section allowed the interviewees to express any related thoughts or concerns that may have occurred during the interview. See Appendix B for the full interview schedule, and for the prompt sheet.

4.2.1 Analysis

The analysis was undertaken in three parts. The first part was ensuring familiarity with the data. I conducted the interviews, and transcribed the audio recordings, which ensured a base level of familiarity with the data. The data was then divided into; A) the data relating to disclosure decisions, and B) the data not relating to disclosure decisions. Before being set aside, data set B was examined to provide context and validity as to the interviewees' suitability for, and engagement with the process.

The second part involved coding the interview data into distinct categories. To do this, the transcripts were examined using an interpretivist approach (Holloway, 1997; Fereday and Muir-Cochrane, 2006), wherein the interviewee constructs the theoretical connection between cue and decision. The analyst is then left to categorise the self-reported 'rules of thumb' (Sundar, 2008).

An analysis challenge stems from people combining heuristics or interweaving heuristic and non-heuristic interpretations to inform decisions (Gigerenzer et al., 2008). To help address this, data set A was first categorised into; A1) heuristic-based, and A2) non-heuristic decisions. Then by matching the language used and the cues mentioned with those outlined in the literature, data set A1 were deductively categorised to align with the heuristics in Table 2.3 as described in (Metzger and Flanagin, 2013).

Data not aligned with Table 2.3 were then inductively coded, as described in (Ryan and Bernard, 2003), to reveal additional heuristics. There is a question however, as to the value of producing an ever-expansive set of heuristics separated only by subtleties. For instance the Intrusiveness heuristic, i.e. unsolicited communications inhibiting disclosure willingness from Gambino et al. (2016), shares similarities with the Persuasive Intent heuristic i.e. a feeling of bias or being pushed that inhibits the willingness to disclose (Metzger and Flanagin, 2013). Likewise, Sundar et al. (2020) expand on the heuristic set from Gambino et al. (2016) in a manner that enriches the discussion yet in a way also complicates it. A balance is required to avoid an ever expanding set of decision cues such as in the MAIN model, and remaining in the scope of pragmatic findings. Subsequently, the focus here is on developing superordinate classes of similarly themed heuristics that might accommodate and frame emerging work. Data set A2 provided a set of non-heuristic decisions that, once coded inductively, acted as deviant cases to counterbalance any confirmatory-bias residing in the efforts to explore heuristics.

Part two teased the results apart. Part three of the analysis serves to recombine them by identifying super-ordinate classes that encapsulate part two. These classes are emergent from the analysis and not formally derived. The formation of these classes allows the extension of the original credibility framework to maintain its richness and keep the overall results concise and pragmatic. For example, while the Reputation and Recognition heuristics involve different interpretations and reasoning, they seem based upon similar underlying cues (for instance, the size of the organisation). Therefore, a parent *prominence* class provides the crux of the discussion, wherein a participant referring to size can potentially be referring to reputation or recognition, yet more certainly, albeit abstractly, referring to the prominence of the subject.

4.2.2 Limitations

The interpretative approach is open to self-confirmation bias. Not only is the interviewee framing their decision in their preferred language, but the coding task also mapped onto the pre-populated list where possible. It is within reason to suggest that overlapping or vague phrases are miss coded. This limitation is an additional advantage of the super-ordinate approach as miss-coding, or miss-interpretation is more likely at the granular level yet less likely in the abstract. Moreover, the recall design on the interviews are subject to miss-remembering, or indeed, favourable post-rationalisation. Openness about less than adequate online security points to candidness, but it remains a valid limitation in design.

Furthermore, while the interview approach to establishing the heuristic framework is a shared choice with similar research, something is appealing with the behavioural experiment that put these decisions, and the cues that invoke them, to the test in a laboratory environment (Sundar et al., 2020). Years of validation is required of the super-ordinate classes and then again for the class components. The limitation is plain that this work is within a growing body of research yet to reach the maturity of the rational decision-making counterpart.

4.3 Interviewee Context and Validity

Interviewees reported habitual engagement with digital living, with typical comments such as; “Oh, literally, all the time”, and “when I wake in the morning, I just look in bed on Facebook.” However, remaining mindful of the temporal nature of the responses Knijnenburg (2013), as summed up by interviewee T6 when they said “I was 15, now I’m 19, so I have different interests.”

The Interviewees were open and candid within the interviews, this was exemplified by a common revelation about having only two or three passwords across all online systems, a finding that mirrors those of Komanduri et al. (2011). For example, one participant admitted “when I am creating a password they say you need a capital or a punctuation so it might vary, but generally I use one of three.” In some cases the participants even admitted to writing them down for ease, “I just created a word document to remember all of my passwords.”

While the similar age and educational status of the participants naturally scopes these findings to a particular demographic, it is a key demographic for the problem of disclosure. The analysis of the first stage of the interview shows that the participants were engaged with the problem of disclosure in their everyday lives, and prepared to give rich answers to the interview questions.

4.4 Findings

Table 4.1 summarises the results of the second and third part of the analysis (coding, and identifying super-ordinate classes), and contains sample dialogue to illustrate the coding process. Six super-ordinate classes are discussed; PROMINENCE, NETWORK, RELIABILITY, ACCORDANCE, NARRATIVE, and MODALITY. A seventh non-heuristics TRADE class was also identified. Within the following sections, the evidence is self-contained within each paragraph, however, throughout I have provided the fuller, encapsulating extract to provide the reader with additional context and supporting evidence. Also, during these findings the analysis tends to follow the decision maker through the full cycle of disclosure, leading an interviewee from an initial assessment of service legitimacy, through the assessment of the service as a trustworthy data controller, and finally to the assessment as to whether the interviewee actually disclosed an item of PII. I include the full quote in slightly greyed out text within the findings, however, this it for extra contextual clarity as the relevant information is integrated into the narrative.

4.4.1 PROMINENCE: Recognition and Reputation

Many of the participants expressed terms aligning with the **Reputation** heuristic by implying that a prestigious service would not knowingly do wrong, as described by Metzger and Flanagin (2013). When asked for location data from a game, Interviewee S3 says; "I mean FIFA is well-known and probably not evil." This is the first example of a simple heuristic, and in this case, having a good reputation was sufficient to complete the full cycle of disclosure. For instance, we know that reputation relates to credibility judgements regarding the legitimacy of an organisation, but also within this interaction with S3 reputation also provided an implicit trustworthiness of FIFA as a data controller, and then, reputation specifically in regards to their willingness to disclose their location data (Q:1).

(Q.1) [...] if it's like a really, really, poorly designed game and it is asking for my location I may say no. Then it is probably not a trustworthy company. But if it is something like FIFA then I wouldn't mind that. I mean FIFA is well-known and probably not evil. *S3.*

Interpreting cues related to size, being low-key, or being a known brand were typical. Such as when Interviewee S5 suggests that "Twitter is such a big company you assume they would not [...] pass your information on." In a related tone, Interviewee T11 associates size with risk (Q:2), as an organisation would want to protect their reputation, and thus protect the user, and the bigger the organisation the better the protection. There is an overall sense that if something has gained prominence then it must be doing something right, whereas lacking prominence suggests otherwise.

(Q.2) Small companies don't have so much security, so giving my data to a company like that is more risky [...]. Because smaller companies don't have as many resources for security. *T11*

TABLE 4.1: Superordinate Classes and Heuristic Coding Reference

CLASS	**Heuristic:** Description	Example Extract -*Interviewee*
PROMINENCE	**Reputation:** Prestigious services would not knowingly do wrong	“FIFA is well-known and probably not evil.”$_{\text{S3}}$
	Recognition: A familiarity with a service, even in name only	“it is just a very small app that I have not heard much about, I think I wouldn’t put my information on it.”$_{\text{S4}}$
NETWORK	**Endorsement:** A recommendation from known others	“My brother has been telling me it is more secure, it is easier, better and safer.”$_{\text{S1}}$
	Authority: A recommendation from official or primary authority	“He was a journalist, so he knows a lot of those sort of things.”$_{\text{T7}}$
	Bandwagon: Perceiving the actions of unknown peers or general population	“I was quite influenced by what everyone was doing.”$_{\text{T7}}$
RELIABILITY	**Consistency:** Interacting with a familiar process	“I tried another website, and also they ask for the same thing. The same questions. You have to sign up first, and it was the same thing. So I signed up.”$_{\text{S1}}$
	Consensus: A normative or standardised process	“Just the normal, name, date of birth and the important one is the mobile number to create an account.”$_{\text{S10}}$
	Expectancy:[1] Inferior site design, errors, poor visual appearance	“this looked fashionable and genuine.”$_{\text{S6}}$
ACCORDANCE	**Intent:**[2] A feeling of bias or being pushed	“they wanted all my details to tell me how much it would cost. So I provided false details.”$_{\text{T3}}$
	Self-confirmation: Feeling a consistency with pre-existing beliefs	“Why do you need ID? I’m only buying make-up.”$_{\text{T11}}$
MODALITY	**Coolness:** Gratifying features of a technology	I like the effect on the photos, I only did it for that, I don’t like the privacy really, but I don’t really use it very often, it is just on there in case.”$_{\text{S6}}$
	Novelty: An new encounter with a technology	“When I actually started, I was so happy about it, that I completed absolutely everything.”$_{\text{S3}}$
NARRATIVE[3]	**Availability:** The ability to recall similar instances	“there has never been a dodgy situation when I don’t want to give [my location data] because it is harmless games like Flappy Bird”$_{\text{S3}}$
	Coherence: The ability to envisage the result of an action	“if someone hacked my Twitter account I honestly wouldn’t care because it is utter rubbish, nothing important, it is only entertainment”$_{\text{S2}}$

1: Violation aspect removed from the Expectancy heuristic to provide neutral label.

2: Persuasive aspect removed from the Intent heuristic to provide neutral label. The Intent heuristic also incorporates the Intrusiveness heuristic from Gambino et al. (2016).

3: Not part of the original credibility heuristics framework as described in Table 2.3

This sentiment is also reflected in the **Recognition** heuristic; trust occurring due to a basic familiarity with an entity (Gigerenzer and Todd, 1999; Metzger et al., 2010). T6 (Q:3) makes the connection from the prominence of a high-street presence, and thus being "well-known" and "not trying to scam me". This may seem similar to reputation, but there is value in the distinction. Reputation seems to involve other people, as in "well-known" compared to an inwards reflection, as in "I have not heard".

(Q.3) If the company wasn't well-known, then A) I wouldn't be keen to register, and B) I would want to read [privacy policy]. [...]. So like high street shops, they are not trying to scam me, or whatever. *T6.*

Perhaps the most notable difference between recognition and reputation, is that reputation extends beyond the original entity towards subsidiaries. Gambino et al. (2016) refer to this as a safety-net heuristic, exemplified by Interviewee T10 when they state "with independent people, you need a barrier" (Q:4). This is a repeated factor in disclosure decisions, yet still a factor of reputation because reputation is acting as a form of collateral for the data exchange, as T1 feels that reputation is something "to live up to". For instance, online organisations acting as a trust intermediary for other associated organisations, because the parent "company image is that valuable" (S11, Q:5). This protection is also inherited by other service users, e.g., S4 finds it "really dodgy" being young and female on "a site that isn't well-known", yet on Twitter they "wouldn't worry too much".

(Q.4) They have a good reputation and loads of people had good experiences, I trust [Ebay] more [than Gumtree]. [...]. They do adverts on TV. With independent people, you need a barrier, that way they don't know my details, they don't know my sort-code or anything, they just see my username. *T10.*

(Q.5) [Apple's] reputation, the fact that they can't put an app on [their App store] that would pull money out of your account, they would have to refund it, their company image is that valuable. *S11.*

Seemingly, being of prominence provides organisations with many cues interpreted towards trust and willingness to disclose. There is then a scenario where credible organisations are "not trying to scam", which attracts users and in turn adds to them being "so well-known and so big you can trust it". This trustworthiness is reinforced by having the "resources for security" against external threats, and being "probably not evil" to be an internal threat. This cocktail of credibility and trustworthiness leading to willing to disclosure is a prime example of the simplicity in the cycle of disclosure. However, from Interviewee T2's perspective, "it is not really about the reputation it is about the price", reminding us that although these are simple heuristics, the decision cues remain diverse, and that when finance is involved it changes the decision further.

Further Prominence Related Extracts:

(Q.6) If they do ask for a lot of information, and it is just a very small app that I have not heard much about, I think I wouldn't put my information on it. *S4.*

(Q.7) I don't think I put my age, but I put that I was female, but I tend not to put my age if it is optional. Obviously, [older men] are on that [photography] site and search for young females. It was really dodgy. [...] I think it has [affected my behaviour], if it is a site that isn't well-known. Like Twitter, I wouldn't worry too much, but if it were a low key, photography or social networking site I wouldn't put too much information on there. *S4.*

(Q.8) [...] that is what people do, you know it is going to be OK, you can trust them, they are so well-known and so big. You can trust [Amazon]. *T12.*

(Q.9) [...] you wouldn't expect [big websites] to pass details to others to annoy you, I assume it is a one off from a small, less trusted site. *T12.*

4.4.2 NETWORK: Endorsement, Authority and Bandwagon

It is evident from the interviews that an individual's interpersonal network has a considerable influence on disclosure decisions. This was also reflected in (Metzger and Flanagin, 2013) through the Endorsement and Bandwagon heuristics. These heuristics are similar to the heuristics in the Prominence class, the difference being that the Prominence class regards a service's place in the world, i.e., a high-street presence, whereas the network class has a personal characteristic, i.e., my friends were doing it.

Focusing first on the **Endorsement** heuristic; testimonial by known others, Interviewee S2 found "that two of my housemates were already there made it seem more comfortable" (Q:10). These findings mirror that of Metzger et al. (2010), suggesting that in some cases individuals prefer recommendations over their own decision. Interviewee S1 reflects on this delegation to others to make decision for them, when admitting; "I am more affected by what people tell me as I am not really an IT person" (Q:11). This type of sentiment, by an educated individual with habitual use of technology, runs counter to the notion that disclosure is to be self-managed.

(Q.10) They advertise that very clearly, when you fill out the form. [...] We will not give this information to third parties. This is under the Data Protection Act. [...]. But the fact that two of my housemates were already there made it seem more comfortable anyway. *S2.*

(Q.11) My brother has been telling me [PayPal] is more secure, it is easier, better and safer, and people say that, and I think I am more affected by what people tell me as I am not really an IT person. I am bad at this kind of thing. *S1.*

The **Authority** heuristic is when trust stems from expert or official authority endorsements (Sundar, 2008). When Interviewee T7's father convinced T7 to allow electoral roll information to be traded, he was the authority, but not as a parent, "he was a journalist, so he knows a lot of those sort of things" (Q:12). In effect it is an endorsement from an individual with reputation, but is not passive like the reputation heuristic. This feature was seldom present in the interviews.

(Q.12) I got a letter [...], saying you are on the [electoral roll] register and [...] unless you opt out, basically, they sell your details to industry or anyone who wants them. But also banks use them to get your credit rating, or find information about you. So I didn't want to be on there, but it feels like no real choice. Well that is what my dad said [...]. He was a journalist, so he knows a lot of those sort of things. *T7.*

Similarly, the **Bandwagon** heuristic involves recommendations and often shares decision cues with the Endorsement heuristic (Metzger et al., 2010). However, the findings here agree with Sundar (2008), that the two are meaningfully different. In the Bandwagon heuristic, Instead of a personal endorsement from friends and family, the recommendations can be from unknown others via less personal factors such as aggregated testimonials or star ratings embedded within the interface. This places the Bandwagon heuristic conceptually close to the Prominence class, illustrated by "many thousands of people have downloaded them they can't be that bad" (S4, Q13). Yet, it also has a socially compelling aspect to it, as T7 (Q:14) explains, "I wasn't 100% satisfied with the [privacy policy]" but was "influenced by what everyone was doing" or as T10 reflects "I thought everyone else was. I assumed you had to fill it in".

(Q.13) [...] if they are the top results and so many thousands of people have downloaded them they can't be that bad. Like you would have heard of horror stories or something. But if it is like quite a new app that not many people have downloaded I might be a bit more sceptical. *S4.*

(Q.14) I wasn't 100% satisfied with [the privacy policy], but I still got a Facebook account anyway. [...] it is sort of impossible to have anonymity any more. [...]. Your option is to sign up or not. [...] Not really [an option], everyone is doing Facebook, I was a teenager at the time, maybe 14 or 15 when I signed up. I was quite influenced by what everyone was doing. *T7.*

Throughout this Network class of heuristics, there is a degree of delegating the decisions within the cycle of disclosure, through direct council and endorsement, or indirectly through the assumed behaviours of peers. There is a free-riding aspect whereby there is an expectation of others doing the risk discovery (Vila et al., 2003). But this is a self-fulfilling 'social proof' (Cialdini and Trost, 1998), whereby a herd mentality can follow without due consideration of the circumstances (Acquisti et al., 2012).

4.4.3 RELIABILITY: Expectancy, Consistency and Consensus

The **Expectancy Violation** heuristic has negative connotations surrounding poor design, central to which is an expectation of professionalism (Metzger and Flanagin, 2013). In this regard, on numerous occasions interviewees were cued by presentation details, with features such as poor layout, inferior design or errors impacting on perceptions of service trustworthiness. To this end T6 trusted their judgement "by looking at their website or social networking site, whether it looks professional or not", or for T7 it was that, "something in my mind saying it is not right". I have chosen to remove the 'violation' label and simply use Expectancy, because the cues can work both ways, in that "you sort of get the feeling that it is not right, but this looked fashionable and genuine" (S6, Q:15), this relabelling reflects that writing inconsistency, non-consensus, or disreputable would be unproductive.

(Q.15) Sometimes you get dodgy websites saying "get money off this", and you sort of get the feeling that it is not right, but this looked fashionable and genuine. I don't know if you can quantify something looking genuine. So trustful of me. *S6.*

This Reliability class also contains the **Consistency** heuristic; trust based on the agreement between two independent sources (Metzger et al., 2010). When S1 (Q:16) says "I didn't want to sign up, so I tried another website, and also they ask for the same thing. The same questions. You have to sign up first, and it was the same thing. So I signed up.", we say they are using the Consistency heuristic. In this case, S1 expected to engage without registering, however, when it became apparent that the seemingly non-standard requirement for registration was a consistent requirement across similar TV services, the user became willing to disclose.

(Q.16) [...] there is this one pop-up and there is this [message] you can watch [TV] if you sign up. I didn't want to sign up, so I tried another website, and also they ask for the same thing. The same questions. You have to sign up first, and it was the same thing. So I signed up. *S1.*

Similarly, the Reliability class includes the **Consensus** heuristic; a normalised and general agreement (Sundar, 2008). Consensus has a broader application than the Consistency heuristic. These situations are exemplified by the use of normative terms, such as Interviewee S4 noting that "obviously name, email address" and "obviously it wanted a photo" to describe an interaction with a social networking site. Likewise, Interview S10 (Q:17) with "[j]ust the normal; name, date of birth", and T6 (Q:18) with "obviously name and email", however, the data within these normal and obvious requests did tend to differ. The result is that "it almost bypasses you because you have done it so many times, but if something unexpected came up like a page you have not seen before that would make you doubt it. [...]. If it is the same process as usual I would assume it was fine" (T12).

(Q.17) Just the normal, name, date of birth and the important one is the mobile number to create an account. *S10.*

(Q.18) My address, a billing address, obviously name and email, that was mostly it, maybe a password for my account. *T6.*

The three heuristics are linked by the idea that if something is broken, has mistakes, or if something changes, it can cue users against disclosure, whereas a professional and as-expected interaction goes unnoticed. Problematically, this manner of thinking could incentivise service providers to request more information than is currently required, because in waiting to do so at a later occasion, the service risks disturbing the user's sense of routine and invoke questions such as S6's when a TV service started asking for information; "why are you doing that? It used to be different. They didn't use to ask for details" (Q:19).

(Q.19) It is not that I necessarily want to give them it, [giving my name] is the norm. I expect to give it, whereas with TV, it is only recently they have asked for more details. I'm like, why are you doing that? It used to be different. They didn't use to ask for details. *S6.*

4.4.4 ACCORDANCE: Self-confirmation and Persuasion

When Interviewee S6 expressed that a TV service "didn't use to ask for details" they were disrupted as a factor of the consistency heuristic, however, when they then reflected on "why are you doing that" this is closer to the **Self-Confirmation** heuristic; when something aligns with one's prior belief (Metzger et al., 2010). S6 later went on to reflect on why the BBC iPlayer "need[s] to check up on me, and my full name? To see what I'm watching?". Then as a result, "I just put my initials in, because I'm just watching TV".

The Accordance class differs from the Reliability class, in that it refers to beliefs and understanding rather than process or interface. Also, the Self-confirmation heuristic does not require a norm to the request, as long as there is an understanding that the request "comes up for good reasons" such as a store requesting a delivery address from S3 (Q:20). Whereas, when asked for ID for a birthday promotion T11 refrains because “that is not a good reason to give my ID especially when I just want to buy make-up. I wasn’t happy, so I didn’t sign up”. They could not justify the disclosure when told the reason, although in contrast T11 (Q:21) did give their ID to a storage company when told it was in case “something happened” despite being vague and not particularly compelling.

(Q.20) [If a game] asks for really personal things like an address, then I would rather not give them that. [...] on deviant art, you can order things from the store, so they need your address, so fine. [...]. So I understand when it comes up for good reasons, but when it is just for demographic reasons. No I don’t want to. *S3.*

(Q.21) Normally if I am asked for ID I ask what it is for. Like when I booked storage, they asked for my ID, I asked why, they explained that [it was for] if something happened. But I asked them, I don’t just give it. But for example, there is a make-up brand [...] they asked for my ID, but I questioned this, why do you need ID, I’m only buying make-up. They were just like it is for our database, oh for your birthday, we can send you a present. I was like, OK, that is not a good reason to give my ID, especially when I just want to buy make-up, I wasn’t happy so I didn’t sign up. *T11.*

Also in the Accordance class is the **Persuasive Intent** heuristic, the underlying principle is that of which perceived manipulation leads to negative judgements (Metzger et al., 2010). For instance, pop-up messages have been shown to produce a negative psychological effect (Fogg et al., 2003; Ward, 2003). Throughout the interviews, such instances related to unsettling aspects of an interaction being “too violent, in your face” (S3, Q22) or annoying features that “as you try and get a page and they are flashing up at you” (T4, Q23). Gambino et al. (2016) recognise such aspects as being an *intrusiveness* heuristics, leading users to “question the integrity” of the service. Removing the ‘persuasion’ part of the name of this heuristic to leave it labelled simply as ‘intent’ serves the purpose of being close to the ‘integrity’ element in Gambino et al., whilst maintaining a neutral description. Intent better describes the grey area between it being “quite helpful if they have picked up on what you are trying to find” and “it seems to be everywhere, [...]. It is annoying and unnecessary. I suppose there is two sides to it”.

(Q.22) [The promotion] was in your face without giving you a chance. [...] I never considered it because it was too much, too violent in your face, and you cant get rid of it. *S3.*

(Q.23) It can be quite helpful if they have picked up on what you are trying to find and shove it in your face. But it seems to be everywhere, as you try and get a page and they are flashing up at you. It is annoying and unnecessary. I suppose there is two sides to it. *T4.*

We learn more about this class when Interviewee T7 (Q:24) implies that paying for prominence on a search result was something to be "wary about", as if it was not in the spirit of things, compared to those who achieve prominence through merit of popularity. Or when Interviewee T2 (Q:25) was deterred by an insurance company because "they wanted all my details to tell me how much it would cost". In this instance, T2 realised it was a 'consistent' process for insurance companies to request this, yet the feeling of being pushed meant they "provided false details".

(Q.24) I tend to never click on the sponsored links, because I fell like they are not the most searched for, they are just paying to be there. If I did click one, I would be be wary about it. But if it is from Google, top result, I wouldn't question it. [...]. a lot of people have searched for it, so you think it is not phony *T7.*

(Q.25) [...] I wanted to find out how much [motorbike] insurance would cost, and they wanted all my details to tell me how much it would cost. So I provided false details. *T2.*

Further Accordance Related Extracts:

(Q.26) I don't really know why I was reluctant to put my age on there. [...]. The way that Facebook tends to target adverts depending on your likes. That annoys me massively.*S4.*

(Q.27) I know it is a clever ploy by advertisers [...] because my relationship status changed and there was loads of adverts about wedding stuff. It just annoys me that they can do that. But [...] I found my wedding photographer through one of those advertisements but most of the time it annoys me. *S4.*

4.4.5 MODALITY: Coolness and Novelty

Sundar (2008) associates the **Coolness** heuristic with new technological features, or the bells and whistles of existing technologies, with positive credibility evaluations. For instance, Interviewee S6 consents [Instagram] access to all their photos, despite that they "don't particularly like to, but you can't download it without giving that permission", and they "like the effect on the photos" (Q:28).

(Q.28) I don't particularly like to but you cant download it without giving that permission, [...] I like the effect on the photos, I only did it for that, I don't like the privacy really, but I don't really use it very often, it is just on there in case. *S6.*

The **Novelty** heuristic is subtly distinct from the Coolness heuristic as it is invoked by a user's initial experience with a technology (Sundar, 2008). S3 describes two instances of “when I first started Facebook I think I got a bit carried away”, and “when I actually started [Deviant Art] I was so happy about it, that I completed absolutely everything” (Q:29). However, that early exuberance waned and “looking back on my profile I used to disclose more information than I do now” (Q:30).

(Q.29) You have this other personal information page and you can choose to display or not, or to complete it or not [...] When I actually started [Deviant Art] I was so happy about it, that I completed absolutely everything. [...] I reactivated [my account] but I didn’t put a profile picture any more, I didn’t complete those details any more. *S3.*

(Q.30) When I first started Facebook I think I got a bit carried away. It was years ago when I actually joined. Yeah looking back on my profile I used to disclose more information than I do now. *S3.*

In Sundar’s MAIN model (Sundar, 2008), these heuristics are seen as a factor of modality. Instances of these heuristics were sparse, and limited to social and entertainment situations, perhaps aligned with the explanation that in these instances individuals are mostly concerned with gains and immediate gratification (Higgins, 1998; Acquisti, 2004a).

4.4.6 NARRATIVE: Availability and Coherence

The framework in Table 2.3 was insufficient to explain all of what the interviewees described. This is mainly because the credibility framework referred to individuals establishing trust, it does not account for individuals considering risk. Instead of “why are you doing that” type questions, interviewees would engage with past examples, and/or hypothetical situations, asking themselves “why would they be interested in me”, or more pertinent, “if I had been affected” type reflections.

To frame these instances of introspection, refer to the description by Tversky and Kahneman (1973, p. 15) of the **Availability** heuristic; a judgement of the likelihood of an event based on the ‘ease with which relevant instances come to mind’. A example is S4’s work insight meaning they “would not sign up for anything like [comparison websites], because I worked in insurance and basically if anyone put information on GoCompare it would come straight to us” (Q:31).

(Q.31) I would not sign up for anything like that [comparison websites], because I worked in insurance and basically if anyone put information on GoCompare it would come straight to us, [...] Then we would call them and [...] they were like how did you get my number and I would say did you go on a comparison website and they would say I think so yeah. Well, it was from that. *S4.*

Interviewee S4's experience was not typical in relation to those less aware, such as Interviewee S3's lack of risk availability in that "there has never been a dodgy situation when I don't want to give [my location data] because it is harmless in games like Flappy Bird". There were many similarities between the perspectives of the interviewees here and those in Balebako et al. (2013). That is, when the extent of data leakages were revealed to their participants they were 'very surprised' by the frequency and the destination of data leakage from mobile games. The overriding difference is that in Balebako et al. (2013) the full extent of data disclosures as a result of playing a game was demonstrated, which in turn, allowed their participants to envisage a list of possible negative outcomes, and therefore they were able to complete a disclosure narrative. In the end, these participants stated a desire to change future behaviours, and one participant even changed from perceiving disclosure as useful for customisation, to later referring to the game as being 'slime'.

It is unsatisfactory to wait for users' negative experiences to instil a more cautious, considered approach to disclosure. Instead, it may be possible to inform users of disclosure risk through a relatable narrative. In this regard, refer to a **Coherence** heuristic; being able to envisage the result of a decision as a plausible outcome. For example, S11 (Q32) does not profess to having been mugged, yet they can reason that "when you post a picture you can add your location then people in the area can look at the picture and they can find you and they could mug you or something like that". Also, S2 (Q:33) can envisage that it is possible to hack a Twitter account, yet "honestly wouldn't care because it is utter rubbish, nothing important, it is only entertainment". Seemingly, the interviewee does not have the available recall or imagination to see the potential negative results, such as those increasingly experienced by victims of facility takeovers (Hoofnagle, 2007; Kahn and Roberds, 2008).

(Q.32) Yeah, location data definitely worries me. [...] I don't have anyone after me, but if there was somebody looking for me they could find it out. I don't like sharing info about where I am. [...] say if someone wanted to steal my phone. Actually, for example, on Instagram when you post a picture you can add your location then people in the area can look at the picture and they can find you and they could mug you or something like that. *S11.*

(Q.33) I think there [is] more of a risk when it is to do with financial stuff because it is a concrete thing. It is your bank account, which is valuable. I'm not saying I don't value my identity because I do, [...] But, if someone hacked my Twitter account I honestly wouldn't care because it is utter rubbish, nothing important, it is only entertainment. *S2.*

Norberg et al. (2007) finds that people abstractly perceive risks in over-disclosure, yet when faced with a specific disclosure decision they most likely disclose. The evidence in this study contributes to that observation, and further suggests that the often missing narrative could play a significant role. In practice however, due to the consent model being 'simple', such narrative is rarely available to the user, and therefore, the resulting behaviour is similar to that observed in Norberg et al. Furthermore, there is an expectation of sorts that this narrative will be brought to them as noted by T1 (Q:34) who "assume[s] that if I had been affected [...] I would be contacted by eBay. [...]. Only at that point if that happened would I care about it a bit more". Or T3 (Q:35) suggesting that "[i]f there was a serious problem I'm sure it would be in the news".

