WMG Service Systems Research Group
Working Paper Series

Homo Dataicus: Correcting the Market for Identity Data

Irene C L Ng and James W C Kingston
University of Warwick

ISSN: 2049-4297
Issue Number: 01/20

About WMG Service Systems Group

The Service Systems research group at WMG works in collaboration with industry researching into market design, economic engineering , new business and economic models and value-creating service ecosystems of people, product, service and technology.

The group conducts research that is capable of solving real problems in practice (ie. how and what do do), while also understanding theoretical abstractions from research (ie. why) so that the knowledge results in high- level publications necessary for its transfer across sector and industry. This approach ensures that the knowledge we create is relevant, impactful and grounded in research.

In particular, we pursue the knowledge of digital and data ecosystems design and innovation for value co-creation that is replicable, scalable and transferable so that we can address some of the most difficult challenges faced by businesses, markets and society. The service systems research group works with the HAT Community Foundation's innovation arm, HATLAB (https://hat-lab.org), the innovation space for the Hub-of-all-Things (HAT) ecosystem.

Homo Dataicus: Correcting the Market for Identity Data

Ng, Irene CL
Kingston, James WC

Service Systems Research Group
Warwick Manufacturing Group (WMG)
University of Warwick, Coventry, CV4 7AL, UK
Tel: +44 (0) 247652 4871
E-mail addresses:
Irene.ng@warwick.ac.uk
James.kingston@warwick.ac.uk

Authors:

Irene C L Ng is the Professor of Marketing and Service Systems at WMG, University of Warwick and Director of HATLAB.
James W C Kingston is the Deputy Director of HATLAB and Portfolio Manager, WMG Service Systems Research Group, University of Warwick

Acknowledgement:

The authors would like to acknowledge the following grants that have supported the research activities that has led to the development of this paper:

- EPSRC Dynamic, Real time, On-demand Personalisation for Scaling (DROPS) EP/R033838/1
- Alan Turing Institute Fellowship

The authors would like to acknowledge the invaluable advice and feedback given by Professor Glenn Parry, HATLAB Research Director and Professor of Digital Transformation at the University of Surrey and Nick McMahon

Citation:

If you wish to cite this paper, please use the following reference:

Ng, Irene C.L. and Kingston, James W.C. (2020) "*Homo Dataicus