(Q.34) I didn't think or consider it really [...] I would assume that if I had been affected [...] I would be contacted by eBay. [...]. Only at that point if that happened would I care about it a bit more. *T1.*

(Q.35) I just trust [Amazon] as a company [...]. They are a very reputable, well-known brand. If there was a serious problem I'm sure it would be in the news and I have not heard it. *T3.*

Further Narrative Related Extracts:

(Q.36) I don't know what they could observe, it is just how good you are at the game, and you have flash-up adverts on the screen so they must have [data collection] for that, but I don't think it is more than that with Angry Birds. *S6.*

(Q.37) I'm not sure how it works, but from my experience nothing has gone wrong yet, so they are pretty efficient about keeping things confidential. *T9.*

(Q.38) I'm not really bothered to be honest, because I couldn't understand why they would want [personal details], if they are collecting from me, they are collecting from millions, so it doesn't bother me. *S11.*

4.4.7 TRADE: Gains and Worth

Despite the primary focus on heuristics, examining the data for deviant cases countered some confirmation bias. From this it was evident that along with heuristics, interviewees were also weighing up their disclosures in terms of trading utility gains versus losses (Acquisti and Grossklags, 2007, 2004). In many ways the Modality class (coolness and novelty) reflects the notion of a trade. Interviewee S3 considers that "it doesn't seem like a good investment" to disclose location data to a poorly designed game (Q:39). Seemingly this is interpreted in a manner associated with the Effort Heuristic (Kruger et al., 2004), in that lack of effort reduces utility. Conversely, S6 explains that it was "quite a lot of details, but I felt like I was getting something back with the [rail] voucher".

(Q.39) Most of the time it is not the actual proper original game, you get a copy. But if the copy is really, really good, then fine, I will give my location and first name, whatever. But if it is really, really bad, then it doesn't seem like a good investment. *S3.*

Trade-type behaviours were often imbalanced in favour of disclosure. For instance, S9 (Q:40) perceived a lack of real option "[w]hen Google linked Gmail and YouTube [...] I didn't have much of a choice, because "I didn't want to lose" my "personal videos" and "amateur stuff". Likewise, S11 described how "[Facebook] force you to have [Messenger] on your phone", with the sentiment that "I sort of need it. I have 100-200 friends on there that I need to contact" (Q:41). S11 also reflected on "why would they be interested in me?" Conveying a common sentiment of insignificance around personal data (Heikkinen et al., 2006), hence apportioning a low overall value to the data disclosed, compared to a clear understanding of Facebook's utility.

(Q.40) When Google linked Gmail and YouTube [...] I didn't have much of a choice, I would have had to close my YouTube account, and I didn't want to do that, I had uploaded a personal video of a character, some amateur stuff, but I didn't want to lose them. I could have gone to Vimeo, but no. *S9.*

(Q.41) Like Facebook Messenger, they force you to have it on your phone. But there is a huge issue. The app can send texts and receive them without your permission, or record audio without your knowledge. [...] But why would they be interested in me. I know they collect data to advertise, but that doesn't bother me. [...] The issue is I sort of need it, I have 100-200 friends on there that I need to contact, and Facebook is the best way because everyone has it. [...] My tennis club uses Facebook [...], my economics society; meeting up with people, they would all be hard. You would rely on mobile phone. I wouldn't forgo it. I wouldn't not use Facebook just for that. *S11.*

Within efforts to disclose in a more calculative manner, the variables underpinning the decision often remain heuristics based. When S2 explains a difference when disclosing "name, phone number, that I need to give, that is fine. But I wouldn't tell them where I'm working or what I study". There seems to be a reliance on the Coherence heuristic when S2 is envisaging someone turning up at their work. Equally, S1 can envisage the risk and therefore caution "in terms of card and bank details. That will put me off subscribing or buying online", but this is only a relative value as "mobile, or equally email, is the least worrying compared to card" (Q:42).

(Q.42) I am more cautious in terms of card and bank details. That will put me off subscribing or buying online [...] Mobile number I would still give to an online forum. I find mobile, or equally email, is the least worrying compared to card. *S1.*

In some instances there was a relative value to single identifiers, i.e., name vs. place of work, whereas, in other instances disclosing a combination of identifiers impacted the valuation. For instance, Interviewee S4 talked of less willingness to disclose their age once gender had already been disclosed (Q:43). Seemingly, being a female was a satisfactory disclosure, but not in conjunction with being young. Not evident is whether age in isolation holds the same value as when in combination with gender, or how this value may change over time. In contrast, from S9's perspective there is a threshold effect;"I consider my phone number a pretty private thing to begin with it, so if someone has it, they already probably know my name."

(Q.43) I put that I was female, but I tend not to put my age if it's optional. *S4.*

Further Trade Related Extracts:

(Q.44) When I give them my information in the first place, I think, do I want it to be up there, yes or no? Name, phone number that I need to give, that is fine. But I wouldn't tell them where I'm working or what I study at [university], or things like that. *S4.*

(Q.45) so when you are shopping and sign up, do you read the terms and conditions Sometimes, but not every time. Why sometimes and not others? If I trust the company I would be less likely to read them. *T6.*

(Q.46) I guess maybe they sort gender or age of people using their websites, so they can target their audience better. So if targeting advertisement are occurring and adverts are relevant, is that something that concerns you? I guess if it is just gender and age, I don't mind, it depends on how much information they are taking, I don't mind. I guess if they are not doing it for anything bad, I understand, it makes sense, like your general location for adverts in your area. So as long as they are not trying to cause you harm it is OK, just a bit annoying. *T12.*

4.5 Conclusion

The interviews conducted have provided a rich qualitative account of users interaction with the disclosure decision points of online systems [RQ2]. Looking at the simplicity of such decisions through the lens of credibility heuristics, validates the prediction that the credibility heuristics are also being used for disclosure decisions. Also, disclosure heuristics outside of the credibility framework were evident. Mainly this was the importance of narrative in how users make disclosure decisions. These results were then encapsulated within superordinate groups (Table 4.1), revealing PROMINENCE, NETWORK, RELIABILITY, ACCORDANCE, NARRATIVE, MODALITY and a seventh non-heuristics TRADE class.

The implication of this work is that the self-managed model, whereby self-informed individuals are responsible for consenting or withholding personal data, is idealistic. The evidence here is that users tend to make impoverished decisions. They evaluate trustworthiness from heuristics formed of prominence and social networks, using decision cues such as popularity, brand exposure or word of mouth, resulting in somewhat of a herd mentality. Alternatively, users evaluate trustworthiness from heuristics formed of accordance with beliefs and a sense of reliability, using cues such as familiarity and regularity. However, this reasoning has inductive pitfalls based on the idea that if nothing negative occurred before as a result of a disclosure, then future like-for-like instances are deemed safe. The results here agree with Sundar et al. (2013), and Gambino et al. (2016) that a reliance on such cognitive heuristics is key to understanding users knowingly consenting to give more than intended (i.e., privacy paradox). Also, this is potentially key to understanding users consenting to give more than they know (i.e., simple consent).

Norberg et al. (2007) calls it a privacy paradox when describing how users base their disclosure intentions on risk, yet base their disclosure behaviours on trust. The resulting behaviour then favours disclosure, because in disclosure environments there are many trust based cues yet scarce information about the risks. Therefore, on occasions when users attempt a considered approach to disclosure, qualitative accounts of what may happen are not adequately portrayed, and users find it difficult to complete a coherent narrative which diminishes their ability to adequately conduct the 'privacy calculus' (Krasnova et al., 2010) required for informed consent. The hope is that these heuristics, and the implication of their susceptibility to bias and manipulation (Krasnova and Günther, 2009), could one day be harnessed so users can benefit from some form of positive nudge and thus mediation of the risks (Balebako et al., 2011; Adjerid et al., 2013).

These new super-ordinate set of heuristics for disclosure is envisaged to allow future research into the heuristics, and also to provide a place for emerging heuristics. For this work however, these heuristics can now be developed into a set of agent behaviours, from which they can be investigated within my social simulation of the identity assurance commons [RQ4].

Chapter 5

Modelling Identity Assurance

5.1 Introduction

This work builds on the premise that within the current self-management approach to identity disclosure, there is an inevitable disclosure escalation. This disclosure can be wilful or ignorantly and can be individuals to organisations, which is the focus here, or organisation to organisation which is an equally exciting issue. If the desire is to manage this escalation, then understanding the behaviours within identity assurance as a system at a fundamental level remains essential. Understanding identity assurance as a system is, however, not simple.

Exploring systems of identity assurance from a disclosure decision perspective encompasses three significant characteristics. First, that identity assurance comprises many elements, e.g., rational versus heuristic disclosure decisions, profit-driven data use, asymmetry of the information within consent, and market incentives for privacy or lack thereof. Second, that within the system, there are multiple stakeholder perspectives; the service perspective of competition, the user perspective of disclosure, privacy, and security, and from the resource perspective of PII sustainability. Then third, that the identity assurance systems necessitates data disclosure, the situation is not one of the services greedily taking data to sell products and or subsidise their provision.

The reason for some of this intertwining is that there is little conceptual difference in identity assurance conducted for security purposes compared to identity assurance for commercial targeting. Both assurance purposes require personal data to determine the identity of an individual, or at least that that individual belongs to a particular profile. Therefore, data extraction and disclosure for security is muddled together with the same systems used to collect data for commercial purposes. This confusion veils any attempts of transparency due to all the problems brought by complex decision circumstances, and users with limited attention, time, and or motivation. If these behaviours incentivise the further extraction of PII from individuals, then it presents a risk to privacy. Moreover, extraction can also amount as a threat to security, the very thing that many PII extractions serve to achieve. The modern digital world requires a modern fit for purpose policies that can balance these characteristics (Anne Toth and Lin, 2018), and this requirement needs to happen within an ever-moving set of social norms and technologies (Beer, 2009; Flanagin et al., 2010).

In this chapter, these elements combine into a single model of identity assurance. The challenge continues in that, even in isolation, these systems take much understanding and analysis; together, they represent a significant problem. To ground this ambition, the following describes an amalgamation of established models. The model in Vila et al. (2003) provides a base in the form of a game-theoretic model outlining a dynamic between services and consumers through the medium of privacy policies. In this context, services are data extractors, and consumers are the users disclosing data. Competition is then added to the model employing work relating to a Bertrand model of competition. Such models examine how competition pressures price towards low cost, break-even services or products, yet tactics can be deployed to maintain profitable enterprises (Vilà, 2008). Finally, sustainability dynamics are incorporated utilising a model which mimics natural resource harvesting and potential overuse. Each aspect is discussed as the chapter proceeds.

The findings of preceding chapters augment the model construction. The work on Chapter 3 provides a personal value metric for PII that compliments the Bertrand competition on price. The heuristics work in Chapter 4 adds realism to the trust-based model in Vila et al. (2003) by justifying the introduction of prominence cues. Equally, evidence within the review in Chapter 2 provides nuance into the system of identity assurance regarding the assurance value of various types of PII.

The outcome of the model is not predictive. Instead, consider this a model for ’illustration (using a simulation) to communicate or make clear an idea, theory or explanation’ (Edmonds, 2017, p2); it is a theory manifest. It is a tool that through its construction, acts as an opaque taught experiment about the complexity surrounding identity systems and the accompanying disclosure decisions. Then post-construction, it is a tool for reasoning about what may or may not drive the escalation of personal identifying information disclosure. This reasoning follows that assuming that PII depletes in extractable confidence, then there is a potential usage dilemma. Subsequently, the construction of a model aids consideration into what factors may speed or slow the degradation of the data upon which identity assurance systems rely. Finally, it is a means to generate ideas and questions for further work (Epstein, 2008).

5.1.1 Summary of Aims

This chapter focuses on my fourth research question, which is, how to do my first three research questions fit together. More aptly, how to these questions fit together within the theory that places identity assurance in a commons type system. The main body of this work is in five parts, and broadly these follow the introduction of five core elements. First, the game-theoretic model described by (Vila et al., 2003) is reassembled and presented as a pilot agent-based model. Second, a cost in terms of users' willingness to disclose, and value in terms of benefits to the service. The third introduction is the depletion factors of PII. Fourth, I introduce prominence as a proxy for trust. Finally, there is a discussion regarding the future work that can capitalise on this model perspective on the identity assurance system.

5.2 Building a Pilot Model

5.2.1 An Equilibrium Game

In a game-theoretic model, (Vila et al., 2003) present an equilibrium analysis of the macro behaviours involved in services exploiting user data. The model, based on a previous economic model of a car 'lemons' market, examines consumer behaviour while buying vehicles with or without checking the quality first. Exploitation occurs during periods of user complacency. The authors used this model to reason why online users don't read privacy policies and why that is harmful to the privacy of personal data. Equation 5.1 summarises this model, albeit with slightly different notations than the original.

$$(I^*, J^*) = \left(\frac{G - V - T}{G - V}, \frac{S + \epsilon}{R + \epsilon} \right) \tag{5.1}$$

where, I^*, a specific ratio of sellers respecting data, and J^*, the ratio of users testing for exploits yields a single equilibrium point, For the consumers, G is the gratification of transaction, T is the cost to test for exploits, and V is the cost of a data violation. Whereas, for the services, R is the base revenue benefit from the transaction, S is the cost to send a privacy signal, and (ϵ) is the extra benefit from exploiting user data. Whilst there exists an equilibrium, the authors note that the sensitivity of the equilibrium, to even small changes, meant that the over-riding behaviour is an oscillation wherein exploitation persists, as depicted in Figure 5.1.

Cost of Consciousness The cost of revealing the negatives associated with the disclosure are covering those very negatives. Besides, this cost to reveal any negatives is problematic, as it is easier borne by those in society with the most resources. Those

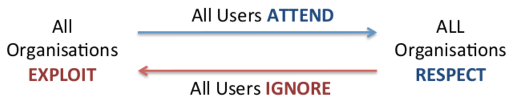


FIGURE 5.1: A privacy free riding dynamic adapted from Vila et al. (2003).

with limited education and or money, cannot readily understand the disclosure, or delegate the protection respectively. Therefore, if a service, for any reason, deems that this cost of consciousness is too high for the majority of users to incur, then it could lead to a moral hazard. This hazard is perhaps more so when the incursion may even happen in the background by third-party companies, or increasingly by the objects with which we interact. For, in the real world, people have less and less opportunity to opt-out in an environment integrating with ambient technologies (Cook et al., 2009). This situation exists for two competing services, but what about multiple services, then the cost of consciousness accumulates within the system. However, this cost spread over multiple services is seemingly below the engagement threshold for most users; hence the attractiveness of the Web. Or,

$$\sum_{i=1}^{n} G > \sum_{i=1}^{n} V > 0 \tag{5.2}$$

or,

$$\sum_{i=1}^{n} Gratification > \sum_{i=1}^{n} Negatives > 0 \tag{5.3}$$

yet,

$$\sum_{i=1}^{n} PerceivedNegatives > 0 \tag{5.4}$$

such that, during any particular moment users can perceive one instance of identity-related issues as being relatively high, at least in the abstract (Norberg et al., 2007). Nevertheless, during specific instances, engagement generally occurs as a consequence stay in the unconsciousness (Beer, 2009; Friedewald et al., 2007; Pollard, 2006). Yet, as any perceived negative increase, with media stories, whistle-blower revelations or personal victimisations, the gratification gained is the first casualty, and markets may fail to reach their potential. That is unless it is possible to move towards an environment that reasonably apportions risk, and the faithful services do not make a veil of respect over those with less ethical operations. Ideally then, a respectful service should want to bear some of the cost of consciousness, S, so that users on a large enough scale can differentiate and thus, punish exploitation.

5.2.2 An Agent-Based Replication

To replicate and adapt the (Vila et al., 2003) model into an agent-based simulation depict a system comprising of two competing, co-located, financially free to use, Wi-Fi services. In this system, there are two types of agent. The set $I = \{1, ..., m\}$ of m hotspot service providers vying for the engagement of a population $J = \{1, ..., n\}$ of n users. These services operate for free despite substantial costs due to their ability to use user data for alternative purposes, e.g., advertisement (Zhang et al., 2012). Users are also generally consenting to overtly and voluntarily disclose personal data to keep the service 'free' to use and or to gain gratification from, for example, a voucher or information (Acquisti, 2004a). Within this consent, the user accepts that along with security reasons there is also a need to disclose some information for the proposes of primary advertisements (Toubiana et al., 2010). A user also knows that there may or may not exist some covert extraction by the service. That is, the service may be collecting extra detail via a misleading, obstructive or confusing privacy policy (Furnell and Phippen, 2012), and or selling personal data to third parties (Laudon, 1996). Disclosure and extraction are distinguished here for emphasis only.

To establish the data cost for the user, hotspot $i \in I$ adopts a pure strategy $s \in S$, where $S = \{RESPECT, EXPLOIT\} = \{\epsilon = 0, \epsilon > 0\}$. Then, $u_i(s) = \rho_s.\delta_s$ is the utility of service i, given strategy s, where $\rho = X + \epsilon$ is the profit per user and $\delta = demand$. Thus, service agents adopt data strategies to maximise a return, and they can either RESPECT user data, i.e., $\epsilon = 0$ or they may EXPLOIT user data by taking more or selling it, i.e., $\epsilon > 0$. All service strategies are profitable, but services will discriminate even the smallest amount of profit gain as advantageous. In practice, profit per user would likely change depending on economies of scale, service overheads or even the residential nation of users. However, here the profit a service achieves is a linear function of user demand and current strategy.

Any user $j \in J$ then has the strategy $\gamma = \{observe, ignore\} = \{o, \tilde{o}\}$ as to whether or not to reveal the true disclosure cost of engagement. If ignoring, a user would be assuming that any service chosen is respecting, i.e., $\epsilon \to 0$, else they observe that $\epsilon = \epsilon$. Thus, while an ignoring user will use any service, an observing user will only use a respecting service. Although this model explicitly uses the cost of disclosure, it is understood that the true cost of exposure is unreasonable to determine due to dormant uncertainties (Weitzner et al., 2006). Moreover, these documents are not accessible for the general reader (Furnell and Phippen, 2012). Nonetheless, in this case, engagement occurs when $(\epsilon \to 0)$. Therefore, a service would expect to grow regardless of respecting or exploiting while $P(j \in J | \gamma_j = o) \to 0$. Likewise, a service would expect to decline if exploiting and $P(j \in J | \gamma_j = \tilde{o}) \to 0$.

Each service sets its strategy (i.e., data policy) This game is zero-sum as users can only engage with one service at a time. There is imperfect information as there is no memory of past choices in the game, e.g., users may re-join a service previously determined to have exploited their data without penalty. Services know the global distribution of user strategies to observe or ignore data policies. Each service knows their opponent's last strategy but otherwise looks forward to judging their move. After a time, one service may change its strategy for no cost.

The basic logistic function, $\Delta = vx * (1 - x/k)$, is used for some key variables, chosen for its validity across many different scenarios, such as, birth and death rates, rumour or infection spread, or relevantly, industry growth (Strogatz, 2001). For example,

$$(Growth_i|s_i = RESPECT) = \big(v * Size_i\big) * \left(1 - \frac{Size_i}{Potential + Size_i}\right) \tag{5.5}$$

represents the expected growth of a respecting service in relation to current market share and the existing pool of potential users in the market; v is some small constant. This function governs other similar variables, such as the populations' likelihood to observe data policies. For instance, the uncovering of exploitation or respect strategies alters the proportion of conscientious users akin to the spread of rumours.

Figure 5.2 illustrates typical results from this *in essence*, agent-based adaption of (Vila et al., 2003). In this version, only two services are vying for users in a population of $n = 15000$. The top and middle panels indicate that the oscillations as described in 5.1 coincide with cycles of disengaged users, i.e., For instance, the central panel showing user reaction to exploitation reflects what could relate to what Westin (2003, p. 24) describes as a 'shifts in public mood', such as during 1999 and 2002 when 'privacy fundamentalists' increased from 24% to 34%. As was in the original equation-based model, the unstable dynamics in Figure 5.2 are largely down to two factors. The first is the number of users in the population that are observing exploitations, and the second is the attraction of data exploitation. These two factors are intrinsically linked, creating feedback in the system that resists a stable system outcome. Equally similar to the original model this model is sensitive to changes in ϵ, i.e., the profits from selling data, as $0 \leftarrow \epsilon$ the oscillations in favour of stable respect strategies, whereas $\epsilon \rightarrow 1$ causes exploitation strategies to dominate. What is not noted by the original authors, yet becomes intuitive under the theory set out in this work, is that each period of exploitation can have a lasting and accumulating exposure on the individual's data. Expanding this model permits the exploration of this sustainability issue.

5.2.3 Pilot Reflections and Limitations

In constructing this replication and adaption, it became evident that much expansion was required to represent the scope of identity assurance, therefore, except for reflection on the overall dynamic, this section does not delve overly deep into the

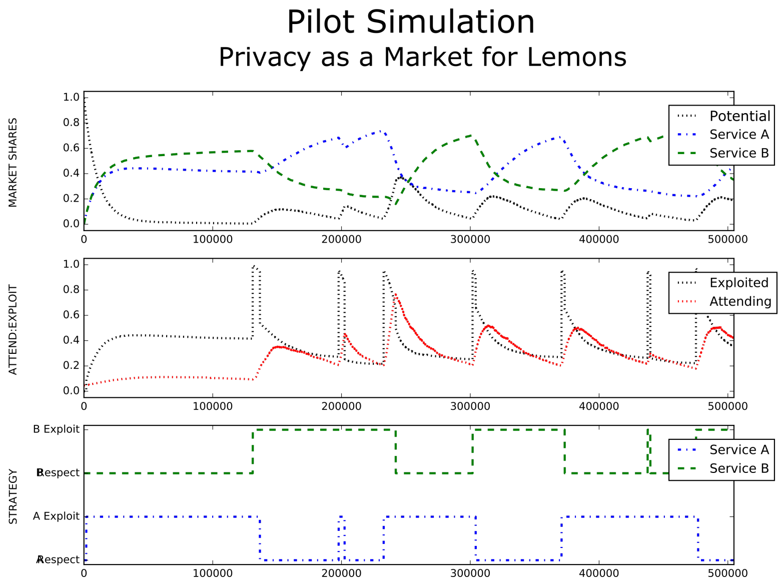


FIGURE 5.2: *Top:* Market shares and potential users. *Middle:* Periods of exploitation and the reactive user strategies. *Bottom:* Oscillation of dominant strategies as user-bases reach tipping points. *Parameters:* $r = 1, \epsilon = 0.005, n = 15000, m = 2, andT = 3000$

results of this pilot. Many of the limitations of the pilot simulation outlined in Table 5.1, cannot be addressed here. Nevertheless, the table stands as a guide for future development. Addressing the limitations represents the most significant challenge because each change reduces our understanding of how the variables interplay (Bullock, 2014). Moreover, the likelihood of errors and artefacts in the code remains beyond what can be recognised by the authors (Axelrod, 1997; Galán et al., 2009). The risk is that the design decisions are ”sufficient to generate a given effect” (Conte et al., 2012). It is for these reasons that this work does not make predictive claims, and instead provides a qualitative account of the system from which further works can emerge.

In addressing some limitations that being sought is to increase the model's realism and generality, (Barandiaran and Moreno, 2006). In the main, this is achieved through added interdependence, interconnectedness and interruptions (Macal and North, 2010; Walby, 2007). From here, inspired by the commons theme, interdependence injects realism, as regardless of any individual's actions, the strategies at play are based on an environment that, in turn, depends on group aggregates. Interconnectedness would be an obvious addition by using network heuristics to proxy for disclosure decisions as is described in Chapter 4, yet it is a step beyond the scope of this chapter. Finally, interruptions would be the malicious (or erroneous) actors from which the systems are in place to protect against, however, instead of introducing said actors, herein, the interruption is discussed in terms of risk.

TABLE 5.1: Potential Enrichments of the Pilot Model

Main Limitations	Possible Enrichment	Addressed
	Service Agents	
Only 2 services	An N-Service Game	No
Limited Pure Strategies	Mixed Extraction Strategy	Yes
	Mixed Exploit/Respect Strategy	Yes
	Diverse PII Veracity Value	Yes
Homogeneous Markets	Markets Types (Social v. Commercial)	No
	User Agents	
Limited Dependence	PII depletion aggregates (interdependent)	Yes
Disconnected Users	Network Effects (interconnectedness)	No
Limited Pure Strategies	Mixed Observe/Ignore Strategy	Yes
	Diverse PII Disclosure Value	Yes
No Memory	Past Experience Influence	No
Rational Users	Contrasting Heuristic-based Users	Yes
	Environment	
Stable system	Added threat factor (interruptions)	No
Closed System	Users can arrive and leave a market	No
	Services can fail and establish	No

Note: Some of the pilot limitations and possible enrichments

5.3 Extending the Model

The system described in the pilot in many ways is a model akin to Bertrand's model of oligopolistic competition (Vilà, 2008). Bertrand's model examines the interdependence between the pricing decisions of a set of non-cooperative rivals, where these rivals offer homogeneous goods to a population of consumers, where they have the same constant marginal cost. Likewise, these hotspots have homogeneous, perfectly substitutable, Wi-Fi functions. They experience fixed costs and marginal costs that are covered by advertisement revenue correlating with the number of current users. Hence, each service is willing to serve any amount of demand.

The main difference from the Bertrand context is that instead of setting price the services set out a data policy (strategy) that details the cost to users in term of the PII required to use the service. Essentially, since the hotspots are 'free' to use, data is the currency of exchange. This distinction makes the interaction between user and service laden with unknowns and speculation. The following sections details just why data makes a weak unit of exchange and how to enrich the model to accommodate these factors.

5.3.1 Enrichment: Data Heterogeneity and The Utilities of PII

The first consideration about data as a currency is that research repeatedly shows that user behaviour in disclosing personal data may be rational, but it is far from simple, and it rarely conforms to calculative economics (Kehr et al., 2015). The value of data is subjective, and there is no agreed unit of exchange. For example, two individuals may value their fingerprints differently, yet their home address similarly. These valuations then can differ across contexts or time and perhaps even current emotional state **?**. This is precisely the premise of Chapter **??**, which outlines a method towards the quantification of personal data while recognising individual and contextual variance.

The results in Chapter 3, Table 5.2 provide an overview of how the properties of PII can be deconstructed into different aspects. In this table, PII classifies into knowledge, physical, and behavioural metrics. Therefore, to extend the model, instead of users having a homogeneous soup of data, a user $j \in J$ beholds a personal data set $PII_j = \{K_j, B_j, H_j\}$. The individual letters $x \in s$ abbreviate as $x = K$ for knowledge, $x = B$ for biometric, and $x = H$ for behavioural. With this change, the model has the variables to enable services to adopt heterogeneous strategies beyond the binary choice to respect or exploit.

TABLE 5.2: Utility Components of PII

Type	Extraction	Veracity	Disclosure Willingness
Knowledge	Overt	Φ_K	v[0..1]
Physical	Overt	Φ_B	v[0..1]
Behavioural	Covert	Φ_H	v[0..1]

Note: This table is simply a guideline for the simulation. It is a framework to illustrate the relative veracities and spread of users' disclosure willingness. The notion of willingness to disclose is derived from the results of Chapter 3. Veracity is assumed from the literature review in Chapter 2 (See Table A.3)

The veracity variable adheres to the evidence that different identifiers provide different assurance levels. For instance, biometrics solve many confidence issues associated with the transferability of PII within knowledge systems; still, despite the advancement, non-revocability issues remaining (Jain and Nandakumar, 2012). Therefore, knowledge-based identifiers, e.g., a password, provide baseline veracity then physical biometric identifiers such as fingerprints are assigned as a progression. Likewise, behavioural biometrics, such as keystroke patterns, gain the highest veracity level as they are hardest to spoof and provide continual verification. Accordingly, services can extract an assurance value from each user, for example, Φ_{K_j} denotes the assurance provided from the knowledge of user j, and with all things being equal, $\Phi_K < \Phi_B < \Phi_H$.

Each user also has a unique willingness to disclose that particular item of data. Thus, v_{x_j} denotes the personal utility assigned to $x \in PII_j$. From Chapter 3, it is evident that as a population, there is overlap in the values placed on items of PII. In general, and agnostic of contextual variance, individuals were less willing to disclosure biometric identifiers and most willing to disclose knowledge-based identifiers, i.e., $v_{K_j} < v_{H_j} < v_{B_j}$. Thus, unlike the ordinal condition placed on the assurance value, the personal utility has no ordinal restriction. An enrichment, therefore, is that users hold a unique value of each category of PII. I discuss these values and distribution per the findings in Chapter 3 in Chapter 6.

Finally, while physical and knowledge metrics mostly require a form of overt extraction, behavioural metrics tend to be extracted covertly. For instance, knowledge, e.g., a password, has relatively low veracity as an identifier as it can be lost or stolen, or weakly constructed. In contrast, behaviours, e.g., a keystroke analyser, have relative high veracity but a trade-off is that the user may not be aware of its covert collection, and it has a higher personal utility in terms of willingness to disclose.

5.3.2 User Enrichment: Dual System Strategies

Given that identity management is deemed to be a self-managed process (Solove, 2012), it is attractive to reduce the population of users to a set of simplistic rational, calculative and economic beings. From that standpoint, one may focus purely on the competition of the services to find the lowest price, i.e., data cost. If such a price exists, then the rational free-market shall find it. This price turns out to be the break-even point; hence it is known as the Bertrand paradox or ‘Trap’ (Cabral and Villas-Boas, 2005). It is a situation in which profit-driven services are motivated to prevent, and they do so through product differentiation, loyalty schemes, value-added service, or underhand small-print tactics (Vilà, 2008). Essentially, the potential of a trap incentivises organisations to add complexity to the decision. Navigating this complexity then requires that the decision-maker invest much effort in analysing the possible decision outcomes (Culnan and Armstrong, 1999). Alternatively, and perhaps more likely given the frequency of such decisions, users depend on ’fast and frugal’ decision heuristics (Gigerenzer, 1999). Therefore, following on from the various utilities of PII that permits more ‘privacy calculus’, the results from Chapter 4 motivate an enrichment of the model to include heuristic behaviours that ignore these inherent values.

Table 5.3 provides an overview of ways to achieve this in the simulation. Along with this table, there are many nuanced heuristics. Thus, not all heuristics are currently addressed, for instance, the EXPECTANCY and INTENT heuristic classes do not readily align with an abstract social model, as there is a certain amount of subjective interpretation performed by the user upon the interface. Questions remain regarding the additive or threshold nature of data value. A situation discussed in Chapter 4 whereby once a high-value item of PII is disclosed, then lower value items tend to

TABLE 5.3: Agent Behaviour Cues for Model Enrichment

Class	Heuristic	Description	Model Behaviour
PROMINENCE	Reputation	A prestigious service would not knowingly do wrong	Years in Service versus Exploitation
	Recognition	Familiarity, even in name only	Advertisement
NETWORK	Endorsement	Recommendation from known others	Friends in Service
	Bandwagon	Recommendations or perceived actions of unknown others	Popularity [size]
RELIABILITY	Consistency	Agreement with another source or procedure	Frequency of Disclosure per identifier
	Self-confirmation	Accord with a pre-existing belief	Willingness to Disclose per identifier
NARRATIVE	Availability	The ability to recall instances	Experience of Threat and Exploitations
	Coherence	Ability to envisage experiences	Network experience of Threat and Exploitations

Note 1. The omission of some heuristics was unavoidable due to the conflict with the model abstraction level.

zero value; recall S9's perspective, "I consider my phone number a pretty private thing to begin with it, so if someone has it, they already probably know my name." Likewise, as items of PII are disclosed some users revealed a higher likelihood to disclose said data the next time in a manner representing changing norms. From here, however, the model only adopts the PROMINENCE of service as a proxy for trust. The remaining will serve well for future research endeavours.

For prominence, a user has a probability of seeing one service first over another. Therefore, they would evaluate this service first and thus, and such an organisation would gain an advantage if higher prominence over a population of users. By adding in a prominence cue, the essence of a Bertrand trap is less likely, as prominence can act as a product differentiator in an otherwise homogeneous service offering (Vilà, 2008). Instead, the model incorporates organisational prominence with a globally fixed ratio concerning each other, i.e., $A/B \rightarrow 50/50$ would mean zero prominences in the market, whereas $A/B \rightarrow 72/25$ suggests a definite prominence advantage for organisation A. While the global ratio is fixed, the individual user will have their own 'local' notion of prominence. Note, that the memory constraint still applies to mean that any aspects of consumer loyalty affecting price, as described in (Deneckere et al., 1992), do not enter user decision-making. Thus, how the following work deals with prominence is closer to the notion of recognition than that of reputation.

5.3.3 User Enrichment: Regulatory Focus

To enact a change towards heuristic agents, I incorporate the notion of regulatory focus as discussed in Higgins (1998); Higgins et al. (2020), and in Chapter 3. Users in the pilot model were consumers that would test for exploits, with some users testing and others ignoring; it was a binary state where $\gamma = \{o, \tilde{o}\} \rightarrow observing, ignoring$. Now, only some users tend to look for detail some of the time. Users have a trait likelihood that they will adopt an observing state, i.e. prevention focus, or an ignoring state, i.e., promotion focus. The likelihood of whether a user j will observe, $P(\gamma_j = o)$, the cost of a service's complete strategy, i.e., read the privacy policy, or otherwise ignore it, $P(\gamma_j = \tilde{o})$, is determined by a mixed strategy $\sigma_j : \gamma_j \rightarrow [0,1]$, where $\gamma_j = \{o, \tilde{o}\}$, and each user j assigns to each $\varsigma \in \gamma$ a probability $\sigma_j(\varsigma) \geq 0$ that it will be played, where $\sum_{\varsigma \in \gamma} \sigma_j(\varsigma) = 1$. Users then assess the cost of the transaction from the information at hand. If there are no tangible data differences in the perceived cost of the interaction, then a user will discriminate on a factor of prominence. Hence, this is a hybrid calculative-heuristics behaviour favouring the calculative. Further research could give more weight to the heuristics. Moreover, for simplicity, prominence will be assigned rather than built through actions or wealth. Thus, the results will only consider the effects of different ratios of prominence between the two services.

5.3.4 Service Enrichment: Service Strategies

Given the new composition of PII whereby $PII_j = \{K, B, H\}$, services in the model have access to more strategies. Now, to set the data cost of using the service, service i adopts a pure strategy s from a finite set S,

where

$$\begin{gathered} S = \{\{\emptyset\},\\ \{K\},\{\hat{K}\},\{B\},\{\hat{B}\},\{H\},\{\hat{H}\},\\ \{K,B\},\{\hat{K},B\},\{K,\hat{B}\},\{\hat{K},\hat{B}\},\\ \{K,H\},\{\hat{K},H\},\{K,\hat{H}\},\{\hat{K},\hat{H}\},\\ \{B,H\},\{\hat{B},H\},\{B,\hat{H}\},\{\hat{B},\hat{H}\},\\ \{K,B,H\},\{\hat{K},B,H\},\{K,\hat{B},H\},\{K,B,\hat{H}\},\\ \{\hat{K},\hat{B},H\},\{\hat{K},B,\hat{H}\},\{K,\hat{B},\hat{H}\},\{\hat{K},\hat{B},\hat{H}\}\} \end{gathered}$$

where $|S| = 27$, and S^m denotes all strategy profiles. Each strategy details the data extracted and the purpose of the extraction. The default purpose of extraction is for identity assurance, while the hat notation, e.g., $\hat{x}$, denotes the alternative use for enhancing revenue. For simplicity, whether extracted for security or revenue purposes each $x \in PII$ provides the same level of assurance, i.e., $\Phi_x = \Phi_{\hat{x}}$. A strategy $s \in S$ can contain $s_\alpha = s \cap \{K, B, H\}$ to represent the strategy's assurance use, and $s_\rho = s \cap \{\hat{K}, \hat{B}, \hat{H}\}$ to represent its revenue use. For example, for $s = K\hat{B}H$, $s_\alpha = \{K, H\}$, and $s_\rho = \{\hat{B}\}$.

A further element to this decision process is policy covertness, because, while each policy truthfully sets out the cost to a user, there are elements of S that are more overt than others. In this schema, knowledge and biometric extraction, e.g., email and fingerprint, is taken as being overt, whereas behavioural, e.g., keystroke, is considered covert. Also, secondary usage of PII to enhance revenue is considered a covert activity. Thus, without self-informing by reading the policy, the uninformed user remains aware of only the overt extraction and purpose. For example, strategy $s = \{\hat{K}, B, H\}$, from which an observing user can determine three extractions; knowledge, biometric, and behavioural, with knowledge being used for revenue purposes., whereas the ignoring user determines $\tilde{s} = \{K, B\}$ as H and the $\hat{hat}$ remain covert. See Table D.17 for all $s \to \tilde{s}$.

The utility of strategy s for service i is a function of identity assurance (α) and profit (ρ), each depending on the number of service users, such that,

$$u_i(s) = \alpha_i . \rho_i \tag{5.6}$$

where,

$$\alpha_i = \frac{1}{|J_i|} \sum_{j \in J_i} \left(\sum_{x \in s_\alpha} \Phi_{x_j} + \sum_{x \in s_\rho} \beta \Phi_{x_j} + \alpha_x \right) \tag{5.7}$$

where,

$$\rho_i = \sum_{j \in J_i} \left(\sum_{x \in s_\alpha} (1 - \beta) \Phi_{x_j} + \sum_{x \in s\rho} (2\beta \Phi_{x_j} + \epsilon) + \rho_x \right) \tag{5.8}$$

and where, α_i and ρ_i are average functions over the number of users, i.e., the mean assurance per person, and the mean profit per person. $J_i \subset J$, such that $j \in J_i$ is a user of hotspot i. Also note that, if s_ρ **and/or** $s_\alpha = \emptyset$ **then** $u_i > 0$ **if** $|J_i| > 0$, i.e., all strategies generate a small profit per user ρ_x, likewise all strategies generate a small amount of assurance α_x. In terms of profits, this would be random adverts and generic customer analysis. In terms of assurance, this would amount to some level of self-authentication, practical obscurity, and trust; risky but valid. The inclusion of ϵ in ρ_i amounts to any bias that an organisation may put of profits, this is a tunable variable in the system, that is, if ϵ is high, then this would increase the utility for strategies that on the balance of things favour exploitation. Conversely, a negative value could represent a social good for a company that wants to favour security over profits, and it could also incorporate something such as a tax on data. For instance, one suggestion is that data held should classify as a liability, and therefore storing and processing any data should incur some cost.

In this model, I also introduce a tunable variable of β that controls the balance of value extracted from an item of PII. For instance, it allows for elements of PII to be used for assurance purposes, not be exploited, but to provide some financial benefit to the beholder in terms of consumer knowledge. For example, the strategy $s = \{K\hat{B}\}$ with $\beta = 0.1$ would have $100\%K$ and $90\%$of $\sum B$ for assurance purposes, along with $10\%K$ and $20\%B$ for profit purposes, whereas $s = \{KB\}$ with $\beta = 0.1$ would have $100\%K$ and $100\%B$ for assurance purposes, along with $10\%K$ and $10\%B$ for profit purposes. Mostly, exploitation strategies are far more profitable for equal demand, especially when considering that ρ_i is accumulative. Whist is the β and ϵ values seem similar, and the difference is that β is relative to the individual user's PII value, whereas, ϵ could be a fixed benefit of cost set outside of considerations as to the user population. Consider β as an internal factor that, in theory, could be within the control of an organisation, and the ϵ factor to be extraneous, i.e., regulation cost, or nefarious gains. Services forecasting demand must first predict the cost to users in terms of the perceived amount of data extraction in return for WiFi use. Therefore, engaging the users is a process of setting a competitive strategy of PII extraction, and obtaining consent from the user for this extraction. What is competitive is a matter of beating, or matching the competition's cost from the perception of the users. Where the cost of a strategy to a user is,

$$c_j(s) = \sum_{x \in s} v_{x_j} \tag{5.9}$$

whereby, $c_j(s)$, is the is the sum of the values, regardless of use, that user j assigns to their PII extracted by a strategy, i.e., $c_j(s|s = \{K, \hat{B}, H\}) = v_{K_j} + v_{B_j} + v_{H_j}$.

Following the terminology of Vila et al., Vila et al. (2003), it is a matter of the service being 'respectful'. In said work, respectful was to not sell consumer data at all, however we soften this definition, doing so to presume that there is a general user understanding that advertisements persist in order to maintain a free service, and that organisations collect some user data to facilitate this. Yet, we hold that the general expectation is that regardless of any often obstructive or confusing privacy policy, the data extracted is not mainly to enhance revenue. Therefore, in this model it is the balance of data collection and its purpose that determines respectfulness. A service is deemed to be respecting if $r_j(s) \geq 0$, where

$$r_j(s) = \left(\sum_{x \in s_\alpha} v_{x_j} (1 - P(\gamma_j = o)) \right) - \sum_{x \in s_\rho} v_{x_j} \tag{5.10}$$

meaning that, the barrier for an organisation's data policy to be considered as being respecting increases for those individuals with a higher regulatory focus.

Once the evaluation has occurred, any ignoring user will choose the 'cheapest' service, and this is because all $r_j(\tilde{s}) \geq 0$, whereas, an observing user will choose the cheapest service if that $r_j(s) \geq 0$, else they will refrain from using any service. Thus, a user j will engage with hotspot i if they believe that this service is being respectful with their data policy and if there exits no cheaper respectful alternative. So user j's possible actions are $A = \langle \phi, 1, 2, ..., m \rangle$, where ϕ means to use no service, and where the numbers map to I, the set of hotspots.

In summary, a service in the extended model must choose a PII extraction strategy that balances the needs for high veracity with user sensitivities. As in the pilot, ignoring users will find any policy attractive. However, some policies are now more attractive than others. Observing users will find 'respectful' policies attractive, and likewise, some respectful policies are more attractive than others.

5.3.5 Enrichment: A Harvesting Model

The final enrichment stems from the intuition that every extraction of PII exposes it in some way. Therefore, with each extraction, there is a potential loss regarding the level of assurance retained by that item of PII. Rather than money that has a relatively stable unit of exchange that is controlled by a central authority, data can be subject to rapid, localised depletion. Liken the situation to the extraction of fish from a fishery, whereby, at a highly abstract level, a fishery is a function of a maximum stock level, a harvest rate, and a re-spawn rate (Hardin, 2009). Therefore, it is possible to examine this interaction within an environment that mimics the dynamics of a common resource pool.

Typically, the logistic function, $\Delta x = gx(1 - x/k)$, can represent the population dynamics of such systems, it can also apply across different yet related scenarios, such as birth and death rates, rumour or infection spread (Strogatz, 2001). Applying this to the context of PII provide a function of a rate of change of assurance,

$$\Delta(\Phi_{x_j}) = g\Phi_{x_j}(1 - \frac{\Phi_{x_j}}{max\Phi_x}) - \sum_{i \in I} E_i(\Phi_{x_j}) \tag{5.11}$$

where x_{Φ_j} represents the current assurance level Φ of user j's metric $x \in \{K, B, H\}$, and where the term $max\Phi_x$ is the maximum assurance possible from processing x such that x has had zero exposure.

The term $\Phi_{max} = K_\Phi + B_\Phi + H_\Phi$, is the total possible assurance, this provides a relative assurance level between it and the absolute assurance of a user's disclosed PII. The E term denotes the extraction of x from user j that occurs when this user engages with one of the hotspots, where $\hat{E}_i(x_\Phi) > E_i(x_\Phi)$, i.e., PII extracted to enhance revenue creates more exposure as the PII are traded and shared with third parties. The natural growth rate, g, in this situation is a constant that represents that not all exposures are long-lasting, they dissipate over time, i.e., data get corrupted, lost, forgotten, overlooked, redundant so that an exposure today may have less effect tomorrow. Nevertheless, sustained overuse of x can lead to overexposure of a user's PII, meaning that $x_{\Phi_j} \to 0$ over time.

5.4 Discussion

Composing a model of the socio-technical system of identity assurance is not a trivial task. But, by building this model, it has been a tool to navigate this opaque system. This process began with a simple mathematical model of organisations in an exploitative equilibrium of consumer data. A replication and simulation of this model (details in Chapter **??**) provide a powerful exercise in understanding how the elements of trust, exploitation, mistrust and respect can fit together. Through codifying these elements, it creates a sense of the assumption edges, and where design decisions conflict with ideals.

After the pilot replication, piece by piece, I built up a more complex model through the combination of a host of mini models that each could be cause for separate explorations. Included, are the concepts of competition and price setting in a manner inspired by Bertrand's model of oligopoly (Vilà, 2008). However, in this model, items of PII are used for the unit of exchange, this makes for a highly uncertain transaction, and one that is unstable across individual personality and within individuals' state and trait behaviours. That is, items of PII are exchanged as a form of cost to a user in exchange for secure use of a service, and organisations compete for users through their service data policies. Moreover, this model allows for covert data extractions, and this is element is something that makes identity assurance an engaging problem environment. In the typical price-setting model, the price is explicit, and both parties can see the cost, and in effect, set out to price find. Using empirical data to infer the value distributions to bring the realism of covert data extraction and potentially misleading use into the model to factor in a cost to reveal the real cost of a data policy.

Add to this that three categories can represent the unit of exchange; knowledge, biometric, and behavioural, each of which can act as a different currency to either party. Already, it is clear that this is a simplified model for what is a highly convoluted transaction. This convolution that can only serve to confuse the situation for the user, and advantage the data extractor. In classic price-setting models, the Bertrand Trap represents the inevitable price equilibrium as equal to break even as competition drives the margin down. To avoid this organisation have to introduce such activities as market differentiation and loyalty programmes ad all layers tot he transaction. In the case of data as a unit of exchange, this break-even point is already unknown, ever-changing, and or permanently beyond a users control of influence as the transaction remains opaque.

Beyond the complexities surrounding PII as a unit of exchange, the model goes on to include heuristic decision elements that stand aside from the notion of rational price finding consumer. First, those users that at that moment decide not to check policy are in some ways trusting that the real cost is acceptable and therefore not worth the price of checking. This behaviour could account for the inductive reasoning explored in Chapter 4 the found that when disclosing, people tend to go through the process with little evaluation. Likewise, the variable of prominence provides one organisation with an advantage over the other. The next chapter explores this variable to understand better if it has an inflationary effect on the price, i.e., escaping the Bertrand trap.

Finally, as set out from the outset in Chapter 1, exposure to PII as it exploited for legitimate, or nefarious tasks is a vital dynamic in this model and one that is central to the analysis in the following chapter.

5.5 Limitations

Exploring just one context of competition is an obvious limitation. WiFi is in focus due to its prevalence in our society, but also because of supplier services are perfect substitutions. Chapter 3 and Chapter 4 strongly indicate that these contextual differences are essential aspects of disclosure decisions. So, how one would extend this model to capture dynamic that also involve exchanges that are not for mediation services such as WiFi that are a means to an end, but instead services that are the end goal and therefore more purely transactional, i.e., online retail, gambling sites, amusement sites. Moreover, the subject of submission is not incorporated here, that is, the notion of being explicitly required by law to disclose the same identifiers that users disclose in other everyday circumstances. Such heterogeneous market types would be a fascinating insight into how even when some industries respect data, they are still subject to the behaviours of other sectors. In tandem, the users that may only engage with one market may see the cost to transact in one market affected by the behaviours of individuals in another market.

Memory is not incorporated. A representation memory would be an easy addition, for instance, by having a base allocation of prominence driven by a reputation cue which could be dynamic based on the individual's experience. An organisation's prominence and therefore, the likelihood of being used can increase and decrease with each interaction.

Finally, and perhaps most substantively, the lack of network effects is a limitation in this version of the model. Network effects would be the adaptation that would ground of the realism via the interconnectedness of the user population. In Chapter 4 it was evident that individuals self-regulate, and also act influenced by the behaviour of their family and peers. It is also possible to add the idea of entities entering and leaving the market to avoid an organisation, or indeed the whole system, becoming entrenched in a suboptimal state. Basing trust in this manner would however push the tractability of this model beyond hat can be achieved in the next chapter that simulates the model

5.6 Summary and Future Work

This chapter focuses on my fourth research question, which is, how to do the complex components of PII and fit together within a model of identity assurance as a commons type system. To my knowledge, this model of competition, multiple currencies types, rational and naturalistic decision making, resourcing harvesting, all within a data policy setting is a unique venture.

Too many components interact for this model to provide predictions within anything close to certainty. Instead, it is plain to say that this model must remain in the qualitative field of analysis. It is a model to tell a story of identity assurance and to discuss the broad dynamic that can inform our collective thinking about how the system comes together. This chapter combines many factors to make this a highly sophisticated research problem.

The intention is that this model can inspire a host of further empirical research endeavour to understand the limitations of the model better ratify some of the variables and parameters that are required to hone the model. To know where to add, and where to subtract. Therefore, the next challenge for this thesis is to start this process. Consequently, I now go on to simulate the model with dual-system decision agents to better understand the parameters that drive the model.

Chapter 6

Simulating Identity Assurance Systems

6.1 Introduction

The previous chapter combines the elements discussed throughout this thesis into a single model of identity assurance., e.g., dual-system disclosures, profit-driven data use, asymmetry of consent information, and market incentives for privacy. Now, in this chapter, it possible to simulate this model and garner insights into the dynamics of identity assurance systems as complex social systems. The underlying premise remains that the current self-management regime of identifier disclosure enables inevitable escalation, wilfully or ignorantly, in the amount of data disclosed by individuals to organisations. Here I explore this premise further, and I do so through the lens of complexity theory because defining identity assurance as a commons puts it into the category of complex systems. It is a complex system comprised of many component systems, each with complexities in their own right. For instance, the stakeholders of identity assurance, i.e., people, organisations, and nations, are well-established complex social systems (Mitchell, 2009), principally because of their interdependent decision-making (Acquisti et al., 2012). Likewise, the technologies of PII hold complexity in the same manner most modern technologies do, in that the techniques of identity assurance are co-dependent, and co-evolve with the behaviours of its stakeholders (Flanagin et al., 2010), including co-evolving with the actions of those with malicious intent.

This work involves agent-based social simulation, perhaps the most explicit division of simulation science that may lend to a philosophical departure, especially as it tends to affect the reduction of human agency into simple rules. Best denoted by the concept that, alongside inductive and deductive reasoning, agent-based social simulation offers an alternative generative way of reasoning (Axelrod, 1997; Epstein, 1999; Frank et al., 2009). The most pertinent of restraints, therefore, is rather than regard the agent-based social simulation method as a tool for inductive or deductive prediction (Conte et al., 2012), instead rely on it as an opaque thought experiment (Di Paolo et al., 2000). It is an opportunity to reason about the behaviours within a complex system, or simply, agent-based social simulation aids intuition to an otherwise intractable problem (Axelrod, 1997).

6.1.1 Summary of Aims

To explore, through an agent-based simulation, the components of PII disclosure that seemingly lead to a dynamic of disclosure escalation. A simulation is constructed as a useful frugal tool that permits the ethical, i.e., non-human laboratory, reasoning of potential interventions. In addition, the simulation, as a theory manifest, can illuminate the situation of the identity assurance commons as a compliment to reason alone.

6.2 Method

6.2.1 Evolution of Technologies

In this method, I simulate (using the Python programming language) the model through two levels of complexity. This path aims to mimic the natural transition between within the real world as technologies matured to enable the acquisition of novel items of PII. In level one, the set represents knowledge-type items of PII, S_K, then, level two incorporates physical biometric PII, S_{KB}, finally, S_{KBH} represents the inclusion of behavioural biometric PII. These sets expand explicitly as; level one,

$$S_K = \{\{K\}, \{\hat{K}\}\}$$

, and level two,

$$S_{KB} = \{\{K\}, \{\hat{K}\}, \{B\}, \{\hat{B}\},$$
$$\{K, B\}, \{\hat{K}, B\}, \{K, \hat{B}\}, \{\hat{K}, \hat{B}\}\}$$

I have removed the empty set, $\{\emptyset\}$, from the strategy set as described in Chapter 5. The empty set was to signify open public networks that don't require authentication. While this is a realistic part of the model when perhaps the contexts changes, in the scenario of free Wi-Fi services, it is a diminished ideal to think that no items of PII are extracted when using these services.

6.2.2 Condition Variables

In the simulation, an individual's value for a service's prominence (i.e., reputation/recognition) and their regulatory focus score are taken independently from seemingly identical distributions, Figure 6.1 and Figure 6.2. It is deemed here that total is that while prominence is between $0 \rightarrow 1$, prominence for service A is constrained between $0 \rightarrow .5$ with $\sigma = 0.1$ and mean increments of $\mu = [0, .1, .2., .3, .4, 5]$ as depicted, and prominence for service B is simply $1 - A$. Accordingly, a value of 0.5 would mean zero prominences of A over B or visa versa, whereas, a value of 0.1 for prominence would give a $10 : 90$ ratio of prominence for B over A. The six distributions depicted indicate the distribution feeding into a particular scenario of the simulation. Likewise, for regulatory focus, any user can have a trait focus score selected from distributions with $\sigma = 0.2$ and mean increments of $\mu = [0, .2, .4, .6, .8, 1]$. Yet, depending on the simulation scenario, the likelihood of what this value could be will be taking from a different distribution. Effectively, the mean regulatory focus and ratio of prominence are the main variables that capture the effect of real-world dynamics. For instance, competition restrictions, i.e., avoiding monopolies, and or in terms of education and nudge whereby users are either more likely to check or to a similar effect, the information is readily available. The two sets of distributions provide 36 different variable sets, at each level. This means that the simulation will run with these variable over both levels of strategies. At level one, this requires $36|S_K| = 144$ simulations, and $36|S_{KB}| = 2304$ simulations at level two.

6.2.3 Personal PII Variables

In Figure 6.4, you can see three different distributions, each corresponding to one type of PII. That is, rather than a particular selection of PII items, I abstract PII as the categorical sets, knowledge, biometric, and behavioural. The sampling uses truncated distributions ranging from 0...1, with standard deviation of $\sigma = sd[KBH] \rightarrow [.28, .18, .19]$, and means of $v[K, B, H] \rightarrow [.33, .72, .60]$. These values reflect the findings from Chapter 3 that place biometrics as the participant's least willing to disclose and that they were most willing to disclose knowledge based-items of PII. Note here that the actual worth to an organisation is unknown. Instead, I apply this as an ordinal progression, i.e., for security assurance, the simulation uses $\Phi[K, B, H] \rightarrow [.2, .3, .4]$. The actual ratios are not apparent.

A Delta Factor The model set out in Chapter 5 incorporates small numbers for various factors. The actual value of these numbers remains speculative, and the likelihood is that that would derive from separate distributions, however, to simplify the simulation, each is selected, randomly, as required, from a single small number distribution, delta. Figure 6.3 indicates this distribution as a highly skewed set with a

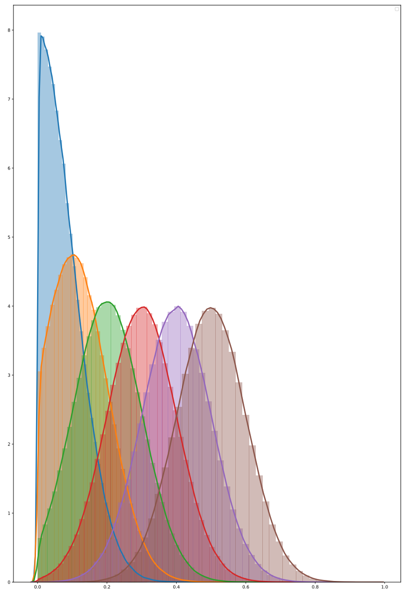


FIGURE 6.1: A distribution, truncated 0 to 0.5, of prominence for service A, with service B as 1-A. $\mu = [0, .1, .2., .3, .4, 5]$ and $\sigma = 0.1$

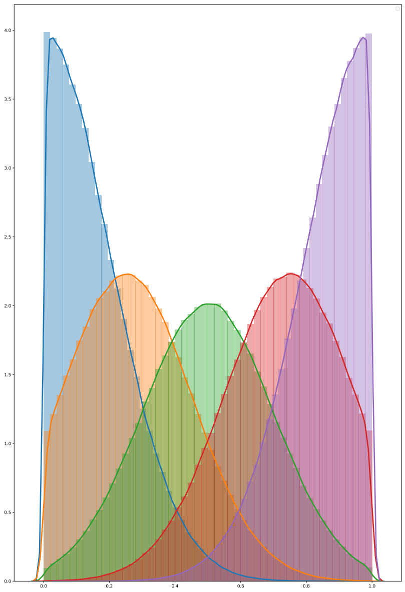


FIGURE 6.2: The regulatory focus distribution, truncated 0 to 1, indicating individual likelihood of policy checking. $\mu = [0, .2, .4, .6, .8, 1]$, and $\sigma = 0.2$

long tail, all between 0-1, albeit for this particular image max = 0.87. The first factor is a variable within the harvesting aspect of the model, EQ 5.11, which is the growth factor for dormant PII, $g > 0$. Likewise, from EQ 5.11, there is a factor that determines the exposure inflicted on an item of PII when part of a transaction, i.e. E. For simplicity, I again draw from the delta distribution, but where, $\hat{E}_i(x_\Phi) > E_i(x_\Phi)$, therefore, for $\hat{E}$, the simulation draws twice from the delta distribution and uses the sum.

Balance of Extraction In the utility function, β is a variable that indicates the balance by which an item of PII can be used for assurance and profit. Essentially, it is a tunable value, where $[0 < \beta \leqslant 1]$. In these results $\beta = 0.1$ for a 90/10 split for regular usage of PII and an extra 10% for exploitations of PII.

6.2.4 Variable Reductions

Another factor set of factors is α_x and ρ_x which as I have removed the empty set, $\{\emptyset\}$ all strategies yield an explicit level of assurance and and an explicit level of profit per user, therefore, these values are not required and can be set to zero. Finally, in the same vein, the factor that provides a hint of a bias toward or against the utility that is gained from using PII for profit over security, i.e., ϵ in (EQ 5.8) is also reduced to zero,

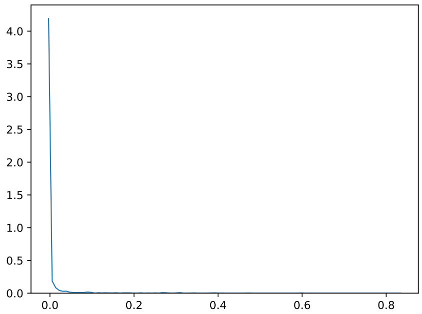


FIGURE 6.3: A delta list of small numbers with a long tail. Nobs=1000000, min max=(0.0, 0.87138), mean=0.00039, variance=0.00011, skewness=40.5, kurtosis=2030.5.

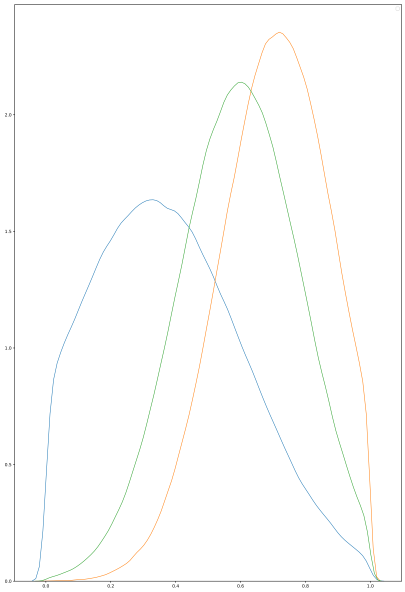


FIGURE 6.4: The distribution of how individuals value their PII. Ranging from 0...1, with means of $v[K, B, H] \rightarrow [.33, .72, .60]$, and with standard deviation of $\sigma = sd[KBH] \rightarrow [.28, .18, .19]$

and therefore in essence to future work. Understanding the effect of these numbers as coming from diverse distributions would be an interesting endeavour, yet a divergent one at this point.

6.2.5 A Reduced Model

$$u_i(s) = \alpha_i . \rho_i \tag{6.1}$$

where,

$$\alpha_i = \frac{1}{|J_i|} \sum_{j \in J_i} \left(\sum_{x \in s_\alpha} \Phi_{x_j} + \sum_{x \in s_\rho} \frac{1}{10} \Phi_{x_j} \right) \tag{6.2}$$

where,

$$\rho_i = \sum_{j \in J_i} \left(\sum_{x \in s_\alpha} \frac{9}{10} \Phi_{x_j} + \sum_{x \in s\rho} \frac{2}{10} \Phi_{x_j} \right) \tag{6.3}$$

6.3 Pseudocode of Simulation

```
for each strategy pairing $(s_i, s_j) \in SxS$ do
    for each time-step do
        for each user in random sample (with replacement) do
            observe full policy sets $(s_i, s_j)$ with P(self-regulatory focus)
            choose a $HS_p$ based on P(prominence $A/B$)
            if observing == false then
                if $HS_p$ is cheapest or equal then
                    use $HS_p$
                else
                    use $HS_{p'}$
            else
                if $HS_p$ cheapest or equal & respects then
                    use $HS_p$
                else if $HS_{p'}$ cheapest or equal & respects then
                    use $HS_{p'}$
                else if $HS_p$ respects then
                    use $HS_p$
                else if $HS_{p'}$ respects then
                    use $HS_{p'}$
                else
                    use alternative means
            if using hotspot then
                Reduce assurance for user's extracted PII as per $s$
        Account for dormant PII with potential inflation
```

6.3.1 Parameters

Alongside the above variables, the simulation uses various parameters. These parameters include **time, population, market size, and the number of runs**. Time, as with population, and market size, were a factor of what was needed to show the system in play. That is, such a complex model can always be affected in unpredictable ways by even seemingly innocuous parameters, especially those parameters at edge-cases where tipping points can occur. However, for the analysis, the parameter bounds were scanned for those that made the model perform in predictable ways. For instance, increasing population numbers provides more noise, and it can sway the profitability of the utility function as profit relies on a summation. In contrast, assurance is a calculation of the average. The population size and running time parameters combine to show near depletion of PII in a market primed to extract as much PII as possible from a population that is the least conscious about outcomes. In general, unless indicated otherwise, the simulations ran with a population of 20k, for 8k time steps, with a random sample (with replacement) of 10%of the population (market size) potentially involved at each time step. With replacement was chosen to provide a little more randomness in term of light and heavy users. The number of runs refers to the number of simulation runs for a fixed set of scenario variables. The results are an average of these runs. The variance between these runs tended to be relatively low, therefore, the number of runs, unless indicated elsewhere was c20.

6.3.2 Scenarios

The results will refer to four scenarios, Table 6.1, to make a narrative sense of the data by relating it to the real world. In scenario one, labelled as **the real world**, the variables follow the overriding evidence that points to a population comprising low regulatory focus. In this scenario, a capitalistic make-up of large versus small organisations represents what we see in the telecom market in the UK. So when referring to the real world, the regulatory focus of the population will be a truncated distributed between 0...1 with a mean of $\sigma = 0.20$, likewise the prominence of organisation A over organisation B will be a 20/80 split. In scenario two, the focus is on the **educated user**. In this scenario, the population's regulatory focus will be a truncated distributed between 0...1 with $\mu = 0.8$, the prominence remains as the real world scenario as an 80/20 split. For scenario three, termed loosely as the **indifferent user**, the regulatory focus is as in the real world situation. Yet, the prominence ratio becomes a 50/50 split to in effect to disembody both services. Finally, scenario four is the perceived **holy grail**. In this world, the educated users and indifferent users markets collide in what on paper would be a utopia for privacy advocatesplenty of market choice, and high coherence of data policy. Qualitative changes in the data between these will indicate where such interventions would be advantageous, and perhaps reveal unintuitive outcomes.

TABLE 6.1: Variables for Four Scenarios in Focus

	Indifference	
Regulatory Focus	Low	High
High	Educated User	Holy Grail
Low	Real World	Indifferent User

Note: The four quadrants of behaviour and intervention. Real World [F0.2 P80/20] Educated User [F0.8 P80/20], Indifferent user [F0.2 P50/50] , Holy Grail [F0.8 P50/50]

6.4 Results

The simulations give the raw data, so the next step for the analysis was to examine which of the strategy pairings could be dominant. To achieve this, I use the Nashpy program (Nisan, 2015) to examine the utility matrix for each set of variables. For each level and cosponsoring set of plays, each service has a utility, i.e., $u_i(s) = \alpha_i.\rho_i$. The program uses a choice of one of two algorithms to find equilibriums in the matrices, support enumeration or Lemke Howson algorithm. In this analysis, the support enumeration option primarily seeks equilibrium and the Lemke Howson algorithm only where the former fails to find an equilibrium. Nisan (2015) notes that this fallback algorithm is not guaranteed to find all equilibriums, and this should be accounted for during any potential replication-extension to this simulation. Where and if multiply equilibriums exist, I take the equilibrium that results in the total combination of utility. Thus the chosen equilibrium is optimal from a collective organisation standpoint, and not necessarily the individual, nor the global system of identity assurance.

The way the results present follows the two levels of simulation, $S_K and S_{KB}$. For each level, I present five varieties of data tables, and each of these tables spans two time-frames. The first is a table showing 36 equilibriums from the $6x6$ combination of regulatory focus and prominence variables. The second set of data tables are the utility matrices for the four scenarios in focus. The third data table contains the breakdown of the environment variables within the four scenarios, i.e. market share, or profit. In terms of time frame, the results fall into two camps, and the first camp is the long-sighted outcome if the services are to truly consider the depletion effect on PII usage and work backwards to set a policy today. The second camp is the short-sighted outcome which would come to pass if the perception and consequently, the calculation of the value of PII considers it as an infinite resource, i.e., disregarding the theory of depletion. The two time-frames colliding provides an overview of what may come to pass if strategy setters do not adopt the message of a shared resource.

The full equilibrium results, Table D.1 and Table **??** read as follows. The first column indicates the four quadrants in focus, the real-world (RW), the educated user (EU), the indifferent user (IU), and the holy grail(). A star* within this column indicates a policy difference between the two equilibriums; long and short-sighted. RF and ProA are the means of the respective distribution for population regulatory focus, and the prominence of Service A. KP refers to the end potential of PII in the system, in this case, PII is knowledge. EQ contains the equilibrium strategies for each set of conditions, along with the probability of playing that strategy, i.e. $(\hat{K}, K, 1.0)$ is a pure strategy set of service A exploiting and service B respecting. KP2 and EQ2 are the short-sighted system results.

For an example of the time frame effect on strategy, examine the two utility matrices in Table 6.2 that arise from the strategy set, $S_K = \{\{K\}, \{\hat{K}\}\}$. The table is calculated for a population of $p20000$, with the likelihood that any particular user observing a policy, $P(\gamma_j = o)$ drawn from a distribution with $\mu = 0, \sigma = 0.2$, along with a user's regard for the prominence of organisation A drawn from a truncated distribution of $\mu = 0$ and $\sigma = 0.1$. The left matrix is the long-term time frame where PII deletes, with use, over time, whereas, the right table is the short-term, zero depletion, time-frame. The pair of utility matrices are,

TABLE 6.2: S_K UTILITY MATRIX

		Service B			Service B	
		K	$\hat{K}$		K	$\hat{K}$
Service A	K	0.46, 5.36	0.27, 1.43	K	0.6, 7.48	**1.74, 11.23**
	$\hat{K}$	**0.48, 4.37**	0.12, 1.35	$\hat{K}$	0.98, 7.68	1.02, 10.83
		Long-Sighted			Short-Sighted	

$p = 20000$, steps of $s = [8000, 1]$, using strategy set one, $S_K = \{\{K\}, \{\hat{K}\}\}$, the likelihood that any particular user will observe a policy, $P(\gamma_j = o)$ drawn from a distribution with $\mu = 0, \sigma = 0.2$ and a users' regard for the prominence of organisation A drawn from a truncated distribution of $\mu = 0$ and $\sigma = 0.1$. **Bold font signifies an equilibrium**.

In the long-sighted version of this world, where very few people check the data policy of a set of services, the more prominent organisation foresees the depletion of PII and therefore plays the respectful and safe strategy. This strategy equals or dominates its counterpart service. Accordingly, the counterpart capitalises on the small population that encounters it first and don't differentiate the policies. Conversely, if short-sighted and disregarding of the depletion of PII, the prominent service is best targeting the people that fail to observe. In contrast, the less prominent service differentiates by respecting and attracting the user that do observe. A problem arises when the depletion is disregarded in the forecasting but is apparent in the system, because, then services are incentivised to extract more, and the value depletion quickens. This example, although not in the four scenarios of focus, is plausibly a version of the world that some markets could experience this profile of user behaviour, e.g., the games industry.

6.4.1 Output Tables

To aid the flow towards the discussion wherein the results are surfaced, the longer result output tables are located in in Chapter D. The tabular results for the level one, 'just knowledge', simulations are as follows. The equilibrium output is in Table D.1, the four scenario utility matrices, for both long and short-term outcomes, in Table D.2, Table D.3, Table D.4, and Table D.5. Then Table D.6 holds the supplementary environment outputs.

Likewise, the output tables for level two, 'knowledge and physical biometrics', are as follows. The equilibrium output is in Table D.1, the four scenario utility matrices in Table D.2, Table D.3, Table D.4, and Table D.5. Then Table D.6 holds the supplementary environment outputs.

Table D.16 for the equilibrium output, then Table D.8, Table D.9, Table D.10, and Table D.11 for the four utility matrices, followed by the environment outputs in Table D.7.

6.5 Discussion

6.5.1 Identity assurance with Just Knowledge

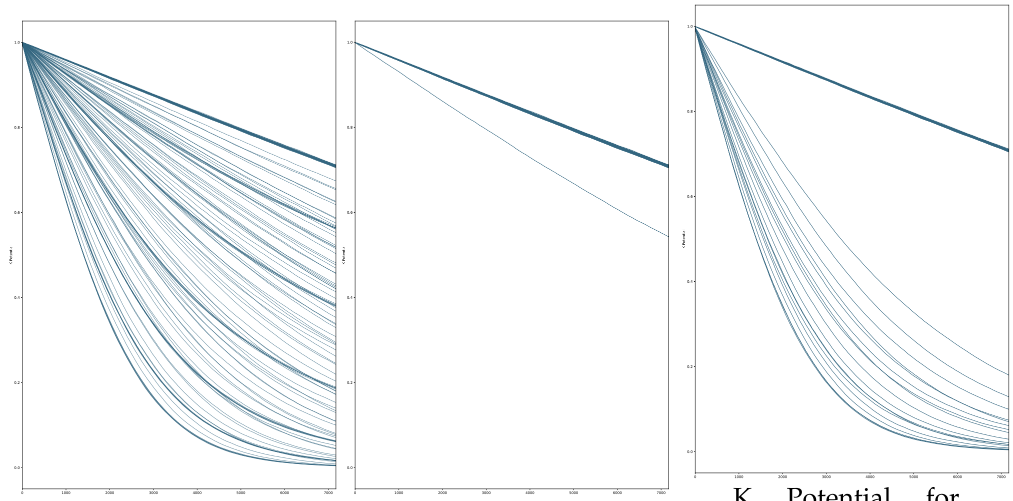
Figure 6.5 illustrates the path of depletion for K in the simulation when using the strategy set $S_K = \{\{K\}, \{\hat{K}\}\}$. From left to right, the first image is the shape of the system for each of the 144 strategy parings, that is $6x6x4$ pairs. The simulation ends where the bottom pairing reaches near zero. The middle and right figures show the value of K within the equilibrium pairs as detailed in Table D.1. As the table suggests, the central figure displays a section of strategies that are respectful in that they focus on security only, with one outlier (discussed earlier, see Table 6.2). In the right image, however, that of the short-sighted version of this 'just knowledge' world, illustrates the impact that choosing a strategy based on immediate gains is a poor outcome for the community value of K.

While possible to determine from Table 6.2, in Figure 6.6 the factor adding incentive to exploitative strategies is apparent. Overwhelmingly, and intuitively, the right image shows that the population with lower regulatory focus permits the least respectful strategies. However, beyond a population's regulatory focus as a critical factor in the sustainability of PII, perhaps most pertinent aspect is co-operation. As seen in the left image, is ensuring that policy setters are fully conscious of the potential depletion of PII as it in their self-interest and the community's to respect data.

6.5.2 Four Scenarios of S_K

In Table D.1, left of the divide is the world with depletion forecasting. In this world, there is consistency, meaning that among the rest, the four variable-based scenarios singled out are equal. The only potentially notable result that they are all the same and that they are all respecting. Right of the divide tells a different story. Just as Figure 6.6

FIGURE 6.5: Knowledge Depletion and Equilibriums



K Potential for all strategy parings ($N = 144$). This figure illustrates the market bounds.

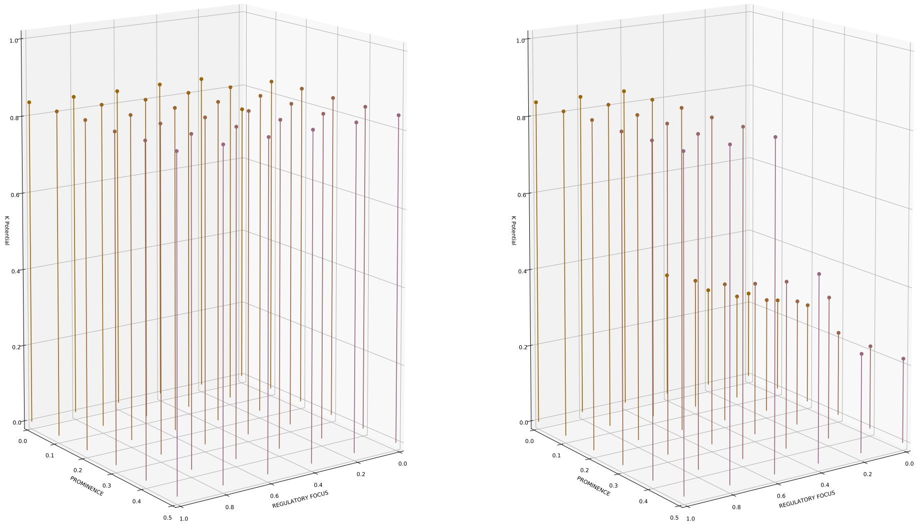
K Potential for Equilibrium parings ($N = 36$) over time in a long-sighted market

K Potential for Equilibrium parings ($N = 36$) over time in a short-sighted market

has shown, regulatory focus plays a key roll in the strategy of the services they disregard the depletion of PII. Intuitively, the higher the attention of individuals, the better behaved the services are. Therefore, both the educated user, and the holy grail scenarios have respectful policy sets, and PII remains sustainable. It stands to reason then that in the real world scenario, with low focus and low competition in the market, services sway towards exploitation. In this RW scenario, the more prominent service chooses the $\hat{K}$ policy while the less prominent service plays safely. It turns out that both services are better off than in the previously worked example Table 6.2. Sightly more focus and increased competition benefit the services, and fortunately, it also has a very slight improvement for PII depletion. Nevertheless, as detailed in Table D.6 the version of the world where services consider depletion as real is the optimal situation for all parties as there is less utility at first, but it is sustainable.

Three of the 36 scenarios lead to a policy pairing of $\hat{K}$ and $\hat{K}$. Included in these is the variable set designated as containing the indifferent users, i.e., the users that are not swayed by any cues of prominence. This result is cause for pause as one would imagine that taking potentially false signals away from the decision would take away the advantage. In the real world scenario that leads to the prominent organisation exploiting, instead of the balance only serves to have both services to co-exist equally sharing exploitation spoils, and therefore, it is the worst-case scenario for PII. Furthermore, with disregarded depletion, the indifferent user scenario is the only scenario whereby an equilibrium exists that involves only 75% of the market. In essence, the outcome is that it is better to exploit most of the market than cater to all.

FIGURE 6.6: 3D plot of K Potential Post Exposure



K Potential when the market does believe in the decay of K assurance through use

K Potential when the market does not believe in the decay of K assurance through use.

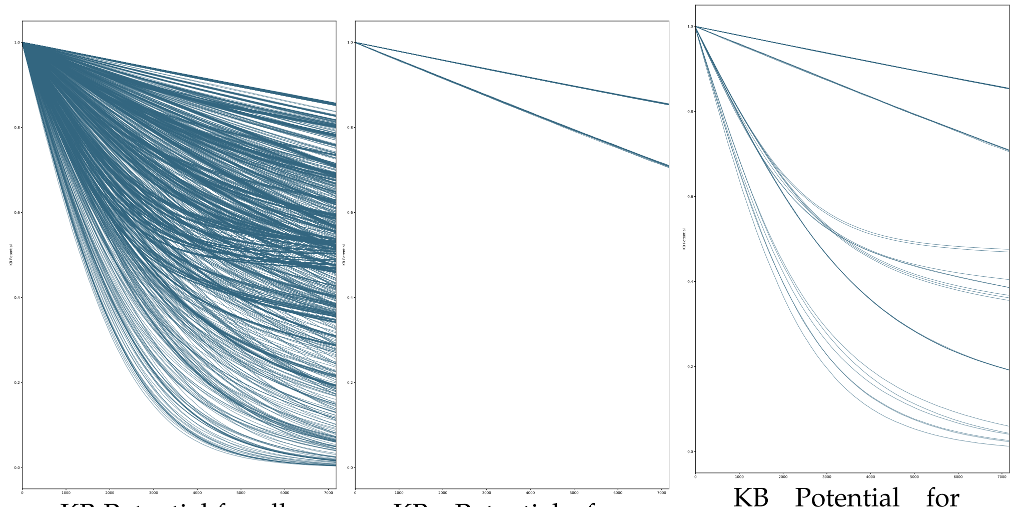
Note: Prominence is along the left axis, regulatory focus along the right, and the vertical z axis indicates the remaining potential, or confidence, for a knowledge-based item of PII to asset an individual's identity.

6.5.3 Adding Physical Biometrics

By adding physical biometrics into the simulation, the results, as one would expect, become complex. This addition provides conflict in terms of what a user may value highly compared to what might be valuable to an organisation. The most significant difference in the results is the lack of pure strategy. Instead, more than half outcomes are mixed strategy equilibriums. This indication is also in Table D.16 and Table D.7, along with the frequency distribution (if higher than 10%) for each strategy pairing. The overall shape of the result is similar to those in the version S_K. However, as many strategies do not use both types of PII, many of the downwards curves flatten in Figure 6.7. The split in the central image is the difference between using a mixed KB strategy verses a pure KB strategy which has the steeper gradient). Even with the added complexity of policy, the trend of short-sighted exploitation versus long-sighted respect continues.

Figure 6.8 provides more insight into what each of the two factors contributes to the system dynamics. In the 'just knowledge' world, it was mainly the population's

FIGURE 6.7: S_{KB} Knowledge and Biometric Depletion and Equilibriums



KB Potential for all strategy parings ($N = 144$). This figure illustrates the market bounds.

KB Potential for Equilibrium parings ($N = 36$) over time in a long-sighted market

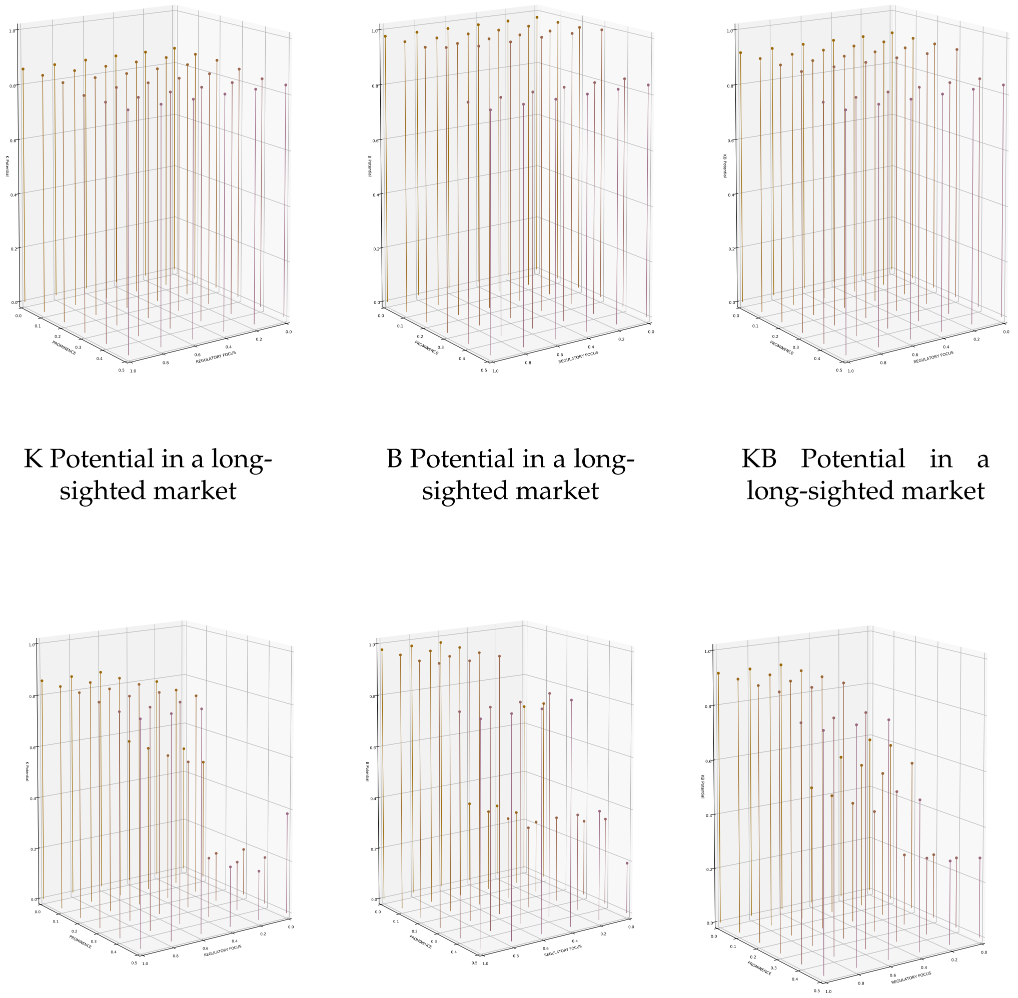
KB Potential for Equilibrium parings ($N = 36$) over time in a short-sighted market

regulatory focus that impacted the systems towards exploitation. Of the top three images, the left image shows a steady depletion of K across all variables in the long-sighted market. In the top central figure, however, there is a step-change in how the data forms. In the long-sighted market, at least in terms of B potential, it is the prominence and not regulatory focus that correlates with this transition. Seemingly, the greater equality of prominence in the market, the more each service can pursue the data. To relate to the real world, this is a like services operating in a comfort zone, neither one needing to carve out a niche market by taking less to attract marginal consumers. Nevertheless, with foresight, comes respect in the market. The extra data extracted, although more costly, is allocated to keeping the system secure, and therefore, has near long term stability in the model. In the short-sighted market, the dynamic is a far more fluid. In this version of the model where there is a greater tendency to take and exploit data. Earlier it was shown that with low focus, similar comparative prominence in the market could push exploitation further than low focus alone. Now, in Figure 6.7 it appears as if the prominence has an impact on strategy even in a highly focused population. This effect is true for knowledge potential, but seemingly more so for biometric potential.

6.5.4 Four Scenarios of S_{KB}

To explore this version of the model further, I return to the four contrasting scenarios. In the complex duel currency world of knowledge and biometric, the four scenarios

FIGURE 6.8: S_{KB} 3D plots of KB Potential Post Exposure



K Potential in a long-sighted market

B Potential in a long-sighted market

KB Potential in a long-sighted market

K Potential in a short-sighted market

B Potential in a short-sighted market

KB Potential in a short-sighted market

Note: Prominence is along the left axis, regulatory focus along the right, and the vertical z axis indicates the remaining potential, or confidence, for a knowledge-based item of PII to asset an individual's identity.

remain divided Table 6.3. consistent with the previous information, the major division between the blinkered, short-sight strategies and the ones with foresight. Other divisions occur in the real world and the indifference scenario. As with the S_K results, the educated user scenario continues to be a stable set of respecting strategies, the same applies to the holy grail scenario. A principle difference comes from the indifferent user sets. When only knowledge as traded, the high indifference led to exploitative strategies that left 25% of the market unfulfilled. However, due to the heuristics of respect outlines in Chapter 5, adding biometrics into the transaction mix gives a service opportunity to take more, but even with taking more the strategy is considered as respectful. This outcome is one of the situations, where there are two pure strategy equilibriums, with the services switching strategy as the other solution.

From these results, one may assume that the combination of low focus and high indifference is the worst-case scenario for the long term potential of PII in the hands of shorts sighted strategy. The external validity conflict here is that if indifference was high, it is because the competition is high, or there is a saturated market of equal good. Yet, in a high competition marketplace, the price drops, whereas here, this model tells the opposite story. The assumption rising from this dynamic is that relying on free competition is perhaps not ideal when heuristics are involved in the decision making as a proxy for trust than it is just too easy for that trust to breed complacent and that complacency leads to exploitation.

The range in the four findings aligns neatly with the original model from which this work emerged (Vila et al., 2003) wherein the respect cycle ebbs and flows with the attention of users. Consider an individual with states form of a central trait of how to make disclosure decisions, now presume that along with the state fluctuating, the trait was also not an immutable value. Then users would adapt, and the organisations would have to predict these adaptions, leading to the ebbs and flows or the original free-riding problem. Our data tells the same story, albeit in a higher abstraction. Suppose that the adaptation of state and trait for each individual moved between the four scenarios rather than a fixed population. from this standpoint you can envisage the bounds of where this system would also ebb and flow.

Notwithstanding that it is notoriously difficult to achieve, it is the regulatory focus of individuals that could best peg the market into a sustainable system of PII use. Alternatively, and perhaps most likely to garner success, is the notion that depleting of PII in the long term lowers the utility for all. All data from this model points to co-operation before it is too late as the optimal solution. However, such co-operation is not readily emergent in the data extraction world. Therefore, instinct hints at external regulation, including mechanisms to punish defectors from the optimal community solution. These are all well-documented strategies in the realm of game theory and would be the logical next steps of the model's development.

One form of regulation would be to add transactional costs to the services that wish to extract such PII from there the prediction is that services will be more refrained in their extraction unless the need warrants the cost. In essence, that means changing the asset definition of personal data, from being a liability. The model in Chapter 5 includes a way for this adaptation. It is possible to introduce liability for items of PII in the utility equation using ϵ in the equations. For simplicity, ϵ is held zero value in this chapter. However, A positive value for ϵ could represent a licence cost to process certain items of PII. The decision, however, was to reduce this variable to zero to contain the outcome. The simulation and model depend on a number of these variables meaning that for predictive circumstance would require great validation (Axelrod, 1997). Like the ϵ variable being reduced to zero, the model and simulation made other compromises and simplifications.

TABLE 6.3: S_{KB}, Variables for Four Scenarios in Focus

	Indifference		
Regulatory Focus	Low	High	Time Frame
	Educated User	Holy Grail	
High	(K, K)(B, K)(K, B)(B, B)	(KB, KB)	Foresight
	(K, K)(B, K)(K, B)(B, B)	(KB, KB)	Blinkered
	Real World	Indifferent User	
Low	(K, K)(B, K)(K, B)(B, B)	(KB, KB)	Foresight
	$(\hat{B},\hat{B})(\hat{B},\hat{K}B)(\hat{K}B,\hat{B})(\hat{K}B,\hat{K}B)$	$(\hat{K}B,\hat{K}\hat{B})$	Blinkered

Note: The four quadrants of behaviour and intervention. Either to educate the users, or to regulate the market. Real World [F0.2 P80/20] Educated User [F0.8 P80/20], Indifferent user [F0.2 P50/50] , Holy Grail [F0.8 P50/50]

6.6 Conclusion

6.6.1 Summary

Users face a diminishing choice between non-disclosure or non-engagement (Staddon et al., 2012). Despite what the laws aim to achieve, the idea that user can opt-out of data extracting services is fast become an archaic one as these systems creep further into the unconsciousness. The complexity of the data extraction ecosystem warrants that the rational thing to do is use heuristics to navigate the uncertainty. Even if possible, consistent, reasonable, informed consent is not plausible. The simulation work in this chapter builds upon the model outlines in Chapter 5. The aim was to combine the components of identity assurance and do so through the lens of complexity theory. To my mind, this work successfully achieves this aim. I have conducted a simulation that explores the role of individual attention on data policies, couples with the rolls of prominence as a heuristic of trust. I have shown that added complexity in the transaction for PII disclosure can lead to greater overall data disclosure. Moreover, the results show that short-term decisions in a two-player strategy game lead to higher extraction of PII, and within exploitation. However, the result here also suggests that with long term planning and foresight, the optimal solution is one of cooperation with cooperation, regulated or emergent, all parties benefit.

Absent of a desire to regulate, solutions within the sphere of the disclosure are to education and nudge. The process by which a user becomes more aware of the outcomes of data disclosure remains a hot topic of research. Primarily, if the goals are influencing the transparency of PII usage, one can approach this goal in terms of more natural to understand policies with salient information surfaced, alternatively, educating users to be more willing to read them in the first place. In the absence of education and straightforward policies, then behavioural nudges can steer a higher proportion of user towards decisions that are in their better interest. What these ′better′ interests are remains outside the remit here, especially in the complex realm of identity assurance where some amount of disclosure is required.

6.6.2 Limitations and Future Work

I believe that this work finds a solid middle ground between the stem mathematical model and an all-encompassing network of social agents with feedback loops. Despite using an agent approach, the model remains close to an analytical solution. There are no network effects that are common in social agent-based models, nor is there animation. Such components are surplus in this cost finding simulation. Likewise, the use of fixed policy pairings to produce 144 different simulations is a simplification in terms of reduced randomness. The pilot detailed in the Appendix **??** used dynamic strategy behaviours from the services. And while that proved useful in terms of how simulation may progress, the outcome is more akin to a random walk. While a random walk through some of the variable of this system would be valuable, the approach taken here abstracts eh ebbs and flows while also showing the edge cases of what may come to pass. The control of this approach is its greatest asset but also its most significant abstraction from the messy real world.

The extend of validation, therefore, regards coherence around the outcomes and face validity Klügl (2008). The inputs for many of the variable come directly from empirical research, and the outputs compare to the findings of the **?** model. Moreover, the analogy approach of the four models provides a level of behavioural validity. The four scenarios that reflect four versions of the world, as it is, and as it could come to be offer narrative checks to which to compare expectations. The results behave coherently. They show the long term optimal solution, the bounds of regulatory focus that warrant services respecting data, and conversely, the levels that permit exploitation. From the outset, the intent was that this non-reductionist approach to simulating the system of identity assurance was for qualitative purposes and not a prediction. This intention still stands. There can be no claim that increasing policy transparency by X will produce behaviour Y. Instead, only applications such as increasing policy transparency is perhaps better than relying on the free market to find an acceptable price. This may not satisfy all stakeholders; nonetheless, it adds to the growing consensus that informed consent is not a form of control for the user when there is so much uncertainty.

The sentiment in this chapter regarding the cooperation of entity for the global good, may not be shared by all parties. Entities that act in a manner disregarding the subjective nature of personal data, and or its unusual depletion quality, may seek short-term gains. These entities benefit from avoiding the conditions of classic price-setting problems, and thus benefit from convoluted decision-making and thus avoiding the Bertrand Trap where the break-even is the equilibrium (Cabral and Villas-Boas, 2005).

Given this work highlights the frugal nature of heuristics and the impoverished decisions that user make when disclosing, it seems unlikely that a break-even or true cost for the transaction could be known. However, advances in the research of negotiating smart agents promise to aid user decisions. Therefore, there remains optimism that despite the recommendation here being that informed consent is beyond reasonable expectation of users and that the self-management of PII s therefore defunct, long term solutions should be person-centric. That is, people have a fundamental right to privacy, so any detraction from self-management perhaps should be under a specific time frame. What is clear, is that left as individual versus the organisation, the power dynamic will always favour the organisation, and if we are to believe that organisations are rational, then the user, and therefore the sustainability of PII is at risk.

Chapter 7

Conclusion

This work argues for treating PII as depleting resources and examining identity assurance as a commons type environment. The existing literature around identity assurance points to persistence in the exploitation of PII by organisations (Vila et al., 2003). Also, evidence points to persistence in the malicious behaviour around identity assurance (Anderson et al., 2014; Kahn and Roberds, 2008). These persistences combine towards an understandable escalation in PII extraction. This view then led to a novel perspective on identity assurance that could have significant implications for the current self-managed regulation approach of PII disclosure within identity assurance. The overall suggestion is that this approach, coupled with the various incentives and strategies for PII exchange, is causing a depletion of the identifying value of PII. Then with that depletion comes a natural escalation in extracting higher veracity PII of greater volume and variety. This work develops this argument through a mix of literature-based analysis and empirical studies.

The aim was to demonstrate depletion of PII's identification value from seemingly harmless yet self-interested, short-term, individual decisions within a system of competing incentives. Combining literature with empirical work into a model of identity assurance as a social system made it possible to simulate the model dynamics via an agent-based representation. In examining the model in motion this way, it was apparent that user behaviour led to a more comprehensive sharing of PII regardless of justification. Agents in the model design, inspired by the empirical work, were too often unconscious to the full disclosure through disinterested self-regulation or reliance on prominence-based heuristics. Moreover, viewing identity assurance as a public good, it is doubtful to presume that what could be reasonable for some informed individuals should be expected of the population.

Therefore, in this representation, it was organisational reasoning and foresight about the long-term depletion of PII's identifying utility that, in theory, leads to a better outcome, encouraging sustainable PII exploitation. This outcome held through four broad scenarios that outlined parameters of the model according to potential interventions. In doing this research, I commit work that calls for a rethink of PII regulation and, ultimately, data processing management in identity assurance. Assured and consistent, user-centric, self-managed PII disclosure is not plausible in our current digital world.

7.1 Overview of Investigatory Work

Chapter 1 introduced the outline of a simple game-theoretic model from Vila et al. (2003), inspired by Akerlof (1970). This outline provided the platform for the incorporation of my empirical findings to enrich this model by using the results from Chapter 3 as a form of disclosure measurement, and those from Chapter 4 as the agent decision rules. For the remaining RQ4, Chapter 5 and 6 represent and simulate these findings within a novel commons-based perspective of identity assurance.

In Chapter 3, I address the question of how to quantify the personal utility of PII (RQ1) when in practice user behaviour betrays intention. Given that users are acting in the face of highly uncertain and complicated decisions, perhaps it is essential to respect a user's intentions over their actions. With the right to erasure laws coming into force internationally, it seems that this is the prevailing sentiment about personal data and how reflection can play an important part in the comfort of a disclosure.

In a novel application of the ELO ranking algorithm, I detail a frugal and scalable method of capturing and combining some of these sensitivities. By conducting this method across three contexts, namely, submitting to government requests, transacting with commercial organisations, or self-expression on social media, this investigation addressed RQ3.

The results of this work indicate how, despite individual differences, as a cohort, the utility of PII forms clusters that indicate distinct levels of personal value. The results also have distinct intra-context and inter-context dependencies with some clusters that users seem relatively willing to disclose but suggesting others that need more protection as they represent a high personal value. This result means that technologies that treat PII as one amorphous group, and those transferring PII across contexts, risk failing to adhere to the sensitivities of the user. Each time an item of PII is transferred and or inferred upon it corrupts the original implied contract between the user and data controller. For instance, users could disclose in arenas of high willingness to reveal when the future use resides in another sphere of control.

These findings indicate that, in an ideal world, organisations would have to justify and bear some cost for the use of higher value PII regardless of consent. Also, by working with these cohort-based clusters in mind, it is plausible that system designers and policymakers may better appropriate system needs with the wants of the individual.

In Chapter 4, I explore how users are making disclosure decisions (RQ2). I show that users evaluate trustworthiness from heuristics. These results were recognised in super-ordinate groups (Table 4.1), revealing PROMINENCE, NETWORK, RELIABILITY, ACCORDANCE, NARRATIVE, MODALITY and a seventh non-heuristics TRADE class.

It seems that prominence and social networks that provide cues such as popularity and brand exposure can replace attempts to self-inform. But, in essence, this describes somewhat of a herd mentality Equally, trustworthiness stems from accordance with beliefs and a sense of reliability. Cues such as familiarity and regularity proving the information proxy This manner of reasoning has inductive blind spots. It boils down to the notion that if nothing negative occurred before then, the trend would continue.

The resulting behaviour favours disclosure, because, in disclosure environments, there are many trust-based cues yet scarce information about the risks. On occasions when users attempt a, considered, self-informed approach to disclosure, qualitative accounts of adverse outcomes remain obscure, leading to incomplete, incoherent, imbalanced narratives. This obscurity, therefore, diminishes our ability to appropriately conduct the 'privacy calculus' (Krasnova et al., 2010) required for informed consent. The main implication is that the self-managed model, whereby self-informed individuals are responsible for consenting or withholding personal data, is idealistic, as they instead tend to make impoverished heuristics-based decisions. The results here agree with Sundar et al. (2013) and Gambino et al. (2016) that explain the why disclosure outpaces people intentions to disclose, i.e., privacy paradox. Likewise, the results demonstrate why disclosure outpaces expectations of exposure, i.e., simple consent.

Chapter 5 revisits the pilot model that preceded the empirical work to merge what is now known with what was known. The pilot fused notions of rational market consumers with rational corporate entities, showing that the default state exploits consumers. Learning from the results in Chapters 3 and 4, it became prudent to consider that these consumers rely on heuristics over rational ideals. PII has a fluid value mapping that expels notions of PII as a standard currency exchange. Moreover, this fluid mapping conducted within informational asymmetry mediums of exchange, many of which obfuscated the exchange, further distorts the consumer decision. This obfuscated complexity leads to the theory of heuristic decision-making and how disclosure processes are one of trust as much as risk.

Alongside the empirical findings brought forward from Chapters 3 and 4, our model construction incorporates features of the commons theme that runs through this work. PII in the model has a decay feature through a harvesting element to illustrate that each use brings risk, something not valid with a one-shot currency transaction. This harvesting feature extends the idea of exploitation and permits reasoning on the sustainability of PII in a system of exploitation. Unfortunately, as these features were lacking in the literature, building the model was as much an end goal as the means to an end. More time and analysis remains to explore the system. It was a backwards for progress situation. By establishing these features into codifiable features of a model, future models can borrow this firmer ground.

Together, the construction of this model from distinct theories and sub-models draws us to a particular portal from which to see the problem environment (RQ4). In reduced models, it is often the search for quantitative outputs and perhaps even a golden lever to transform controllable inputs towards favourable outcomes. Instead, in permitting a model of complex layers, the notion of a golden lever is moot. It is the direction of change or the extent of change in the outputs occurring from different environmental stimuli that point upstream to a potentially pivotal interaction of elements.

The key work in Chapter 6 is simulating the model of identity assurance. This simulation helps to ratify the component led construction of the model in Chapter 5 to further understand how the utilities of PII and decisions making approaches of individuals combine (RQ4). The simulation is conditioned through the lens of four situational schemes, each representing a caricature of the world, however plausible, if certain conditions held. The scenarios are real-world (low indifference and low regulatory focus), highly educated (high focus), highly indifferent, and a holy grail environment of high focus and high indifference. On the surface, the populations' regulatory focus may be a pivotal element towards keeping data extraction to the required as in the Bertrand model towards marginal cost or break-even. However, it is inconceivable in practice to motivate a population of users to consistently apply energy (time, money, attention) to a complex decision that returns during most technological interactions. It is easier for individuals to adopt a heuristic-based trust model. Such a model would let the individual focus on peer behaviours as a form of social proof or induce outcomes from what has since occurred. That is, this is easier, at least in the absence of any technical delegation.

Zooming out and what seems to be more promising is that enhancing the organisational understanding of the value of PII as a public good would be the parameter to explore. With a longer-term view onto the depletion of PII, organisations in the simulation took more PII, but exploited it less. Perhaps this win-win scenario is plausible. If corporate responsibility is implausible to policy influencers, then legislation and restriction of data through liabilities and stricter repercussion seem to be the straightest road towards sustainable use of PII.

7.2 Contribution

Essentially, this work combines quantitative, qualitative methods to provide extra support to the generative simulation findings. Each of the empirical and simulation experiments has distinct contributions, yet the cumulative contribution is less obvious. Firstly, it is worth pointing to the multi-method approach. This approach is a template for others, perhaps those more inclined to stay in the coding and simulation realm of complexity science, to step into the qualitative research paradigm rather than relying solely on normative assumptions which seems to be the template. It is far from a comfortable experience. The temptation to create parameters from educated guesswork is high. However, the satisfaction of building agents with the voices and sentiments gained from the interview and transcribing process is an immeasurable positive. It was a means to self-check bias that often plagues abstract modelling.

A more overt contribution is the complexity itself, that is, resisting the reduction of a complex identity assurance system into a simple input-output model. The result means some loose edges and some compromises in rigour, but this is a simulation towards a qualitative output. It is a theory manifest. It represents one perspective, and albeit less so than the world it describes, that perspective is chaotic. Therefore, it is a work that edges the lines between the expectations of reductionist science, complex systems, and the pseudoscientific romance of coherent artificial systems.

Finally, the narrative contribution points to regulatory needs. The notion of informed consent and information self-determinism in the face of networked digital systems fails the individual in favour of those best able to exploit data. The result is one of identity assurance decay and subsequent identifier extraction expansion. This work converges to recommend further regulatory efforts on the extraction side of the equation rather than regulating self-disclosure.

7.3 Limitations

The limitations of this research have two levels of abstraction. In the weeds, as it were, there are limits to the methods chosen to conduct the work. For instance the Elo ranking algorithm from Chapter 3, or the reliance on the credibility of the credibility heuristics in Chapter 4. Each of these were discussed in the respective Chapters. At a higher abstraction, the complexity of the model of models is a limitation. Using different techniques involves moving away from the equation-based approach in Vila et al. (2003) towards an agent-based simulation. Principally the agent approach permits the introduction of factors such as adaptation, spatial interaction, signalling and decisional inconsistencies that reduce to population averages in equation-based models (Davidsson, 2000). This approach introduces flexibility yet reduces the tractability of the model. The approach removes the norm for equilibrium analysis in game theory that can remain intractable in real-world situations (Axelrod, 1997; Bonabeau, 2002). Agent simulation can have an advantage over game theory principles that express the system's stable states, by instead allowing one to peer into a system in motion (Gray and Bullock, 2014). For instance, in Vila et al. (2003) the equilibrium analysis provides an insight into the persistence of data exploitation, yet the implication of any long-term sustainability is not easily addressed with this method. As previously stated, the lack of rigour for such a simulation is open to criticism. Each element of the model, such as the harvesting element or how we code the two-system decision process between value and trust, is subject to debate and should be treated with caution and probed. As with any complex system, even a small change in these conditions can alter the outputs significantly. Therefore, each chapter's limitations could be additive, or, problematically non-linear.

Moreover, the approach is inter-disciplinary at its core. The literature spans legal, economic, social, psychological, technical remits, not as an aside, as a necessity. Identity assurance in the digital age is a vast topic that looks set to come ever-closer to the forefront of research, policy and regulation. This work treats these remits carefully, but it is certainly a compromise of depth for breadth from this complex goal.

7.4 Future Work and Applications

7.4.1 Layers of Regulation

Behaviour biometrics provides a frictionless yet somewhat covert method of authenticating a user. However, this frictionless approach is not without pitfalls; it only serves to put PII disclosure further into the unconsciousness of users, and therefore further away from scrutiny. Then, with similar techniques, a service can profile a user for some unilateral gain.

In what Jin (2013) describes as a 'privacy onion', it is possible to infer private information from public information. The onion analogy puts that low-value pieces of information such as daily activities exist at the outermost layer. In contrast, higher value identity information such as health details remain close to the inner layer. Who we are, impacts the decisions we make, including choices of how to represent ourselves, meaning that even an anonymous online identity often entail remnants of self-presentation (Lee and Wagenmakers, 2005). For instance, what a person 'likes' on Facebook would be the outer layer, but this can reveal sensitive knowledge about a person's sex, race or religious views within the inner layers (Kosinski et al., 2013).

In such a scenario, Jin's notion of a privacy onion starts to strip away layer by layer. Those individuals believing that they are revealing only the superficial may one-day cause distress (Attrill and Jalil, 2011). In the realm of identity assurance wherein instances of profiling and access abide by similar technologies, it is not a great leap to consider that my classification in Chapter 2 of three types of authentication from knowing and having, presence and being and living and doing, could progress to include beliefs and intentions.

The conclusion to Chapter 3 states that, in using a clustered approach to defining PII as a protected set of data, policymakers could deploy a multi-level licence approach to data management. Allowing organisation the opportunity and the hurdle of justifying the extent of data they extract, infer, and process. Each layer of this ′onion′ of data should require ever more justification. Moreover, to realign the incentives of consuming PII because it is there. This justification of PII extraction could then couple with liabilities in data insurance that better reflect the onion layer. It is perhaps time to level the terms rather than persist with the current externalities.

7.4.2 Data Quarantine

Obtaining consent, especially in the social aspects of digital living, involves many users in the pursuit of gratification. That is, a user′s primary goals often reside beyond the authentication process. There is consequently a lack of engagement with the process. In these promotion focuses transactions, one protection could be in the form of a data quarantine period. Like the cooling-off period on a purchase, it could be that a service can only process data for non-security reasons if, and only if, after a particular set of time, the individual has consented through opt-in a second time. From Chapter 3 we can take the recommendation that PII could be treated as distinct sets, with some ‘special’ status sets of PII, i.e., a licence to collect biometric data, or that this unique set could be subject to additional quarantine.

Federated management systems are a halfway house towards this. Instead of disclosing PII to each system, we engage with gatekeepers ‘vouching’ for our identity and therefore, overall, we disclose less. However, half is not sufficient, and indeed not satisfactory when at least in the UK, the federation entities, e.g., Experian, entrusted by the UK government for secure access, are also data controllers for peoples’ daily transactions and behaviours. Thus, creating a honey pot of PII.

One promising approach to bridge the gap might be to employ ′smart contracts′ that rely on integrity centric computer code, code that resists modification and subversion to execute transactions (Kosba et al., 2016). This way, the actual decision to permit data processing is made well beyond the initial gratification and can even be automated. This method would for those legitimately wanting the service to go through what may amount to simple consent and, later, more informed consent. Whereas for the transient user. For example, a user enters a digital service but doesn′t engage with the service in any meaningful manner. This encounter may suffice with only cryptographic proof of a few credentials (Bünz et al., 2020). If quickly leaving the system, there is no long-lasting exposure to their data as the contract would not ratify non-encrypted details. Moreover, it could be that such zero-knowledge proofs could suffice for many interactions in their entirety.

7.4.3 Recommender Systems

I believe a bright future for PII disclosure takes the majority of decisions out of user's hands but remains within their control. Like smart contracts that can conduct, without emotion, the administration of a decision. But, users also need assistance in deciding the first place. What is the fair disclosure price for a system? To find this price, it is not ideal to use the economic mantra that everything is worth what people are willing to pay because there exist many hidden costs. Future technological solutions should not discount intent as a standpoint. For instance, future users may delegate the disclosure of PII to automated negotiation agents, whereby the agents can manage the multitude of permissions required in the IoT (Baarslag et al., 2017). One barrier to the success of such methods is the initial training period for preferences. Using techniques such as those in Chapter 3 provides a way to harness the willingness of crowds. Systems could aggregate the partial preferences of individuals into a cohort result, and then an individual can bootstrap their decision from this baseline training. An obvious trap would be that this could formalise a herd mentality regardless, but it could be that a group of informed individuals give or trade their baseline results.

7.5 Final Remarks

Exploiting personal identifying information (PII) is critical for secure access to digital and web-based systems. It is also a significant element of the online social media business model. However, how this exploitation relates to users' valuation of their PII is poorly understood as an individual's willingness to disclose items of PII in different situations is unknown. Moreover, the actual cost of disclosure gets obfuscated within dense and lengthy policies designed to exploit additional data. Thus, an individual may not understand that systems such as facial recognition can be a gateway to infer further PII. Even with respectful intentions, identity-dependent technologies face many challenges to transparently balance users' sensitivities with their own need for high veracity PII.

As the regulation of PII is by consent, the use of heuristics rather than a robust evaluation is cause for concern. But there is also an opportunity. These heuristics, and the implication of their susceptibility to bias and manipulation (Krasnova and

Günther, 2009), could one day be harnessed so users can benefit from some form of a positive nudge and thus mediation of the risks (Balebako et al., 2011; Adjerid et al., 2013). Perhaps a more liberal paternalism whereby users tendency of heuristic-based disclosure adheres with cognitive nudges that may elicit more thoughtful consent (Acquisti et al., 2012; Balebako et al., 2011). Failing this, if we accept that the informed user is an unrealistic notion, our collective efforts can move towards more paternalistic regulations. Such regulations, e.g., the financial liability of data storage, could realign the risks and benefits that currently favour the data controllers (Varian, 1996).

Together, these findings indicate an inadequacy within current PII regulations, which treat PII as a homogeneous set. In contrast, a more nuanced set of rules would perhaps help avoid any over-use of high-value items. Moreover, this work implies that the self-managed model, whereby self-informed individuals are responsible for consenting or withholding personal data, is ideological.

To design new regulatory models is not an easy task, nor one that warrants haste. Yet, there is a clear need to understand the system of identity assurance better and understand what may temper usage escalations in identity assurance. My approach is agent-based simulation. Others will favour alternatives such as economic models, experimental psychology, ethical and legal debates. The key is exploring these ideas through all arenas in what is an increasingly multidisciplinary endeavour.

While current headlines contain synonyms of the words privacy and demise, these issues are far from new (Warren and Brandeis, 1890). Given the debates' longevity and that they maintain vigour, then the significant concerns around the demise of privacy have and will not come to pass. Nevertheless, the rate of personal data disclosure enabled by the increasingly integrated nature of the Web (Debatin et al., 2009; Edmonds and Gershenson, 2015) seems to have injected adrenaline into the discourse. Accordingly, for these debates to remain spirited, opposing sides are required, and perhaps the middle ground is open ground. That middle ground or balance, and mainly the motivation for this thesis, is that identity assurance systems of the future can thrive within privacy-enabled societies.

As a society with the potential for solutions to thrive, we must strive to protect PII for all users. Those leading technological curve must carefully tread the right path because the global and ubiquitous reach of modern technologies means that in terms of privacy and identity, we need to think collectively of PII as a public and a private good (Fairfield and Engel, 2015). If others are increasingly sharing their PII, however personal their own motivations, then the pressure increases on us to share our own PII. The goalposts move, and increasingly we will be expected to share PII that we consider being of high value. A public good indicates a common problem. We are all in this together.

Appendix A

Appropriating the ELO Ranking Algorithm

Ranking of subjective data items might conventionally be achieved through a card sort method. Whereby, a group of participants independently sort a set of cards, where each card represents some concept. It is a cheap method requiring relatively few participants, because the task includes verbal feedback with a facilitator. However, results can be overly specific to the participants, and therefore, are sensitive to biases within individual participants. Moreover, early grouping of the cards can become entrenched in the participants minds and influence the outcomes (Faiks and Hyland, 2000). Additionally, there is no clear way of handling conflicts between participant results. Perhaps most important however, is how quickly the card sort task becomes cognitively challenging as the number of cards increases past a relatively low threshold (Nielson, 1994).

For this study, the number of 'cards' is 30+, and this is a conservative number from what it could be in future work. The design must consider scalability. Therefore, I turn to an algorithmic design; the ELO ranking method. This method uses the results of direct competition to determine if a card should rise or fall in the ranking. In our scenario, this competition will be the pairwise comparison of two cards made by a single participant (the judge). With one winner and one loser, based on the participants interpretation of a simple judgement question, i.e., which of the cards would you prefer to disclose if given the option.

This is a distributed method that avoids the need for participants to see every item. That feature negates the cognitive load con from the card sort method. Instead, each participant contributes to a proportion of the overall work. In addition, participants are not aware, and therefore, not influenced by any emerging results. This means that increasing the number of cards only necessitates that more participants are required, while the actual work by each participant can remain constant.

Using the ELO algorithm, requires a number of participants $[P]$, making a number of pairwise comparisons $[C]$ on a number of cards (data items) $[D]$. What is unclear, is the sufficient and feasible magnitude of these variables. Here we approach these questions with two factors in mind, the ideal, and the feasible.

$[C]$ **Number of Pairwise Comparisons:** If the aim as to reduce burden on the participants, then $C = 1$ could be considered an ideal. However, in a controlled experiment setting, the cost of recruitment and time devoted by participant makes this an unattractive low bar. Intuition and a quick practice suggest that due to tedium the number of pair-wise comparisons should be limited at $C = 50$.

$[D]$ **The Number of Data Items:** This study includes 33 cards representing 33 different data items. This subset of all potential identifiers are chosen as they are likely to be recognised as within the bounds of normal online use, or at least cause little confusion.

$[P]$ **Pairs of Data Items:** The traditional card sort method had task consistency, whereby every participant sorts every item. This is not a feasible constraint to have on our design. Becasue if D is the number of data items then P is the number of possible pairings, and if we wish to keep $C <= 50$, then the maximum number of items is $D <= 10$. For example, if each participant judges each pair, ($C = P$), using permutation $(nPr, A : B! = B : A)$ arrives at $nPr = 42$, when $D = 7$. Alternatively, using combination $(nCr, A : B = B : A)$, then $nCr = 45$, when $D = 10$. Given that the variety of data items available that we wish to choose from is $D > 30$, it is obvious that task consistency is inappropriate, and therefore that compromising task consistency, at least in terms of which pairs are presented, is unavoidable.

Consequently, setting $D = 35$ and $C = 50$ gives $nPr = 1190$ and $cPr = 595$. The combination result is obviously more attractive for the sake of keeping numbers manageable[1]. Now, the question is how best to divide 595 pairs over a cohort of participants.

A.1 Baseline

One way to proceed is to simply divide P over C $(595/50) = 11.9$. Then dividing all pairs into 11 sets of 50 (plus 1 set of 45), and then with only 12 participants each working on a single set, the 35 items and 595 pairs are covered, with each item playing $N - 1 = 34$ times. However, this design is highly dependant on the preferences of individual participants, i.e., if one participant was adjudicator for most of a particular item's games, it could sway the result. Moreover, if the list does not stabilise, and 34 is

[1] It is prudent that in the experiment data items are presented side by side in a neutral way, and results remain aware of potential bias or anchoring resulting from an item appearing on the left (read first) compared to on the right side of the screen.

not a sufficient number games for a data item to find a true rank using our population, then this design does not provide enough flexibility to add more participants. Nor is it flexible enough to simply increase the number of item games, or add to the list of identifiers. Instead, in this design, adding participants would result in lesser work per participant, but the distribution and combined effect would remain unaffected. Nevertheless, this approach provides a good baseline from which to improve our method (summarised in Table Table A.1).

A.1.1 Random Allocation

To avoid possible participant bias, and to allow for a greater coverage of the data, we introduce random pair allocation. We now design a task where each participant receives a randomly chosen set of data pairs. Therefore, with $P = 595$ and retaining $C = 50$, what is the value of N that will adequately cover the list. This design however, introduces its own problems. An immediate trade-off is that with randomisation there is no guarantee that a pair will be presented at all. This design was simulated to help confirm the the spread of item pairs to better understand how the chosen parameters effect the likelihood of pairs being played and the number of times a particular items plays. For the simulation;

- Reverse pairs were considered equal (i.e. $A : B = B : A$).
- Items were equally as likely to appear on the left or right of the pairs.
- The simulation was repeated 5000 times

Using the same parameters as in the baseline $[N = 12, D = 35, C = 50]$, the first effect of random pair allocation is that (although an outlier) a pair was played 9 times, whereas in the baseline all pairs play once and once only. However, a significant proportion (34%) of pairs do not play a single game. While this is not essential, as it is similar to competitive sports, the first application of our chosen algorithm, nevertheless, the distribution could certainly improve. Table A.2 provides an overview for a quick comparison with the baseline. Looking at the individual item distributions, as one may expect the distribution oscillates either side of the baseline value of (D-1) 34 games, with a range 13 to 63. Given that our baseline was 34, a low-end range of 13 is far from acceptable. Therefore, we use our added flexibility to explore the parameter space in order to obtain a suitable balance.

A.1.2 Final Parameters.

Conducting a parameter sweep, with some intuitions) provided the final choice of parameters. Seeking a balance between number of participants and work to complete, maintaining $C = 50$ and $D = 35$, $N = 40$ seems to work well. These parameters have the desirable properties of a mean average of 97% of pairs playing at least once. In addition, the minimum item frequency more than doubles the baseline value (minimum of 72 item plays) and a stronger mean of 114.

TABLE A.1: Summary of The Baseline Design

Parameter	#	Distribution	Pros	Cons
Participants (N)	12	Each Pair plays exactly once	Simple design, Equal pair comparisons	Participant Bias, Inflexible
Comparisons (C)	50			
Data Items (D)	35	Items play D-1 games		
Pairs using nCr (P)	595			

TABLE A.2: Summary of Random allocation using baseline parameters

Parameter	Value	Distribution	Pros	Cons
Number of Participants (N)	12	Mean: 1.01 SD: 0.93 Range: 0 - 9	Flexible Experiment design	Some items play considerably more than others Lowest item play range too low 34% of pairs do not play
Number of Pairwise comparisons (C)	50			
Number of Data Items (D)	35	Mean: 34.3 SD: 5.66 Range: 13 - 63		
Number of possible Pairs using nCr (P)	595			

TABLE A.3: Summary of Final Parameters

Parameter	Value	Distribution	Pros	Cons
Number of Participants (N)	40	Mean: 1.01 SD: 0.93 Range: 0 - 9	Flexible Experiment design 97% of pairs play High minimum item play (72)	Some items play considerably more than others
Number of Pairwise comparisons (C)	50			
Number of Data Items (D)	35	Mean: 34.3 SD: 5.66 Range: 13 - 63		
Number of possible Pairs using nCr (P)	595			

Appendix B

Interview Materials for Disclosure Heuristics Study

B.1 Interview Schedule

Introduction and Consent: Before we begin, I have an information sheet that outlines the interview process and your involvement, I also have a form for you to sign which provides consent for me to record this interview and use the information for the research purposes outlined in the information sheet. The most important things I would like to bring to your attention, is that you have the right to stop and withdraw your involvement at any time, without consequence, including withdrawal of the recording even after this interview. Here is your Participant Information Sheet: please take the time to read the contents. Here is the Consent Form for you to review and sign if consent is granted. I will also ask you to reconfirm this consent at the end of the interview. Do you have any questions?

Warm-up Questions: Online Engagement We will now begin with the interview and start the audio recording.

Q1a: On an average day, what devices do you connect to the Internet with?

Q2a: Try to visualise your inbox. What subscription emails do you receive?

Q2b: Can you think of an occasion when you unsubscribed to an email list?

Q3a: If any, what types of items do you purchase online?

Q3b: Groceries? Entertainment? Clothes?

Q3c: Is there anything you would not purchase online? Why not?

Q4a: Are you a member of any social networking sites?

Q4b: How frequently do you use these services?

Q4c: What do you use these services for?

The Walk Through: Promotion Focus The Internet has many social applications and products available for free. Games such as Candy Crush, Angry Birds or Farmville are very popular whilst social media sites such as Facebook and Twitter are used extensively around the world.

Q5a: Think of a time when you registered for a free online service or application.

Q5b: From the very b eginning when you first considered the service, tell me about this registration in as much detail as possible. Take your time and try to visualise the process.

Q6: See prompt questions in Appendix B

The Walk Through: Prevention Focus People are shopping increasingly online for products such as clothes, flights or insurance, meaning that credit card checks and security processes are an important aspect of keeping people secure online.

Q6a: Think of a recent online purchase you have made.

Q6b: In your own time and from the very beginning when you first considered the item, tell me about the purchase in as much detail as possible. Try to visualise the process as you go.

Closing Questions and Final Consent Additional to the data we manually enter online, many organisations collect and store information about users including location data from GPS applications or your browser history via cookies. Currently the majority of this activity seems to be for commercial gains such as targeted advertisements, however other uses exist.

Q7a: Do you have any general concerns about extensive data collection?

Q7b: How do you feel this affects you when disclosing personal information?

Q7c: How much control of your personal information do you believe you have?

Before we end the interview and end the recording, is there anything you wish to add? Here is the Consent Form once again, please take your time to review the final section of the form and sign to indicate your consent

B.2 Interview Prompt Set

Set of possible follow up questions from which the participants can be prompted in a consistent manner across all interviews, however not all questions may be asked, the order may alter and additional prompts were used. **Searching for a service:** How did you find this product? Search? Advert? Website?

- Were alternative retailers available? Did you consider any of them?
- What made you choose this product/service? Friends? Colleagues?
- What device did you use and why?
- How familiar was the retailer? Was it a national chain? Is this important?
- What made you choose this retailer?
- Did you look at any reviews? How important are reviews?
- How confident were you about quality of the product/service?
- How did you know you were getting the right product at the right price?

Committing to a service:

- Was a registration required? Tell me about this process
- Can you recall the information required for registration?
- Did you provide all the information requested? Why?
- How do you feel about falsifying personal details?
- Would you use a third party to log in? Why?
- Describe the password process. Would you use a password manager?
- Describe the terms and conditions process
- Did you read the terms and conditions? Why?
- At what point had you committed? What would have deterred you?
- Was there anything unexpected about the process?
- Did you need to download any software? Tell me about this.
- Did you feel secure during the payment process? Why?

Post Commitment:

- Was this service linked to any social media sites?
- Did you post a review or recommendation? Why?
- What problems did you anticipate from registering?

- What do you expect will happen to your information?
- Do you think your data will be transferred to a third party? Why?
- What data protection assurances were present?

Appendix C

Pilot Model and Simulation

C.1 Defining the Model

Service A and Service B are two financially free to use Web services. These services operate for free despite substantial costs due to their ability to use user's data for the purposes of advertisement (Zhang et al., 2012).

[SPS] Service Pure Strategy: A service has two pure strategic options $[SPS_{Respect} \bigoplus SPS_{Exploit}]$. A service can adopt a pure respectful data strategy and limit data collection to overtly volunteered disclosures. Alternatively, a service can choose a pure exploiting strategy to covertly extract additional personal data.

Overt Disclosure: Regardless of the service strategy, a user consents to overtly and voluntarily disclose personal data. Within this consent, the user accepts that apart from security reasons there is a need to disclose some information for the proposes of basic advertisements (Toubiana et al., 2010).

Covert Extraction: A user knows that there may or may not exist some covert extraction by the service. That is, the service may be collecting extra detail via a misleading, obstructive or confusing privacy policy (Furnell and Phippen, 2012), and or selling personal data to third parties (Laudon, 1996). Note, disclosure and extraction are distinguished here for emphasis.

[PPU] Profit Per User: In practice, profit per user would likely change depending on economies of scale, service overheads or even the residential nation of users. Here, the profit a service achieves is a function of number of users and current strategy. In this scenario, all strategies are profitable.

Gratification: The user gains a level of gratification from engagement with a service, for instance, a voucher, information or a map (Acquisti, 2004a).

Value of Disclosure: Although this model explicitly uses the cost of disclosure, it is understood that the true cost of disclosure is unreasonable to determine due to dormant uncertainties, as discussed in (Weitzner et al., 2006). Therefore, the value of disclosure here is merely a simplification of the knowledge that each identifier disclosed is a non-zero exchange.

[UPS] User Pure Strategy: Each user has two pure strategic options $UPS_{Attend} \bigoplus UPS_{Ignore}$. A user may ignore the strategy of a service and proceed without judgement. Alternatively, a user may attend to the strategy of the service and reveal any covert extraction practices, if any do indeed exist (Furnell and Phippen, 2012).

Join/Remain or Avoid/Leave: Engagement occurs when the gratification outweighs the costs. A user adopting a UPS_{Ignore} strategy will decide to join or decide to remain. A user adopting a UPS_{Attend} strategy will join or remain with a service with $SPS_{Respect}$, else, they will avoid or leave the service.

C.2 Defining the Game

We refer to this as a game, because there is an opponent to this strategy. The opponent in question is the service that also has a strategy to maximise a return. That is, they would naturally wish to engage a number of users with their service, at a minimal internal cost and maximum revenue. We focus on a game that is non-zero-sum, as the gaining actions of one player do not necessarily lead to loss with the other players. This game is sequential in that each play is allocated to play a move in isolation, this player knows the previous action of the opponent player, and therefore, may assume that the opponents move is temporarily stable. The game is of imperfect information because there is no lasting memory of past choices in the game. Each service looks back to see their opponent's last move, but otherwise forward to judge their own move. Any one snapshot of the game play could be considered as symmetrical, however the ever-changing circumstances for each service created by the iterated game led more to asymmetric description. If one service has a dominant strategy to respect they choose to respect, regardless of the other service's strategy.

Memory: Users can re-join a service at a later turn. Moreover, users may re-join a service that may previously have exploited their data. That is, reputation damage and user grudges are not incorporated into this model, although these would be relevant for future work.

[ML] Market Latency: The number of user currently not engaged with a service plays an important role, because it is for this pool of disengaged population that a service can gain new users. Note that users can decide to engage with either service or opt out altogether however, they implicitly wish to engage.

Forecasting: Each service only considers the immediate future to assess best strategy gains for that step rather than a trend. Additionally services will discriminate even the smallest amount of profit gain as advantageous. In practice the utility of profits are likely to count for more complexity than this such as cost to change strategy, ethics, reputation and a longer-term view of the market.

Growth: Each service knows what the other service did last turn and assumes that they will not change in the next turn. They also know the global distribution of users strategies $[UPS]$ and the current size of the latent market [ML]. Each service therefore, looks at their expected growth depending on their chosen strategy, changing strategy and what their opponent is currently doing.

Change Dynamics: In the simulation, the basic logistic function $[vx * (1 - x/k)]$, is used for some key variables. This function was chosen for its remarkable validity across many different yet related scenarios, such as, birth and death rates, rumour or infection spread, or indeed industry growth (Strogatz, 2001). The following three components of the simulation use this function.

Population Consciousness: The first use is the rate at which the population of users become more or less conscientious towards checking for covert data extraction. Therefore, depending on the strategies of each service and the number of currently conscious users $[UPS_{Attend}]$ the number of conscious users will grow or decline analogous to the spread of a rumour.

Expected Change in Market Share: The second use governs how a service's market share will grow or decline in each time step T. This dynamic is dependant of two factors, the consciousness of the population UPS and the balance between current market share and the latent market. A Service would expect to grow if respecting and $ML > 0.0$, or if exploiting and $ML > 0.0$ and $UPS_{Attend} \rightarrow 0.0$. Likewise, a service would expect to decline is exploiting and $UPS_{Attend} \rightarrow 1.0$.

Therefore, if a service has a respect strategy they would forecast the following growth.

$$Growth = \left(vS * MarketShare\right) * \left(1 - \frac{MarketShare}{LatencyShare + MarketShare}\right) \tag{C.1}$$

When both service share the same strategy then the latent market is considered as a shared potential.

$$LatencyShare = \frac{MarketLatency}{2} \tag{C.2}$$

Yet, when the strategies are contrasting, then for the respecting service, they get the share of the ignorant latent market, and the whole share of the conscious latent market. So,

$$LatencyShare = MarketLatency * \left(\frac{UPS_{Ignore}}{2} + UPS_{Attend}\right) \tag{C.3}$$

Else, if a service has exploit strategy, then growth depends on the ignorance of the latent market,

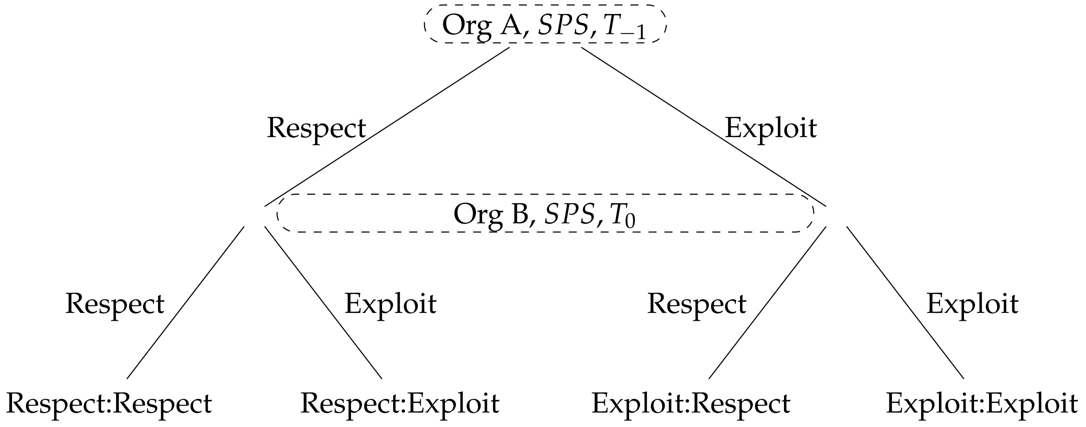
$$LatencyShare = \frac{MarketLatency * UPS_{Ignore}}{2} \tag{C.4}$$

$$Decline = \left(vS * MS\right) * \left(1 - \frac{UPS_{Ignore} * MarketShare}{MarketShare)}\right) \tag{C.5}$$

Market Movement: Whilst each organisation has a forecast regarding the growth or decline of their market share, the actual movement of the market users is less predictable. The number $[x]$ is calculated in a similar way using the predicted growth dynamics of a market depending on current size and maximum size, however, the variable of number of user each turn is chosen between $[x/2]$ and $[x + x/2]$. This is the number of user action per time step.

C.3 Extensive Form, Sub Game

In our repeated sequential game, each decision can be represented in the following form. Decision two being the active decision, with the knowledge of what decision one already was. Then as decision two is made it becomes the subsequent decision one.



C.4 Initial Variables

- $[V : Service] = 0.01$
- $[V : User] = 0.001$
- $[PPU_{Respect}] = 1.0p/u$
- $[PPU_{Exploit}] = 1.05p/u$
- $[P]Population = 10000$
- $[T]TimeSteps = 250000$
- $UPS_{Attend} = 1\%$
- $SPS_{Respect} = 2$
- $MarketShare =$ Random allocation of 200 users

C.5 Initial Observations and Discussion

The oscillating dynamics in Figure C.1 are largely down to two factors. The first if the cost of consciousness and the second is the attraction of data exploitation. It is not hard to reason that with these two opposing factors at play, a stable system outcome is never reached. This is where agent simulation has an advantage over static game theoretic analysis, as the agent approach allows us to peer into a system that may not benefit from equilibrium analysis (Gray and Bullock, 2014).

It is the cost of revelling the negatives associated with disclosure that are covering those very negatives. In addition, this cost to reveal any negatives is problematic, as it is easier borne by those in society with the most resources. Those with limited education and or money, cannot as ready understand disclosure, or delegate the protection respectively. Therefore, if a service for any reason deems that this cost of consciousness is to high for the majority of users to incur then it could lead to a moral hazard. This is perhaps more so when the incursion may even happen in the background by third party companies, or increasingly by the objects with which we interact with. For in the real world, people have less and less opportunity to opt out in an environment integrating with ambient technologies (Cook et al., 2009). This situation exits for two competing services, but what about multiple services, then the cost of consciousness accumulates within the system. However, this cost spread over multiple services is seemingly below the engagement threshold for most users, hence the attractiveness of the Web. Or,

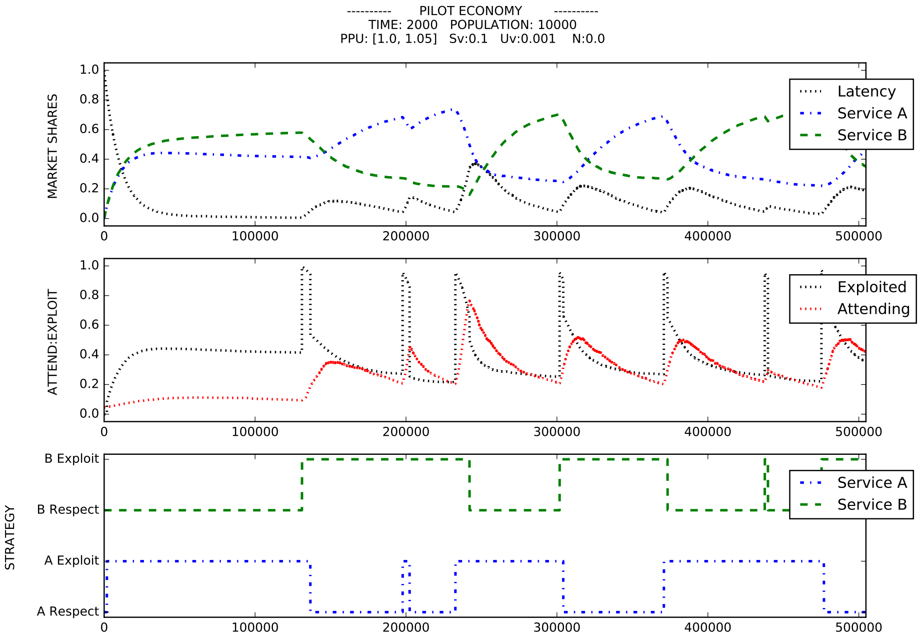


FIGURE C.1: Initial observations of an oscillation with the system.

$$\sum_{i=1}^{n} Gratification > \sum_{i=1}^{n} PercievedNegatives \tag{C.6}$$

Yet,

$$\sum_{i=1}^{n} PerceivedNegatives > 0 \tag{C.7}$$

As often cited, users during any particular moment can perceive one instance of identity related issues as being relatively high, at least in the abstract (Norberg et al., 2007). Nevertheless, during specific instances, engagement generally occurs as the consequence stay in the unconsciousness (Beer, 2009; Friedewald et al., 2007; Pollard, 2006). Yet, as any perceived negative increase, with media stories, whistle-blower revelations or personal victimisations, the gratification gained is the first casualty, and markets may fail to reduce their latency. That is, unless it is possible to move towards an environment whereby risks are reasonably apportioned and the respectful services do not make a veil of respect over those with less ethical operations. Ideally then, a respectful service should want to bear some of the cost of consciousness so that users on a large enough scale can differentiate and thus, punish exploitation.

While useful, the pilot model is not adequate to explore some of these thoughts let alone some of other complexities of the system. Therefore, the next step in the process is to develop the model and subsequently intricacy and realism of the simulation.

Appendix D

Simulation Tables & Overview

TABLE D.1: S_K EQUILIBRIUM RESULTS

Scenario	RF	ProA	KP	EQ	KP2	EQ2
*	0	0	0.74	($\hat{K}$, K, 1.0)	**0.23**	(K, $\hat{K}$, 1.0)
*	0.2	0	0.84	(K, K, 1.0)	**0.26**	(K, $\hat{K}$, 1.0)
*	0.4	0	0.84	(K, K, 1.0)	**0.32**	(K, $\hat{K}$, 1.0)
	0.6	0	0.83	(K, K, 1.0)	0.83	(K, K, 1.0)
	0.8	0	0.83	(K, K, 1.0)	0.83	(K, K, 1.0)
	1	0	0.84	(K, K, 1.0)	0.84	(K, K, 1.0)
*	0	0.1	0.83	(K, K, 1.0)	**0.24**	(K, $\hat{K}$, 1.0)
RW*	0.2	0.1	0.84	(K, K, 1.0)	**0.27**	(K, $\hat{K}$, 1.0)
*	0.4	0.1	0.84	(K, K, 1.0)	**0.34**	(K, $\hat{K}$, 1.0)
	0.6	0.1	0.83	(K, K, 1.0)	0.83	(K, K, 1.0)
EU	0.8	0.1	0.84	(K, K, 1.0)	0.84	(K, K, 1.0)
	1	0.1	0.84	(K, K, 1.0)	0.84	(K, K, 1.0)
*	0	0.2	0.84	(K, K, 1.0)	**0.26**	(K, $\hat{K}$, 1.0)
*	0.2	0.2	0.84	(K, K, 1.0)	**0.3**	(K, $\hat{K}$, 1.0)
*	0.4	0.2	0.84	(K, K, 1.0)	**0.36**	(K, $\hat{K}$, 1.0)
	0.6	0.2	0.84	(K, K, 1.0)	0.84	(K, K, 1.0)
	0.8	0.2	0.83	(K, K, 1.0)	0.83	(K, K, 1.0)
	1	0.2	0.84	(K, K, 1.0)	0.84	(K, K, 1.0)
*	0	0.3	0.84	(K, K, 1.0)	**0.22**	(K, $\hat{K}$, 1.0)
*	0.2	0.3	0.84	(K, K, 1.0)	**0.32**	(K, $\hat{K}$, 1.0)
*	0.4	0.3	0.84	(K, K, 1.0)	**0.39**	(K, $\hat{K}$, 1.0)
	0.6	0.3	0.84	(K, K, 1.0)	0.84	(K, K, 1.0)
	0.8	0.3	0.84	(K, K, 1.0)	0.84	(K, K, 1.0)
	1	0.3	0.83	(K, K, 1.0)	0.83	(K, K, 1.0)
*	0	0.4	0.84	(K, K, 1.0)	**0.22**	($\hat{K}$, $\hat{K}$, 1.0)
*	0.2	0.4	0.84	(K, K, 1.0)	**0.36**	(K, $\hat{K}$, 1.0)
*	0.4	0.4	0.84	(K, K, 1.0)	**0.43**	(K, $\hat{K}$, 1.0)
	0.6	0.4	0.84	(K, K, 1.0)	0.84	(K, K, 1.0)
	0.8	0.4	0.84	(K, K, 1.0)	0.84	(K, K, 1.0)
	1	0.4	0.84	(K, K, 1.0)	0.84	(K, K, 1.0)
*	0	0.5	0.84	(K, K, 1.0)	**0.22**	($\hat{K}$, $\hat{K}$, 1.0)
IU*	0.2	0.5	0.84	(K, K, 1.0)	**0.25**	($\hat{K}$, $\hat{K}$, 1.0)
*	0.4	0.5	0.84	(K, K, 1.0)	**0.48**	(K, $\hat{K}$, 1.0)
	0.6	0.5	0.83	(K, K, 1.0)	0.83	(K, K, 1.0)
HG	0.8	0.5	0.83	(K, K, 1.0)	0.83	(K, K, 1.0)
	1	0.5	0.84	(K, K, 1.0)	0.84	(K, K, 1.0)

p20000, steps of s8000, using strategy set one, $S_K = \{\{K\}, \{\hat{K}\}\}$. RF : Regulatory Focus, ProA : Prominence of Service A, KP : Knowledge Potential, EQ : Strategy Equilibrium.

TABLE D.2: S_K REAL WORLD UTILITY MATRIX

		Service B	
		K	$\hat{K}$
Service A	K	**0.75,5.05**	0.5, 1.39
	$\hat{K}$	0.62, 3.85	0.19, 1.27

Long-Sighted

		Service B	
		K	$\hat{K}$
	K	1.2, 7.11	**3.05,8.97**
	$\hat{K}$	1.33, 7.32	1.35, 9.02

Short-Sighted

TABLE D.3: S_K EDUCATED USERS UTILITY MATRIX

		Service B	
		K	$\hat{K}$
Service A	K	**0.75,5.08**	2.45, 0.09
	$\hat{K}$	0.27, 5	0.15, 1

Long-Sighted

		Service B	
		K	$\hat{K}$
	K	**1.07,7.05**	6.4, 3.09
	$\hat{K}$	0.49, 8.09	0.67, 3.08

Short-Sighted

TABLE D.4: INDIFFERENT USERS UTILITY MATRIX

		Service B	
		K	$\hat{K}$
Service A	K	**2.91,2.92**	1.32, 1.24
	$\hat{K}$	1.23, 1.33	0.73, 0.73

Long-Sighted

		Service B	
		K	$\hat{K}$
	K	3.92, 4.25	5.05, 5.85
	$\hat{K}$	5.67, 4.89	**5.28,5.52**

Short-Sighted

TABLE D.5: S_K HOLY GRAIL UTILITY MATRIX

		Service B	
		K	$\hat{K}$
Service A	K	**2.88,2.9**	3.45, 0.73
	$\hat{K}$	0.73, 3.42	0.58, 0.58

Long-Sighted

		Service B	
		K	$\hat{K}$
	K	**4.08,3.98**	7.22, 1.63
	$\hat{K}$	1.74, 7.05	1.96, 1.64

Short-Sighted

TABLE D.6: S_K Four Scenarios : Environment Results

Scenario	RF	ProA	Type	Value	EQ1	Value 2	EQ2
RW	0.2	0.1	A Utility	0.75	(K, K, 1.0)	0.5	(K , $\hat{K}$, 1.0)
			A Market Share	0.13		0.35	
			A Assurance	0.17		0.06	
			A Profit	4.43		4.52	
			B Utility	5.05		1.39	
			B Market Share	0.87		0.65	
			B Assurance	0.17		0.05	
			B Profit	29.81		13.87	
EU	0.8	0.1	A Utility	0.75	(K, K, 1.0)	0.75	(K, K, 1.0)
			A Market Share	0.13		0.13	
			A Assurance	0.17		0.17	
			A Profit	4.43		4.43	
			B Utility	5.08		5.08	
			B Market Share	0.87		0.87	
			B Assurance	0.17		0.17	
			B Profit	29.88		29.88	
IU	0.2	0.5	A Utility	2.91	(K, K, 1.0)	0.73	($\hat{K}$, $\hat{K}$, 1.0)
			A Market Share	0.5		0.37	
			A Assurance	0.17		0.04	
			A Profit	17.13		7.35	
			B Utility	2.92		0.73	
			B Market Share	0.5		0.37	
			B Assurance	0.17		0.04	
			B Profit	17.18		7.35	
HG	0.8	0.5	A Utility	2.88	(K, K, 1.0)	2.88	(K, K, 1.0)
			A Market Share	0.5		0.5	
			A Assurance	0.17		0.17	
			A Profit	17.03		17.03	
			B Utility	2.9		2.9	
			B Market Share	0.5		0.5	
			B Assurance	0.17		0.17	
			B Profit	17.16		17.16	

p20000, steps of s8000, using strategy set one, $S_K = \{\{K\}, \{\hat{K}\}\}$.

TABLE D.7: S_{KB} EQUILIBRIUM RESULTS

S	RF	PA	KP	BP	KBP	EQ	KP2	BP2	KBP2	EQ2
*	0	0	0.89	0.95	0.92	(K,K, 0.5)(B,K, 0.25)(K,B, 0.16)(B,B)	0.56	0.57	**0.57**	($\hat{K}$, $\hat{B}$, 0.4)($\hat{k}$, $\hat{K}$B, 0.23)($\hat{B}$, $\hat{B}$, 0.23)($\hat{B}$, $\hat{K}$B, 0.13)
*	0.2	0	0.89	0.95	0.92	(K,K, 0.49)(B, K, 0.26)(K,B, 0.17)(B,B)	0.62	0.39	**0.51**	($\hat{B}$, $\hat{B}$, 0.5)($\hat{B}$, $\hat{K}$B, 0.37)($\hat{K}$B, $\hat{B}$)($\hat{K}$B, $\hat{K}$B)
*	0.4	0	0.89	0.95	0.92	(K,K, 0.5)(B,K, 0.26)(K,B, 0.16)(B, B)	0.57	0.31	**0.44**	($\hat{B}$, $\hat{B}$B, 1.0)
	0.6	0	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.25)(K, B, 0.17)(B, B)	0.89	0.95	0.92	(K,K, 0.5)(B, K, 0.25)(K, B, 0.18)(B, B)
	0.8	0	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.26)(K, B, 0.16)(B, B)	0.89	0.95	0.92	(K,K, 0.5)(B, K, 0.26)(K, B, 0.16)(B, B)
	1	0	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.26)(K, B, 0.16)(B, B)	0.89	0.95	0.92	(K,K, 0.5)(K, B, 0.17)(B, K, 0.25)(B, B)
*	0	0.1	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.26)(K, B, 0.15)(B, B)	0.43	0.69	**0.56**	(K, $\hat{B}$, 0.4)($\hat{B}$, $\hat{B}$, 0.2)(K,K)($\hat{K}$, $\hat{K}$B, 0.19)($\hat{B}$, $\hat{K}$)($\hat{B}$
RW*	0.2	0.1	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.26)(K, B, 0.15)(B, B)	0.62	0.39	**0.5**	($\hat{B}$, $\hat{B}$, 0.47)($\hat{B}$, $\hat{K}$B, 0.37)($\hat{K}$B, $\hat{B}$)($\hat{K}$B, $\hat{K}$B)
*	0.4	0.1	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.26)(K, B, 0.16)(B, B)	0.56	0.31	**0.44**	($\hat{B}$, $\hat{B}$B, 1.0)
	0.6	0.1	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.26)(K, B, 0.16)(B, B)	0.89	0.95	0.92	(K,K, 0.51)(B, K, 0.27)(K, B, 0.15)(B, B)
EU	0.8	0.1	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.26)(K, B, 0.16)(B, B)	0.89	0.95	0.92	(K,K, 0.51)(B, K, 0.28)(K, B, 0.13)(B, B)
	1	0.1	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.27)(K, B, 0.15)(B, B)	0.89	0.95	0.92	(K,K, 0.5)(B, K, 0.28)(K, B, 0.14)(B, B)
*	0	0.2	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.27)(K, B, 0.14)(B, B)	0.67	0.37	**0.52**	($\hat{B}$,$\hat{B}$, 0.52)($\hat{K}$B, $\hat{B}$, 0.23)($\hat{B}$,$\hat{K}$B, 0.17)($\hat{K}$B,$\hat{K}$B)
*	0.2	0.2	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.27)(K, B, 0.13)(B, B)	0.62	0.39	**0.51**	($\hat{B}$, $\hat{B}$, 0.45)($\hat{B}$, $\hat{K}$B, 0.4)($\hat{K}$B, $\hat{B}$)($\hat{K}$B, $\hat{K}$B)
*	0.4	0.2	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.27)(K, B, 0.13)(B, B)	0.56	0.31	**0.44**	($\hat{B}$, $\hat{B}$B, 1.0)
	0.6	0.2	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.28)(K, B, 0.12)(B, B)	0.89	0.95	0.92	(K,K, 0.52)(B, K, 0.27)(K, B, 0.14)(B, B)
	0.8	0.2	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.27)(K, B, 0.14)(B, B)	0.89	0.95	0.92	(K,K, 0.52)(B, K, 0.27)(K, B, 0.13)(B, B)
	1	0.2	0.89	0.95	0.92	(K, K, 0.50)(B, K, 0.28)(K, B, 0.13)(B, B)	0.89	0.95	0.92	(K,K, 0.5)(B, K, 0.27)(K, B, 0.16)(B, B)
*	0	0.3	0.89	0.95	0.92	(K, K, 0.61)(B, K, 0.34)(K, B)(B, B)	0.18	0.29	**0.24**	($\hat{K}B$,$\hat{K}\hat{B}$, 1.0)
*	0.2	0.3	0.89	0.95	0.92	(K, K, 0.62)(B, K, 0.36)(K, B)(B, B)	0.19	0.33	**0.26**	($\hat{K}B$,$\hat{K}\hat{B}$, 1.0)
*	0.4	0.3	0.89	0.95	0.92	(K, K, 0.63)(B, K, 0.36)(K, B)(B, B)	0.57	0.31	**0.44**	($\hat{B}$, $\hat{B}$B, 1.0)
	0.6	0.3	0.89	0.95	0.92	(K, K, 0.62)(B, K, 0.36)(K, B)(B, B)	0.89	0.95	0.92	(K,K, 0.58)(B, K, 0.34)(K, B)(B, B)
	0.8	0.3	0.89	0.95	0.92	(K, K, 0.62)(B, K, 0.36)(K, B)(B, B)	0.89	0.95	0.92	(K,K, 0.55)(B, K, 0.34)(K, B)(B, B)
*	1	0.3	0.83	1	0.92	(K,K, 1.0)	0.89	0.95	0.92	(K,K, 0.61)(K, B, 0.025)(B, K, 0.34)(B, B)
*	0	0.4	0.83	0.83	0.83	(KB, KB, 1.0)	0.18	0.33	**0.26**	($\hat{K}B$,$\hat{K}\hat{B}$, 1.0)
*	0.2	0.4	0.84	0.84	0.84	(KB, KB, 1.0)	0.19	0.37	**0.28**	($\hat{K}B$,$\hat{K}\hat{B}$, 1.0)
*	0.4	0.4	0.84	0.84	0.84	(KB, KB, 1.0)	0.23	0.85	**0.54**	($\hat{K}B$,$\hat{K}\hat{B}$, 1.0)
	0.6	0.4	0.84	0.84	0.84	(KB, KB, 1.0)	0.84	0.84	0.84	(KB, KB, 1.0)
	0.8	0.4	0.83	0.83	0.83	(KB, KB, 1.0)	0.83	0.83	0.83	(KB, KB, 1.0)
	1	0.4	0.84	0.84	0.84	(KB, KB, 1.0)	0.84	0.84	0.84	(KB, KB, 1.0)
*	0	0.5	0.84	0.84	0.84	(KB, KB, 1.0)	0.38	0.19	**0.29**	(K$\hat{B}$,$\hat{K}\hat{B}$, 1.0)
IU*	0.2	0.5	0.84	0.84	0.84	(KB, KB, 1.0)	0.19	0.42	**0.3**	($\hat{K}B$,$\hat{K}\hat{B}$, 1.0)
*	0.4	0.5	0.84	0.84	0.84	(KB, KB, 1.0)	0.23	0.85	**0.54**	($\hat{K}B$,$\hat{K}\hat{B}$, 1.0)
	0.6	0.5	0.84	0.84	0.84	(KB, KB, 1.0)	0.84	0.84	0.84	(KB, KB, 1.0)
HG	0.8	0.5	0.84	0.84	0.84	(KB, KB, 1.0)	0.84	0.84	0.84	(KB, KB, 1.0)
	1	0.5	0.83	0.83	0.83	(KB, KB, 1.0)	0.83	0.83	0.83	(KB, KB, 1.0)

p20000, steps of s8000, using strategy set one, $S_K = \{\{K\}, \{\hat{K}\}, \{B\}, \{\hat{B}\}, \{K, B\}, \{\hat{K}, B\}, \{K, \hat{B}\}, \{\hat{K}, \hat{B}\}\}$. % < 10 not shown.

TABLE D.8: S_{KB} **Real World - Long Term**

	Service B							
SA	K	B	$\hat{K}$	$\hat{B}$	KB	$K\hat{B}$	$\hat{K}B$	$\hat{K}\hat{B}$
K	**0.8, 5.1**	**5, 1.8**	0.5, 1.4	5.3, 0.5	5.8, 0	5.8, 0	5.8, 0	5.8, 0
B	**1.8, 5**	**1.7, 11.3**	5.4, 1.3	1.1, 3.1	12.9, 0	13, 0	12.9, 0	13, 0
$\hat{K}$	0.6, 3.9	1.3, 5.4	0.2, 1.3	1.3, 0.5	1.4, 6.2	1.4, 2.5	1.1, 5.7	1.5, 0
$\hat{B}$	0.5, 5.3	1.4, 8.6	0.5, 1.3	0.4, 2.8	3.1, 4.3	2.8, 2.1	3.2, 3.4	3.3, 0
KB	0, 5.8	0, 12.9	6.2, 1.4	4.3, 3.1	4.6, 31.2	3.8, 13	2.9, 18.2	3.1, 8.6
$K\hat{B}$	0, 5.8	0, 13	2.6, 1.4	2.1, 2.8	4.2, 25.7	1.8, 12.4	2.4, 14.7	2, 7.5
$\hat{K}B$	0, 5.8	0, 12.9	5.7, 1.1	3.4, 3.2	4.6, 26.7	3.1, 10.6	2.6, 17.4	3, 7.4
$\hat{K}\hat{B}$	0, 5.8	0, 13	0, 1.4	0, 3.3	3.8, 23.7	1.8, 11.1	2.2, 14.8	1.1, 7.8

TABLE D.9: S_{KB} **Educated User - Long Term**

	Service B							
SA	K	B	$\hat{K}$	$\hat{B}$	KB	$K\hat{B}$	$\hat{K}B$	$\hat{K}\hat{B}$
K	**0.8, 5.1**	**5, 1.8**	2.5, 0.9	5.6, 0.4	5.8, 0	5.8, 0	5.8, 0	5.8, 0
B	**1.8, 5**	**1.7, 11.3**	10.8, 1	5.4, 2.1	12.9, 0	12.9, 0	12.9, 0	13, 0
$\hat{K}$	0.3, 5	1, 10.7	0.2, 1	1, 0.4	1, 21.4	1.1, 0.2	0.7, 4.7	1.2, 0
$\hat{B}$	0.4, 5.6	0.6, 11.2	0.4, 1	0.3, 2.2	2.2, 18.5	2.5, 0.3	2.4, 3	2.5, 0
KB	0, 5.8	0, 13	21.4, 1	18.5, 2.2	4.6, 31.2	19.1, 6.7	17.1, 9.9	15.1, 5.7
$K\hat{B}$	0, 5.8	0, 12.9	0.3, 1.1	0.2, 2.5	1.6, 32.3	1.1, 7.5	1.2, 9.9	0.9, 6.2
$\hat{K}B$	0, 5.8	0, 13	4.6, 0.7	3, 2.4	2.4, 31.2	3.5, 5.8	1.5, 10.3	3.1, 5.3
$\hat{K}\hat{B}$	0, 5.8	0, 13	0, 1.2	0, 2.6	1.7, 30.9	1.3, 7.1	1.3, 9.8	0.9, 6.2

TABLE D.10: S_{KB} **Indifferent User - Long Term**

	Service B							
SA	K	B	$\hat{K}$	$\hat{B}$	KB	$K\hat{B}$	$\hat{K}B$	$\hat{K}\hat{B}$
K	2.9, 2.9	5, 1.7	1.3, 1.2	5.3, 0.5	5.8, 0	5.8, 0	5.8, 0	5.8, 0
B	1.8, 5	6.5, 6.5	5.4, 1.3	3, 2.8	13, 0	13, 0	13, 0	13, 0
$\hat{K}$	1.2, 1.3	1.3, 5.4	0.7, 0.7	1.3, 0.4	1.4, 6.2	1.5, 2.5	1.1, 5.6	1.5, 0
$\hat{B}$	0.4, 5.3	2.8, 3	0.4, 1.3	1.6, 1.6	3.1, 4.2	2.8, 2.1	3.1, 3.4	3.2, 0
KB	0, 5.8	0, 12.9	6.2, 1.4	4.3, 3.2	**17.9, 17.9**	10.6, 10.1	11.4, 12.7	8.1, 7.6
$K\hat{B}$	0, 5.8	0, 12.9	2.6, 1.4	2.1, 2.8	10.1, 10.6	7.1, 7.1	6.4, 7.3	5.4, 5.1
$\hat{K}B$	0, 5.8	0, 12.9	5.6, 1.1	3.4, 3.2	12.7, 11.4	7.3, 6.5	10, 10.1	6.7, 5.5
$\hat{K}\hat{B}$	0, 5.8	0, 13	0, 1.5	0, 3.2	7.6, 8.1	5.1, 5.5	5.5, 6.7	4.5, 4.5

TABLE D.11: S_{KB} **Holy Grail - Long Term**

	Service B							
SA	K	B	$\hat{K}$	$\hat{B}$	KB	$K\hat{B}$	$\hat{K}B$	$\hat{K}\hat{B}$
K	2.9, 2.9	5, 1.8	3.4, 0.7	5.7, 0.4	5.8, 0	5.8, 0	5.8, 0	5.8, 0
B	1.7, 5	6.5, 6.5	10.8, 1	7.6, 1.6	13, 0	12.9, 0	13, 0	13, 0
$\hat{K}$	0.7, 3.4	1, 10.8	0.6, 0.6	1, 0.4	1, 21.5	1.1, 0.3	0.7, 4.7	1.1, 0
$\hat{B}$	0.4, 5.7	1.7, 7.6	0.4, 1	1.3, 1.3	2.2, 18.6	2.5, 0.3	2.4, 3	2.6, 0
KB	0, 5.8	0, 13	21.7, 1	18.7, 2.2	**17.9,17.9**	24.4, 4.7	22.4, 6.8	21.1, 4.6
$K\hat{B}$	0, 5.8	0, 12.9	0.3, 1.1	0.3, 2.5	4.8, 24.3	4.3, 4.3	3.7, 6.2	3.5, 4.2
$\hat{K}B$	0, 5.8	0, 13	4.7, 0.7	3, 2.4	6.8, 22.3	6.1, 3.7	5.9, 5.9	5.7, 3.8
$\hat{K}\hat{B}$	0, 5.8	0, 12.8	0, 1.2	0, 2.6	4.5, 21.1	4.2, 3.5	3.8, 5.7	3.6, 3.6

TABLE D.12: S_{KB} **Real World - Short Term**

	Service B							
SA	K	B	$\hat{K}$	$\hat{B}$	KB	$K\hat{B}$	$\hat{K}B$	$\hat{K}\hat{B}$
K	1.7, 6.6	7.2, 2.7	3.3, 8.6	7.4, 3.2	8.1, 0	8.1, 0	8.1, 0	8.3, 0
B	2.4, 7.1	3.9, 14.3	6.3, 9.5	7.7, 18.5	18.3, 0	18.1, 0	18.3, 0	18.5, 0
$\hat{K}$	2.3, 6.9	9.2, 6.6	2.4, 8.3	9.2, 3.5	10.5, 13.3	10.8, 9.7	10.9, 14.6	10.8, 0
$\hat{B}$	3.2, 7.4	4.8, 16.1	3.3, 9.3	**5.1, 18.9**	24.5, 13	23.8, 9.3	**24.2, 15.7**	24.3, 0
KB	0, 8.3	0, 18.3	12.7, 10.7	13, 23.9	10.2, 40.7	15.1, 52.3	11.3, 52.2	20.7, 53.1
$K\hat{B}$	0, 8.1	0, 18.1	8.3, 10.9	9.3, 24.1	13.2, 41.6	13.7, 51.9	12.9, 54.9	20.3, 53.8
$\hat{K}B$	0, 8.2	0, 18.1	15.8, 10.7	**15.7, 24.2**	13.5, 40.6	20, 51.3	**13.3, 52.1**	27.4, 50.9
$\hat{K}\hat{B}$	0, 8.1	0, 18.2	0, 10.5	0, 24.4	15, 41.8	14.1, 53.4	14.1, 55.3	13.9, 52.3

TABLE D.13: S_{KB} **Educated User Short Term**

	Service B							
SA	K	B	$\hat{K}$	$\hat{B}$	KB	$K\hat{B}$	$\hat{K}B$	$\hat{K}\hat{B}$
K	**1.8, 6.3**	**7.3, 2.4**	6.6, 2.7	7.9, 1.2	8.2, 0	8.1, 0	8, 0	8.3, 0
B	**2.6, 6.9**	**4, 14.5**	14.2, 3.1	14.6, 6.7	18.1, 0	18.2, 0	18.2, 0	18.4, 0
$\hat{K}$	0.7, 7.6	3.3, 13.9	0.7, 2.8	3.3, 1.5	3.6, 37.7	3.9, 1	3.5, 14.6	3.8, 0
$\hat{B}$	1.1, 8	1.6, 17.5	1.1, 3.1	1.6, 6.6	7.5, 38.3	8.1, 1.2	8.7, 15	8.1, 0
KB	0, 8.1	0, 18.2	37.2, 3.7	38.1, 8.2	10.5, 39.9	39.5, 16	32.1, 24.7	40, 18.1
$K\hat{B}$	0, 8.3	0, 18.1	1.1, 3.6	1.1, 7.7	4.5, 47.2	5.1, 15.9	4.6, 28.1	4.7, 19.1
$\hat{K}B$	0, 8.3	0, 18.3	15.7, 3.7	14, 8.4	6.4, 45.8	18.2, 15.9	7, 24.7	16.1, 20
$\hat{K}\hat{B}$	0, 8.2	0, 18.2	0, 3.7	0, 8.5	4.6, 47.7	5, 16	4.8, 28.2	4.6, 18.2

TABLE D.14: S_{KB} **Indifferent User - Short Term**

	Service B							
SA	K	B	$\hat{K}$	$\hat{B}$	KB	$K\hat{B}$	$\hat{K}B$	$\hat{K}\hat{B}$
K	4, 4.2	7.1, 2.4	5, 5.5	7.3, 3.4	8.3, 0	8.4, 0	8.2, 0	8.1, 0
B	2.8, 6.8	9.3, 9	6.2, 9.5	10.9, 12.8	18.5, 0	18.3, 0	18.6, 0	18.1, 0
$\hat{K}$	5.2, 5.1	9.3, 6.4	5.3, 5.3	9.3, 3	10.8, 12.9	10.8, 9.4	10.9, 15.1	10.7, 0
$\hat{B}$	3.2, 7.4	13, 10.8	3, 9.6	12.3, 11.8	23.9, 13.2	24.1, 9.1	24.3, 15.5	23.8, 0
KB	0, 8.2	0, 18.3	12.5, 10.8	12.6, 24.1	25.3, 25.2	28.3, 33.3	25.7, 33.2	31.7, 33.2
$K\hat{B}$	0, 8.5	0, 18.2	9.6, 10.6	9.1, 24.3	32.5, 28.9	31.8, 32.2	32.1, 37.1	37.4, 33.2
$\hat{K}B$	0, 8.1	0, 18.3	15.2, 10.9	14.8, 24.3	32.2, 26.4	36.9, 32.6	31.5, 34.6	**39.7, 34.7**
$\hat{K}\hat{B}$	0, 8.3	0, 18.2	0, 10.9	0, 23.6	33.2, 31.9	33.3, 37.6	34.2, 40	33.3, 33.8

TABLE D.15: S_{KB} **Holy Grail - Short Term**

	Service B							
SA	K	B	$\hat{K}$	$\hat{B}$	KB	$K\hat{B}$	$\hat{K}B$	$\hat{K}\hat{B}$
K	4.2, 4	6.9, 2.7	7, 2	8, 1.2	8.3, 0	8.3, 0	8.1, 0	8, 0
B	2.4, 7.3	9, 9.1	14.1, 3.3	15.8, 4.3	18.4, 0	18, 0	18.2, 0	18.2, 0
$\hat{K}$	1.9, 7.1	3.1, 14.4	1.9, 1.9	3.4, 1.2	3.6, 37.8	3.7, 1.3	3.5, 13.5	3.6, 0
$\hat{B}$	0.5, 5.3	1.4, 8.6	0.5, 1.3	0.4, 2.8	3.1, 4.3	2.8, 2.1	3.2, 3.4	3.3, 0
KB	0, 8.1	0, 18.5	37.3, 3.7	37.4, 8.2	**25.4, 24.7**	43.9, 9.6	38.7, 15.4	43.8, 11.7
$K\hat{B}$	0, 8.1	0, 18.4	0.8, 3.4	1.2, 7.7	10, 43.5	10.3, 10.5	12, 23.7	11.6, 11.1
$\hat{K}B$	0, 8.2	0, 18.1	14.3, 3.8	15, 8.4	16.1, 38.3	22.9, 9.7	15.3, 16.9	23.8, 11.7
$\hat{K}\hat{B}$	0, 8.1	0, 18.3	0, 3.7	0, 8.3	10.4, 44.5	11.7, 11	12.4, 22.7	11.7, 12.7

TABLE D.16: S_{KB} Four Scenarios : Environment Results

	RF	ProA	Type	Value	EQ1	Value 2	EQ2
RW	0.2	0.1	A Utility	2.3		2.39	
			A Market Share	0.31		0.3	
			A Assurance	0.21	(K, K, 0.50)	0.15	($\hat{B}$, $\hat{B}$, 0.47)
			A Profit	11.88	(B, K, 0.26)	11.45	($\hat{B}$, $\hat{K}$B, 0.37)
			B Utility	5.79	(K, B, 0.15)	6.7	($\hat{K}$B, $\hat{B}$)
			B Market Share	0.69	(B, B)	0.62	($\hat{K}$B, $\hat{K}$B)
			B Assurance	0.21		0.15	
			B Profit	27.73		27.37	
EU	0.8	0.1	A Utility	2.3		2.3	
			A Market Share	0.31		0.31	
			A Assurance	0.21	(K, K, 0.50)	0.21	(K,K, 0.51)
			A Profit	11.87	(B, K, 0.26)	11.87	(B, K, 0.28)
			B Utility	5.79	(K, B, 0.16)	5.79	(K, B, 0.13)
			B Market Share	0.69	(B, B)	0.69	(B, B)
			B Assurance	0.21		0.21	
			B Profit	27.71		27.71	
IU	0.2	0.5	A Utility	17.89	(KB, KB, 1.0)	6.71	($\hat{K}B$, $\hat{K}\hat{B}$, 1.0)
			A Market Share	0.5		0.61	
			A Assurance	0.42		0.16	
			A Profit	42.24		24.6	
			B Utility	17.9		5.52	
			B Market Share	0.5		0.37	
			B Assurance	0.42		0.14	
			B Profit	42.23		22.93	
HG	0.8	0.5	A Utility	17.88	(KB, KB, 1.0)	17.88	(KB, KB, 1.0)
			A Market Share	0.5		0.5	
			A Assurance	0.42		0.42	
			A Profit	42.22		42.22	
			B Utility	17.86		17.86	
			B Market Share	0.5		0.5	
			B Assurance	0.42		0.42	
			B Profit	42.18		42.18	

$S_{KB} = \{\{K\}, \{\hat{K}\}, \{B\}, \{\hat{B}\}, \{K, B\}, \{\hat{K}, B\}, \{K, \hat{B}\}, \{\hat{K}, \hat{B}\}\}$. % $<$ 10 not shown.

TABLE D.17: Service Policy Characteristics

$s \in S$	Overt $\hat{s}$	Security	Revenue
–	–	0_0	0_0
K	K	1_1	0_0
B	B	2_2	0_0
H	–	3_0	0_0
KB	KB	3_3	0_0
KH	K	4_1	0_0
BH	B	5_2	0_0
KBH	KB	6_3	0_0
$\hat{K}B$	KB	2_3	1_0
$\hat{B}H$	B	3_2	2_0
$\hat{K}H$	K	3_1	1_0
$K\hat{B}H$	KB	4_2	2_0
$\hat{K}BH$	KB	5_3	1_0
$KB\hat{H}$	KB	3_3	3_0
$\hat{K}\hat{B}H$	KB	3_3	3_0
$K\hat{B}$	KB	1_3	2_0
$K\hat{H}$	K	1_1	3_0
$B\hat{H}$	B	2_2	3_0
$\hat{K}B\hat{H}$	KB	2_3	4_0
$K\hat{B}\hat{H}$	KB	1_3	5_0
$\hat{K}$	K	0_1	1_0
$\hat{B}$	B	0_2	2_0
$\hat{H}$	–	0_3	3_0
$\hat{K}\hat{B}$	KB	0_3	3_0
$\hat{K}\hat{H}$	K	0_1	4_0
$\hat{B}\hat{H}$	B	0_2	5_0
$\hat{K}\hat{B}\hat{H}$	KB	0_3	6_0

These policies indicate the full set of a service's discrete pure strategy options. For completeness, the subscripts indicates the overt levels of $\Psi\&\Phi$, however, note that these are all drawn from the first three strategies, K, B, KB.

References

Acquisti, A. (2004a). Privacy and security of personal information. In Economics of Information Security in Advances in Information Security, volume 12, pages 179–186. Springer.

Acquisti, A. (2004b). Privacy in electronic commerce and the economics of immediate gratification. In 5th ACM conference on Electronic Commerce, page 21, New York, New York, USA. ACM Press.

Acquisti, A. (2012). Nudging privacy: The behavioral economics of personal information. Digital Enlightenment Yearbook 2012, pages 193–197.

Acquisti, A. (2014). The Economics of Privacy. The Economics of Privacy - Resources on financial privacy, economics, anonymity, page 16.

Acquisti, A. and Grossklags, J. (2003). Losses, gains, and hyperbolic discounting: An experimental approach to information security attitudes and behavior. In 2nd Annual Workshop on "Economics and Information Security", volume 3, pages 1–27. Citeseer.

Acquisti, A. and Grossklags, J. (2004). Privacy attitudes and privacy behavior. In Economics of Information Security, pages 1–15. Springer.

Acquisti, A. and Grossklags, J. (2007). What Can Behavioral Economics Teach Us about Privacy ? In Digital Privacy, pages 363–377. Auerbach Publications.

Acquisti, A., John, L., and Loewenstein, G. (2012). The impact of relative standards on the propensity to disclose. Journal of Marketing Research, 49(2):160–174.

Adjerid, I., Acquisti, A., Brandimarte, L., and Loewenstein, G. (2013). Sleights of privacy. In Proceedings of the Ninth Symposium on Usable Privacy and Security - SOUPS '13, page 1. ACM.

Afroz, S., Islam, A. C., Santell, J., Chapin, A., and Greenstadt, R. (2013). How privacy flaws affect consumer perception. In Workshop on Socio-Technical Aspects in Security and Trust, STAST, pages 10–17. IEEE.

Aïmeur, E. and Schonfeld, D. (2011). The ultimate invasion of privacy: Identity theft. In 2011 9th Annual International Conference on Privacy, Security and Trust, PST 2011, pages 24–31. IEEE.

Akerlof, G. A. (1970). The market for" lemons": Quality uncertainty and the market mechanism. The quarterly journal of economics, pages 488–500.

Alhadeff Jones, M. (2008). Three generations of complexity theories: Nuances and ambiguities. Educational Philosophy and Theory, 40(1):66–82.

Anderson, J. G. (2007). Social, ethical and legal barriers to E-health. International Journal of Medical Informatics, 76(5-6):480–483.

Anderson, K. B., Durbin, E., Salinger, M. A., and Anderson, B. (2014). Identity Theft. The Journal of Economic Perspectives, 22(2):171–192.

Anne Toth, William Hoffman, E. K. T. D. and Lin, J. (2018). Data policy in the fourth industrial revolution: Insights on personal data. Technical report, World Economic Forum.

Attrill, A. and Jalil, R. (2011). Revealing only the superficial me: Exploring categorical self-disclosure online. Computers in Human Behavior, 27(5):1634–1642.

Axelrod, R. (1997). Advancing the art of simulation in the social sciences. In Simulating social phenomena, pages 21–40. Springer.

Baarslag, T., Alan, A. T., Gomer, R., Alam, M., Perera, C., Gerding, E. H., and schraefel, m. (2017). An automated negotiation agent for permission management. In Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems, AAMAS '17, pages 380–390, Richland, SC. International Foundation for Autonomous Agents and Multiagent Systems.

Balebako, R., Jung, J., Lu, W., Cranor, L. F., and Nguyen, C. (2013). "Little brothers watching you". In Proceedings of the Ninth Symposium on Usable Privacy and Security - SOUPS '13, page 1. ACM.

Balebako, R., Leon, P. G., Almuhimedi, H., Kelley, P. G., Mugan, J., Acquisti, A., Cranor, L. F., and Sadeh, N. (2011). Nudging users towards privacy on mobile devices. In CEUR Workshop Proceedings, volume 722, pages 23–26.

Barandiaran, X. and Moreno, A. (2006). ALife models as epistemic artefacts. In Artificial Life X: Proceedings of the Tenth International Conference on the Simulation and Synthesis of Living Systems, pages 513–519. The MIT Press Bradford Books Cambridge, MA.

Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. New Media & Society, 11(6):985–1002.

Beres, Y., Baldwin, A., Mont, M. C., and Shiu, S. (2007). On identity assurance in the presence of federated identity management systems. In Proceedings of the 2007 ACM workshop on Digital identity management - DIM '07, number 1, page 27, New York, New York, USA. ACM Press.

Bhargav-Spantzel, A., Squicciarini, A. C., Modi, S., Young, M., Bertino, E., and Elliott, S. J. (2006). Privacy preserving multi-factor authentication with biometrics. Proceedings of the second ACM workshop on Digital identity management, 15(5):63–72.

Bizer, C., Heath, T., and Berners-Lee, T. (2009). Linked data-the story so far. Semantic Services, Interoperability and Web Applications: Emerging Concepts, pages 205–227.

Black, S. M., Creese, S., Guest, R. M., Pike, B., Saxby, S. J., Fraser, D. S., Stevenage, S. V., Whitty, M. T., Stanton Fraser, D., Stevenage, S. V., and Whittty, M. T. (2012). Superidentity: Fusion of identity across real and cyber domains. ID360: Global Identity.

Böhme, R., Koble, S., and Dresden, T. U. (2007). On the viability of privacy-enhancing technologies in a self-regulated business-to-consumer market: Will privacy remain a luxury good? In WEIS.

Bohn, J., Coroama, V., Langheinrich, M., Mattern, F., and Rohs, M. (2005). Social, economic, and ethical implications of ambient intelligence and ubiquitous computing. In Ambient Intelligence, pages 5–29. Springer.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences, 99(suppl 3):7280–7287.

Bonneau, J., Anderson, J., Anderson, R., and Stajano, F. (2009). Eight friends are enough. In Proceedings of the Second ACM EuroSys Workshop on Social Network Systems - SNS '09, pages 13–18. ACM.

Boult, T. E., Scheirer, W. J., and Woodwork, R. (2007). Revocable fingerprint biotokens: Accuracy and security analysis. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 1–8. IEEE.

Brandstätter, E., Gigerenzer, G., and Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. Psychological Review, 113(2):409–432.

Brede, M. and Boschetti, F. (2009). Commons and anticommons in a simple renewable resource harvest model. ecological complexity, 6(1):56–63.

Brighton, H. (2006). Robust Inference with Simple Cognitive Models. In AAAI spring symposium: Cognitive science principles meet AI-hard problems, number Hutter, pages 17–22.

Brown, B., Chui, M., and Manyika, J. (2011). Are you ready for the era of ‘big data’. McKinsey Quarterly, 4(1):24–35.

Brunton, F. and Nissenbaum, H. (2015). Obfuscation: A User's Guide for Privacy and Protest. Mit Press.

Buckland, M. K. (1991). Information as thing. Journal of the American Society for Information Science (1986-1998), 42(5):351.

Bullock, S. (2014). Levins and the lure of artificial worlds. The Monist, 97(3):301–320.

Bünz, B., Agrawal, S., Zamani, M., and Boneh, D. (2020). Zether: Towards privacy in a smart contract world. In International Conference on Financial Cryptography and Data Security, pages 423–443. Springer.

Cabral, L. M. and Villas-Boas, M. (2005). Bertrand supertraps. Management Science, 51(4):599–613.

Cavoukian, A. (2008). Privacy in the clouds. Idis, 1(December):89–108.

Cavoukian, A. (2011). Privacy by Design: Origins, Meaning, and Prospects. Privacy Protection Measures and Technologies in Business Organizations: Aspects and Standards: Aspects and Standards, 170.

Chikkerur, S. (2008). Generating Registration-free Cancelable Fingerprint Templates. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 29(4):0–5.

Choe, E. K., Jung, J., Lee, B., and Fisher, K. (2013). Nudging people away from privacy-invasive mobile apps through visual framing. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 8119 LNCS, pages 74–91. Springer.

Cialdini, R. and Trost, M. (1998). Social influence: Social norms, conformity and compliance. In The Handbook of Social Psychology, Vol. 2, pages 151–192.

CIFAS (2017). Fraudscape 2017: External and internal fraud threats. Technical report.

Cifas (2019). Fraudscape 2019: Idenity fraud and money mules rise again.

Cioffi-Revilla, C. (2010). A methodology for complex social simulations. Journal of Artificial Societies and Social Simulation, 13(1):7.

Cioffi-Revilla, C. (2014). Computation and Social Science. In Introduction to Computational Social Science, pages 23–66. Springer.

Clarke, R. (1988). Information technology and dataveillance. Communications of the ACM, 31(5):498–512.

Clarke, R. and Wigan, M. (2011). You are where you've been: the privacy implications of location and tracking technologies. Journal of Location Based Services, 5(3-4):138–155.

Conger, S., Pratt, J. H., and Loch, K. D. (2013). Personal information privacy and emerging technologies. Information Systems Journal, 23(5):401–417.

Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertesz, J., Loreto, V., Moat, S., Nadal, J.-P., and Sanchez, A. (2012). Manifesto of computational social science. The European Physical Journal Special Topics, 214(1):325–346.

Cook, D. J., Augusto, J. C., and Jakkula, V. R. (2009). Ambient intelligence: Technologies, applications, and opportunities. Pervasive and Mobile Computing, 5(4):277–298.

Correa, D., Silva, L., Mondal, M., Benevenuto, F., and Gummadi, K. (2015). The many shades of anonymity: Characterizing anonymous social media content. In Proceedings of the International AAAI Conference on Web and Social Media, volume 9.

Coulom, R. (2007). Computing elo ratings of move patterns in the game of go. In Computer games workshop.

Coventry, L. M., Jeske, D., Blythe, J. M., Turland, J., and Briggs, P. (2016). Personality and Social Framing in Privacy Decision-Making: A Study on Cookie Acceptance. Frontiers in Psychology, 7.

Cozby, P. C. (1972). Self-disclosure, reciprocity and liking. Sociometry, 35(1):151–160.

Craciun, G. (2018). Choice defaults and social consensus effects on online information sharing: The moderating role of regulatory focus. Computers in Human Behavior, 88:89–102.

Culnan, M. J. and Armstrong, P. K. (1999). Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation. Organization Science, 10(1):104–115.

Das, S., Kramer, A. D., Dabbish, L. A., and Hong, J. I. (2015). The role of social influence in security feature adoption. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work &#38; Social Computing, CSCW '15, pages 1416–1426, New York, NY, USA. ACM.

Davidsson, P. (2000). Multi agent based simulation: beyond social simulation. In Multi-Agent-Based Simulation, pages 97–107. Springer.

Debatin, B., Lovejoy, J. P., Horn, A. K., and Hughes, B. N. (2009). Facebook and online privacy: Attitudes, behaviors, and unintended consequences. Journal of Computer-Mediated Communication, 15(1):83–108.

Deneckere, R., Kovenock, D., and Lee, R. (1992). A model of price leadership based on consumer loyalty. The Journal of Industrial Economics, pages 147–156.

Dhamija, R. and Dusseault, L. (2008). The seven flaws of identity management: Usability and security challenges. IEEE Security and Privacy, 6(2):24–29.

Dhamija, R., Tygar, J. D., and Hearst, M. (2006). Why phishing works. In Proceedings of the SIGCHI conference on Human Factors in computing systems, pages 581–590. ACM.

Di Paolo, E. A., Noble, J., and Bullock, S. (2000). Simulation models as opaque thought experiments. In Artificial Life VII: The Seventh International Conference on the Simulation and Synthesis of Living Systems, pages 497–506.

Dickenson, H. (2015). Cross-Sector Fraud Findings Report. Technical report, CIFAS.

Donna L. Hoffman, Thomas P. Novak, (1999). Information Privacy in the Marketspace: Implications for the Commercial Uses of Anonymity on the Web. The Information Society, 15(2):129–139.

Eaton, J. J. and Bawden, D. (1991). What kind of resource is information? International journal of information management, 11(2):156–165.

Eckersley, P. (2010). How Unique Is Your Web Browser? In Proc. of the Privacy Enhancing Technologies Symposium (PETS), pages 1–18. Springer.

Edmonds, B. (2017). Different Modelling Purposes, pages 39–58. Springer International Publishing, Cham.

Edmonds, B. and Gershenson, C. (2015). Modelling complexity for policy: opportunities and challenges. Handbook on Complexity and Public Policy, page 205.

Edwards, W., Hopkins, T. J., Hop, T. J., Flood, M. M., Papandreou, A., Savage, L. J., and Coombs, C. H. (1954). The theory of decision making. Psychological Bulletin, 51(4):380–417.

Elo, A. E. (1978). The Rating of Chess Players, Past and Present. Arco Pub.

Emanuel, L., Neil, G. J., Bevan, C., Fraser, D. S., Stevenage, S. V., Whitty, M. T., and Jamison-Powell, S. (2014). Who am I? Representing the self offline and in different online contexts. Computers in Human Behavior, 41:146–152.

Epstein, J. M. (1999). Agent-based computational models and generative social science. Complexity, 4(5):41–60.

Epstein, J. M. (2008). Why model? Journal of Artificial Societies and Social Simulation, 11(4):12.

Faiks, A. and Hyland, N. (2000). Gaining user insight: A case study illustrating the card sort technique. College and Research Libraries, 61(4):349.

Fairfield, J. A. T. and Engel, C. (2015). Privacy as a public good. Duke LJ, 65:385.

Fereday, J. and Muir-Cochrane, E. (2006). Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. International Journal of Qualitative Methods, 5(1):80–92.

Fischer-Hübner, S., Hoofnagle, C. J., Krontiris, I., Rannenberg, K., and Waidner, M. (2011). Online Privacy: Towards Informational Self-Determination on the Internet. Dagstuhl Manifestos, 1(1):1–20.

Flanagin, A. J., Flanagin, C., and Flanagin, J. (2010). Technical code and the social construction of the internet. New Media & Society, 12(2):179–196.

Foddy, W. H. and Finighan, W. R. (1980). The Concept of Privacy from a Symbolic Interaction Perspective. Journal for the Theory of Social Behaviour, 10(1):1–18.

Fogg, B. J. (2003). Prominence-interpretation theory: Explaining how people assess credibility online. In Conference on Human Factors in Computing Systems - Proceedings, pages 722–723. ACM.

Fogg, B. J., Soohoo, C., Danielson, D. R., Marable, L., Stanford, J., and Tauber, E. R. (2003). How Do Users Evaluate the Credibility of Web Sites? A Study with Over 2,500 Participants. In Proceedings of the 2003 conference on Designing for user experiences (DUX'03), pages 1–15. ACM.

Foley, L., Barney, K., Gordon, S., Foley, J., Lee, J., and Fergerson, J. (2009). Identity theft: The aftermath 2008.

Frank, U., Squazzoni, F., and Troitzsch, K. G. (2009). EPOS-Epistemological perspectives on simulation: an introduction. Epistemological Aspects of Computer Simulation in the Social Sciences, page 1.

Fridman, I. and Higgins, E. T. (2017). Regulatory focus and regulatory fit in health messaging. In Oxford Research Encyclopedia of Communication.

Friedewald, M., Vildjiounaite, E., Punie, Y., and Wright, D. (2007). Privacy, identity and security in ambient intelligence: A scenario analysis. Telematics and Informatics, 24(1):15–29.

Frigg, R. and Reiss, J. (2009). The philosophy of simulation: hot new issues or same old stew? Synthese, 169(3):593–613.

Furnell, S. and Phippen, A. (2012). Online privacy: A matter of policy? Computer Fraud and Security, 2012(8):12–18.

Galán, J. M., Izquierdo, L. R., Izquierdo, S. S., Santos, J. I., del Olmo Martínez, R., López-Paredes, A., and Edmonds, B. (2009). Errors and artefacts in agent-based modelling. J. Artificial Societies and Social Simulation, 12.

Gambino, A., Kim, J., Sundar, S. S., Ge, J., and Rosson, M. B. (2016). User Disbelief in Privacy Paradox: Heuristics that determine Disclosure. In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, pages 2837–2843. ACM.

Gaw, S. and Felten, E. W. (2006). Password management strategies for online accounts. In Proceedings of the second symposium on Usable privacy and security - SOUPS '06, page 44. ACM.

Gigerenzer, G. (1999). Simple Heuristics That Make Us Smart. In Simple Heuristics That Make Us Smart, pages 3–34. Oxford University Press.

Gigerenzer, G. and Gaissmaier, W. (2011). Heuristic decision making. Annual Review of Psychology, 62:451–482.

Gigerenzer, G., Hoffrage, U., and Goldstein, D. G. (2008). Fast and frugal heuristics are plausible models of cognition: reply to Dougherty, Franco-Watkins, and Thomas (2008). Psychological review, 115(1):230–239.

Gigerenzer, G. and Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In Simple heuristics that make us smart, pages 3–34. Oxford University Press.

Goodman, M. (2015). Future crimes: Everything is connected, everyone is vulnerable and what we can do about it. Anchor.

Goodspeed, R. (2017). Research note: An evaluation of the elo algorithm for pairwise visual assessment surveys. Landscape and Urban Planning, 157:131–137.

Gray, J. and Bullock, S. (2014). Deciding to Disclose : Pregnancy and Alcohol Misuse. (2006):1–4.

Grossklags, J., Hall, S., and Acquisti, A. (2007). When 25 Cents is too much : An Experiment on Willingness-To-Sell and Willingness-To-Protect Personal Information. In Information Security, pages 7–8.

Guest, R., Miguel-Hurtado, O., Stevenage, S. V., Neil, G. J., and Black, S. (2014). Biometrics within the SuperIdentity project: a new approach to spanning multiple identity domains. In 2014 International Carnahan Conference on Security Technology (ICCST), pages 1–6. IEEE.

Halamish, V., Liberman, N., Higgins, E. T., and Idson, L. C. (2008). Regulatory focus effects on discounting over uncertainty for losses vs. gains. Journal of Economic Psychology, 29(5):654–666.

Hall, S., Crowe, E., and Higgins, E. T. (1997). Regulatory focus and strategic inclinations: Promotion and prevention in decision-making. Organizational Behavior and Human Decision Processes, 69(2):117–132.

Halperin, R. and Backhouse, J. (2008). A roadmap for research on identity in the information society. Identity in the Information Society, 1(1):14–15.

Hansen, P. G. and Jespersen, A. M. (2013). Nudge and the Manipulation of Choice: A Framework for the Responsible Use of the Nudge Approach to Behaviour Change in Public Policy. European Journal of Risk Regulation, 1:3–28.

Hardin, G. (1968). The tragedy of the commons. science, 162(3859):1243–1248.

Hardin, G. (2009). The Tragedy of the Commons. Journal of Natural Resources Policy Research, 1(3):243–253.

Heikkinen, A., Wickström, G., and Leino-Kilpi, H. (2006). Understanding Privacy in Occupational Health Services. Nursing Ethics, 13(5):515–530.

Heller, M. A. (1998). The tragedy of the anticommons: property in the transition from Marx to markets. Harvard law review, pages 621–688.

Hess, C. and Ostrom, E. (2003). Ideas, artifacts, and facilities: information as a common-pool resource. Law and contemporary problems, 66(1/2):111–145.

Higgins, E. (1998). Promotion and prevention. Regulatory focus as a motivational principle.pdf. Advances in Experimental Social Psychology, 30:1–46.

Higgins, E. T., Nakkawita, E., and Cornwell, J. F. (2020). Beyond outcomes: How regulatory focus motivates consumer goal pursuit processes. Consumer Psychology Review, 3(1):76–90.

Hildebrandt, M. (2009). The Future of Identity in the Information Society Challenges and Opportunities. Springer Science & Business Media.

Hollingsed, T. and Novick, D. G. (2007). Usability inspection methods after 15 years of research and practice. In Proceedings of the 25th annual ACM international conference on Design of communication, pages 249–255. ACM.

Holloway, I. (1997). Basic concepts for qualitative research. Wiley-Blackwell.

Hong, J. (2012). The state of phishing attacks. Communications of the ACM, 55(1):74–81.

Hoofnagle, C. J. (2007). Identity Theft: Making the Known Unknowns Known. Harvard Journal of Law & Technology, 21:98–122.

Humphreys, P. (2009). The Philosophical Novelty of Computer Simulation Methods.

Hvattum, L. M. and Arntzen, H. (2010). Using ELO ratings for match result prediction in association football. International Journal of forecasting, 26(3):460–470.

Idson, L. C., Liberman, N., and Higgins, E. (2000). Distinguishing Gains from Nonlosses and Losses from Nongains: A Regulatory Focus Perspective on Hedonic Intensity. Journal of Experimental Social Psychology, 36(3):252–274.

Inglesant, P. G. and Sasse, M. A. (2010). The true cost of unusable password policies. In Proceedings of the 28th international conference on Human factors in computing systems - CHI '10, page 383. ACM.

Jain, A. K., Bolle, R., and Pankanti, S. (1999). Biometrics: personal identification in networked society, volume 1. Springer Science & Business Media.

Jain, A. K. and Nandakumar, K. (2012). Biometric authentication: System security and user privacy. Computer, 45(11):87–92.

Jain, A. K., Ross, A., and Prabhakar, S. (2004). An Introduction to Biometric Recognition 1. IEEE Transactions on Circuits and Systems for Video Technology, 14(1):4–20.

Jamieson, K. G. and Nowak, R. D. (2011). Active ranking using pairwise comparisons. arXiv preprint arXiv:1109.3701.

Jin, S.-A. A. (2013). Peeling back the multiple layers of Twitter's private disclosure onion: The roles of virtual identity discrepancy and personality traits in communication privacy management on Twitter. New Media & Society, page 1461444812471814.

Johnson, R. E., King, D. D., Lin, S.-H. J., Scott, B. A., Walker, E. M. J., and Wang, M. (2017). Regulatory focus trickle-down: How leader regulatory focus and behavior shape follower regulatory focus. Organizational Behavior and Human Decision Processes, 140:29–45.

Kahn, C. M. and Roberds, W. (2008). Credit and identity theft. Journal of Monetary Economics, 55(2):251–264.

Kahneman, D. (2011). Thinking, fast and slow. Macmillan.

Kallemeyn, L. M., Hall, J. N., and Gates, E. (2020). Exploring the relevance of complexity theory for mixed methods research. Journal of Mixed Methods Research, 14(3):288–304.

Katos, V., Stowell, F., and Bednar, P. M. (2010). Surveillance, Privacy and the Law of Requisite Variety. In Dpm/Setop, volume 6514, pages 123–139. Springer.

Kehr, F., Wentzel, D., and Kowatsch, T. (2015). Rethinking Privacy Decisions : Pre-Existing At- Titudes , Pre-Existing Emotional States , and a Situational Privacy Calculus. In ECIS 2015 Proceedings, pages 1–15.

Kehr, F., Wentzel, D., and Mayer, P. (2013). Rethinking the Privacy Calculus: On the Role of Dispositional Factors and Affect. The 34th International Conference on Information Systems, (1):1–10.

Klügl, F. (2008). A validation methodology for agent-based simulations. In Proceedings of the 2008 ACM symposium on Applied computing, pages 39–43.

Knight, A. and Saxby, S. (2014). Identity crisis: Global challenges of identity protection in a networked world. Computer Law & Security Review, 30(6):617–632.

Knijnenburg, B. P. (2013). On The Dimensionality Of Information Disclosure Behavior in Social Networks. International Journal of Human-Computer Studies, 71(12):1144–1162.

Knijnenburg, B. P. and Kobsa, A. (2013a). Making Decisions About Privacy: Information Disclosure in Context-Aware Recommender Systems. ACM Trans. Interact. Intell. Syst., 3(3):20:1—-20:23.

Knijnenburg, B. P. B. and Kobsa, A. (2013b). Helping users with information disclosure decisions: potential for adaptation. In Proceedings of the 2013 international conference on Intelligent user interfaces - IUI '13, volume 3, page 407. ACM.

Komanduri, S., Shay, R., Kelley, P. G., Mazurek, M. L., Bauer, L., Christin, N., Cranor, L. F., and Egelman, S. (2011). Of Passwords and People. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11), pages 2595–2604. ACM.

Kosba, A., Miller, A., Shi, E., Wen, Z., and Papamanthou, C. (2016). Hawk: The blockchain model of cryptography and privacy-preserving smart contracts. In 2016 IEEE symposium on security and privacy (SP), pages 839–858. IEEE.

Kosinski, M., Stillwell, D., and Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. Proceedings of the National Academy of Sciences of the United States of America, 110(15):5802–5.

Krasnova, H. and Günther, O. (2009). Privacy concerns and identity in online social networks. Identity in the Information Society, 2(1):39–63.

Krasnova, H., Spiekermann, S., Koroleva, K., and Hildebrand, T. (2010). Online social networks: Why we disclose. Journal of Information Technology, 25(2):109–125.

Kruger, J., Wirtz, D., Van Boven, L., and Altermatt, T. W. (2004). The effort heuristic. Journal of Experimental Social Psychology, 40(1):91–98.

Kruglanski, a. W., Thompson, E. P., Higgins, E. T., Atash, M. N., Pierro, a., Shah, J. Y., and Spiegel, S. (2000). To "do the right thing" or to "just do it": locomotion and assessment as distinct self-regulatory imperatives. Journal of personality and social psychology, 79(5):793–815.

Lancaster, W. and Lancaster, J. (1982). Rational Decision Making. JONA: The Journal of Nursing Administration, 12(9):23???28.

Lash, S. (2007). Power after Hegemony: Cultural Studies in Mutation? Theory, Culture & Society, 24(3):55–78.

Laudon, K. C. (1996). Markets and privacy. Communications of the ACM, 39(9):92–104.

Leary, M. R. and Allen, A. B. (2011). Self-presentational persona: simultaneous management of multiple impressions. Journal of personality and social psychology, 101(5):1033–49.

Lee, M. D. and Wagenmakers, E.-J. (2005). Bayesian statistical inference in psychology: comment on Trafimow (2003). Psychological review, 112(3):662–668; discussion 669–674.

Legendre, P. (2005). Species associations: the Kendall coefficient of concordance revisited. Journal of agricultural, biological, and environmental statistics, 10(2):226–245.

Lin, C. L., Sun, H. M., and Hwang, T. (2001). Attacks and solutions on strong-password authentication. IEICE Transactions on Communications, E84-B(9):2622–2627.

Lipshitz, R., Klein, G., Orasanu, J., and Salas, E. (2001). Taking stock of naturalistic decision making. J. Behav. Decis. Mak., 14(5):331–352.

Lodge, J. (2007). Biometrics: A Challenge For Privacy Or Public Policy-Certified Identity And Uncertainties. Minority, Politics, Society, 1:193–206.

Loewenstein, G. F., Hsee, C. K., Weber, E. U., and Welch, N. (2001). Risk As Feelings. Psychological bulletin, 127(2):267.

LoPucki, L. M. (2002). Did privacy cause identity theft. Hastings Lj, 54:1277.

Macal, C. M. and North, M. J. (2010). Tutorial on agent-based modelling and simulation. Journal of Simulation, 4(3):151–162.

MacCarthy, M. (2010). New directions in privacy: Disclosure, unfairness and externalities. ISJLP, 6:425.

MacCrimmon, K. and Larsson, S. (1979). Utility theory: Axioms versus paradoxes. In Expected utility hypotheses and the Allais paradox, volume 21, pages 333–409. Springer.

Manshaei, M. H. (2011). Game Theory Meets Network Security and Privacy. ACM Computing Surveys (CSUR), 45(December):1–45.

Marmion, V., Bishop, F., Millard, D. E., and Stevenage, S. V. (2017a). The Cognitive Heuristics Behind Disclosure Decisions, pages 591–607. Springer International Publishing.

Marmion, V., Millard, D. E., Gerding, E. H., and Stevenage, S. V. (2017b). The tragedy of the identity assurance commons. In Proceedings of the 2017 ACM on Web Science Conference, WebSci '17, pages 397–398, New York, NY, USA. ACM.

Marmion, V., Millard, D. E., Gerding, E. H., and Stevenage, S. V. (2019). The willingness of crowds: Cohort disclosure preferences for personally identifying information. In Proceedings of the International AAAI Conference on Web and Social Media, volume 13, pages 358–368.

Martin, K. and Nissenbaum, H. (2017). Privacy interests in public records: An empirical investigation. Harv. JL & Tech., 31:111.

Metzger, M. J. (2006). Privacy, Trust, and Disclosure: Exploring Barriers to Electronic Commerce. Journal of Computer-Mediated Communication, 9(4):1–29.

Metzger, M. J. and Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. Journal of Pragmatics, 59:210–220.

Metzger, M. J., Flanagin, A. J., and Medders, R. B. (2010). Social and Heuristics Approaches to Credibility Evaluation Online. Journal of Communication, 60(3):413–439.

Millen, D. R. and Patterson, J. F. (2003). Identity disclosure and the creation of social capital. In CHI'03 extended abstracts on Human factors in computing systems, pages 720–721. ACM.

Mitchell, M. (2009). Complexity: A guided tour. Oxford University Press.

Morin, E. (1992). From the concept of system to the paradigm of complexity. Journal of social and evolutionary systems, 15(4):371–385.

Murtagh, F. and Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: Which algorithms implement ward's criterion? Journal of Classification, 31(3):274–295.

Nandakumar, K., Jain, A. K., and Nagar, A. (2008). Biometric template security. Eurasip Journal on Advances in Signal Processing, 2008:113.

Newman, G. R. and Mcnally, M. M. (2005). Identity theft literature review. Citeseer.

Nielsen, J. (1994). Usability inspection methods. In Conference companion on Human factors in computing systems, pages 413–414. ACM.

Niinuma, K. and Jain, A. (2010). Continuous user authentication using temporal information. In SPIE Defense, Security, and . . . , page 11. International Society for Optics and Photonics.

Nisan, N. (2015). Algorithmic mechanism design: Through the lens of multiunit auctions. In Handbook of Game Theory with Economic Applications, volume 4, pages 477–515. Elsevier.

Nissenbaum, H. (2004). Privacy as contextual integrity. Wash. L. Rev., 79:119.

Norberg, P. A., Horne, D. R., and Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. Journal of Consumer Affairs, 41(1):100–126.

Office, C. (2019). National cyber security strategy: Progress report. Technical report, Cabinet Office and National security and intelligence.

Office, I. C. (2010). Ico report 2010.

O'Hara, K. (2010). Intimacy 2.0: Privacy Rights and Privacy Responsibilities on the World Wide Web. World Wide Web Internet And Web Information Systems.

Olivero, N. and Lunt, P. (2004). Privacy versus willingness to disclose in e-commerce exchanges: The effect of risk awareness on the relative role of trust and control. Journal of Economic Psychology, 25(2):243–262.

Ostrom, E. (2008). Tragedy of the commons. The new palgrave dictionary of economics, 2.

Ostrom, E. (2014). Collective action and the evolution of social norms. Journal of Natural Resources Policy Research, 6(812):235–252.

Parliament, B. (1998). Data protection act of 1998.

Perez, B., Musolesi, M., and Stringhini, G. (2018). You are your metadata: Identification and obfuscation of social media users using metadata information.

Pollard, J. (2006). Knowing Capitalismt, volume 37. Sage.

Preibusch, S., Krol, K., and Beresford, A. R. (2013). The privacy economics of voluntary over-disclosure in web forms. In The Economics of Information Security and Privacy, pages 183–209. Springer.

Preuveneers, D. and Joosen, W. (2015). Smartauth: Dynamic context fingerprinting for continuous user authentication. In Proceedings of the 30th Annual ACM Symposium on Applied Computing, pages 2185–2191. ACM.

Ratha, N., Connell, J., Bolle, R. M., and Chikkerur, S. (2006). Cancelable biometrics: A case study in fingerprints. In Proceedings - International Conference on Pattern Recognition, volume 4, pages 370–373. IEEE.

Ratha, N. K., Chikkerur, S., Connell, J. H., and Bolle, R. M. (2007). Generating cancelable fingerprint templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(4):561–572.

Ratha, N. K., Connell, J. H., and Bolle, R. M. (2001). Enhancing security and privacy in biometrics-based authentication systems. IBM Systems Journal, 40(3):614–634.

Ratha, N. K. N., Bolle, R. R. M., Pandit, V. D. V., and Vaish, V. (2000). Robust fingerprint authentication using local structural similarity. In Fifth IEEE Workshop on Applications of Computer Vision, pages 29–34. IEEE.

Rathgeb, C. and Uhl, A. (2011). A survey on biometric cryptosystems and cancelable biometrics. EURASIP Journal on Information Security, 2011(1):3.

Resnick, M. (1996). Beyond the centralized mindset. The journal of the learning sciences, 5(1):1–22.

Resnick, M. (1997). Turtles, Termites, and Traffic Jams: Explorations in Massively Parallel Microworlds. Number 1997. Mit Press.

Roberts, C. (2007). Biometric attack vectors and defences. Computers and Security, 26(1):14–25.

Rose, J. and Kalapesi, C. (2012). Rethinking personal data: strengthening trust. BCG Perspectives, 16(05):2012.

Rust, R. T., Kannan, P. K., and Peng, N. (2002). The Customer Economics of Internet Privacy. Journal of the Academy of Marketing Science, 30(4):455–464.

Ryan, G. W. and Bernard, H. R. (2003). Techniques to Identify Themes. Field Methods, 15(1):85–109.

Saari, D. G. (1999). Explaining all three-alternative voting outcomes. Journal of Economic Theory, 87(2):313 – 355.

Sandström, M. (2004). Liveness detection in fingerprint recognition systems. page 128.

Sarma, A. D., Sarma, A. D., Gollapudi, S., and Panigrahy, R. (2010). Ranking Mechanisms in Twitter-like Forums.

Schauer, F. (1978). Fear, Risk and the First Amendment: Unraveling the Chilling Effect. Bul Rev., 58:685.

Schildkrout, E. (2004). Inscribing the body. Annual Review of Anthropology, 33(1):319–344.

Schwab, K., Marcus, A., Oyola, J. O., Hoffman, W., and Luzi, M. (2011). Personal data: The emergence of a new asset class. In An Initiative of the World Economic Forum.

Simoens, K., Tuyls, P., and Preneel, B. (2009). Privacy weaknesses in biometric sketches. In Proceedings - IEEE Symposium on Security and Privacy, pages 188–203. IEEE.

Simon, H. A. (1955). A Behavioral Model of Rational Choice. The Quarterly Journal of Economics, 69(1):99–118.

Simpson, W. R. and Foltz, K. E. (2017). Assured identity for enterprise level security. JSTOR.

Solove, D. J. (2007). 'I've Got Nothing to Hide' and Other Misunderstandings of Privacy. San Diego Law Review, 44(May):1–23.

Solove, D. J. (2012). Introduction: Privacy Self-Management and the Consent Dilemma. Harvard Law Review, 126:1880–1903.

Solove, D. J. and Schwartz, P. M. (2011). The PII Problem: Privacy and a New Concept of Personally Identifiable Information.

Squazzoni, F. (2014). A social science-inspired complexity policy: Beyond the mantra of incentivization. Complexity, 19(6):5–13.

Staddon, J., Huffaker, D., Brown, L., and Sedley, A. (2012). Are privacy concerns a turn-off?: engagement and privacy in social networks. In Proceedings of the Eighth Symposium on Usable Privacy and Security, pages 1–13. ACM.

Steeves, V. and Pinero, V. (2008). Privacy and Police Powers: Situating the Reasonable Expectation of Privacy Test. Canadian Journal of Criminology & Criminal Justice, 50(3):263–269.

Strack, F. and Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. Personality and social psychology review : an official journal of the Society for Personality and Social Psychology, Inc, 8(3):220–247.

Strahilevitz, L. (2013). Toward a positive theory of privacy law. Harvard Law Review, 113(1).

Strogatz, S. H. (2001). Nonlinear Dynamics And Chaos: With Applications To Physics, Biology, Chemistry, And Engineering. Westview press.

Sullivan, C. (2013). Digital identity, privacy and the right to identity in the United States of America. Computer Law and Security Review, 29(4):348–358.

Sullivan, C. L. (2010). Digital Identity, an Emergent Legal Concept. University of Adelaide Press.

Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. Digital Media, Youth, and Credibility, pages 73–100.

Sundar, S. S., Kang, H., Wu, M., Go, E., and Zhang, B. (2013). Unlocking the privacy paradox: do cognitive heuristics hold the key? CHI'13 Extended Abstracts on Human Factors in Computing Systems, pages 811–816.

Sundar, S. S., Kim, J., Rosson, M. B., and Molina, M. D. (2020). Online privacy heuristics that predict information disclosure. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pages 1–12.

Szczecinski, L. and Djebbi, A. (2019). Understanding and pushing the limits of the elo rating algorithm. arXiv preprint arXiv:1910.06081.

Taitsman, J. K., M, G. C., and Agrawal, S. (2013). Protecting Patient Privacy and Data Security Julie. New England Journal of Medicine, Electronic(11):977–979.

Tene, O. and Polonetsky, J. (2012). Privacy in the age of big data: a time for big decisions. Stanford Law Review Online, 64:63.

Tene, O. and Polonetsky, J. (2013a). A Theory of Creepy: Technology, Privacy and Shifting Social Norms. Yale Journal of Law & Technology (. . . , 16:1–32.

Tene, O. and Polonetsky, J. (2013b). Big data for all: Privacy and user control in the age of analytics. Northwestern Journal of Technology and Intellectual Property Volume, 11(5):240–273.

Tibbs, H. (2013). The Global Cyber Game: The Defence Academy Cyber Inquiry Report. Swindon, UK: Defence Academy of the United Kingdom, page 127.

Toubiana, V., Narayanan, A., Boneh, D., Nissenbaum, H., and Barocas, S. (2010). Adnostic : Privacy Preserving Targeted Advertising. In Proceedings of the NDSS Symposium 2010, pages 1–21.

Tsakalakis, N., Stalla-Bourdillon, S., and O'Hara, K. (2017). Identity assurance in the uk: technical implementation and legal implications under eidas. The Journal of Web Science, 3(3):32–46.

Tsesis, A. (2014). Right to Erasure: Privacy, Data Brokers, and the Indefinite Retention of Data, The. Wake Forest L. Rev., 49:433.

Turner, J. R. and Baker, R. M. (2019). Complexity theory: An overview with potential applications for the social sciences. Systems, 7(1):4.

Tversky, a. and Kahneman, D. (1973). Availability: A Heuristic for Juudging Frequency and robability. Cognitive Psychology, 5(2):207–232.

Tversky, A. and Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In Utility, probability, and human decision making, pages 141–162. Springer.

Tversky, A. and Kahneman, D. (1986). Rational Choice and the Framing of Decisions. Journal of Business, 59(4):S251–S278.

Tversky, A. and Kahneman, D. (1992). Advances in Prospect-Theory - Cumulative Representation of Uncertainty. Journal of Risk and Uncertainty, 5(4):297–323.

Udo, G. J. (2001). Privacy and security concerns as major barriers for e-commerce: a survey study. Information Management & Computer Security, 9(4):165–174.

UK-Cabinet-Office (2016). National Cyber Security Strategy 2016 to 2021. Technical report.

Uludag, U., Pankanti, S., Prabhakar, S., and Jain, A. K. (2004). Biometric cryptosystems: Issues and challenges. Proceedings of the IEEE, 92(6):948–959.

Ur, B., Leon, P. G., Cranor, L. F., Shay, R., and Wang, Y. (2012). Smart, Useful, Scary, Creepy: Perceptions of Online Behavioural Advertising. In Symposium on Usable Privacy and Security, page 4. ACM.

van der Velden, M. and El Emam, K. (2013). "Not all my friends need to know": a qualitative study of teenage patients, privacy, and social media. Journal of the American Medical Informatics Association : JAMIA, 20(1):16–24.

Van Noort, G., Kerkhof, P., and Fennis, B. M. (2008). The persuasiveness of online safety cues: The impact of prevention focus compatibility of Web content on consumers' risk perceptions, attitudes, and intentions. Journal of Interactive Marketing, 22(4):58–72.

Van Zoonen, L. and Turner, G. (2014). Exercising identity: agency and narrative in identity management. Kybernetes, 43(6):935–946.

Varian, H. (1996). Economic aspects of personal privacy. Topics in Regulatory Economics and Policy, (3):1–12.

Vidalis, S. and Olga, A. (2014). Assessing Identity Theft in the Internet of Things. IT CoNvergence PRActice (INPRA), 2(1):15–21.

Vila, T., Greenstadt, R., and Molnar, D. (2003). Why we can't be bothered to read privacy policies models of privacy economics as a lemons market. In Proceeding

ICEC '03 Proceedings of the 5th International Conference on Electronic Commerce, pages 403–407. ACM.

Vilà, X. (2008). A model-to-model analysis of bertrand competition. J. Artificial Societies and Social Simulation, 11.

Walby, S. (2007). Complexity theory, systems theory, and multiple intersecting social inequalities. Philosophy of the social sciences, 37(4):449–470.

Walters, W. and Betz, A. (2012). Medical identity theft. Journal of Consumer Education, page 75.

Ward, R. (2003). Physiological responses to different WEB page designs. International Journal of Human-Computer Studies, 59(1-2):199–212.

Warren, S. D. and Brandeis, L. D. (1890). The right to privacy. Harvard law review, pages 193–220.

Wayman, J. L. (2008). Biometrics in identity management systems. IEEE Security and Privacy, 6(2):30–37.

Weitzner, D. J., Abelson, H., Berners-Lee, T., Feigenbaum, J., Hendler, J., and Sussman, G. J. (2008). Information accountability. Communications of the ACM, 51(6):82–87.

Weitzner, D. J., Abelson, H., Hanson, C., Hendler, J., Mcguinness, D. L., Jay, G., Waterman, K. K., Berners-lee, T., Kagal, L., and Sussman, G. J. (2006). Transparent Accountable Data Mining: New Strategies for Privacy Protection. pages 1–12.

Westin, A. F. (2003). Social and political dimensions of privacy. Journal of social issues, 59(2):431–453.

Whiteside, S. P., Lynam, D. R., Miller, J. D., and Reynolds, S. K. (2005). Validation of the UPPS impulsive behaviour scale: a fourfactor model of impulsivity. European Journal of Personality, 19(7):559–574.

Whitley, E. A. (2009). Informational privacy, consent and the "control" of personal data. Information Security Technical Report, 14(3):154–159.

Whitney, S. and Mccullough, L. B. (2004). A Typology of Shared Decision Making , Informed Consent , and Simple Consent. Annals of Internal Medicine, 140(1):54–59.

Williams, M.-a. (2008). Privacy Management , The Law and Global Business Strategies : A Case for Privacy Driven Design. In Computational Science and Engineering, 2009. CSE'09. International Conference on, volume 3, pages 71–79. IEEE.

Wirtz, J. and Lwin, M. O. (2009). Regulatory focus theory, trust, and privacy concern. Journal of Service Research, 12(2):190–207.

World Economic Forum (2011). Personal data : The emergence of a new asset class. In Forum American Bar Association, pages 1–40.

Yakowitz, J. (2011). Tragedy of the data commons. Harv. JL & Tech., 25:1.

Yampolskiy, R. V. and Govindaraju, V. (2007). Direct and indirect human computer interaction based biometrics. Journal of computers, 2(10):76–88.

Yampolskiy, R. V. and Govindaraju, V. (2008). Behavioural biometrics: a survey and classification. International Journal of Biometrics, 1(1):81–113.

Yang, K., Du, E. Y., and Zhou, Z. (2013). Consent biometrics. Neurocomputing, 100:153–162.

Zafeiropoulou, A. M., Millard, D. E., Webber, C., and O'Hara, K. (2013). Unpicking the privacy paradox. In Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13, pages 463–472. ACM.

Zhang, L., Gupta, D., and Mohapatra, P. (2012). How expensive are free smartphone apps? ACM SIGMOBILE Mobile Computing and Communications Review, 16(3):21–32.

Zhu, F. and Zhu, W. (2009). Rational exposure: A game theoretic approach to optimize identity exposure in pervasive computing environments. In 2009 IEEE International Conference on Pervasive Computing and Communications, pages 1–8. IEEE.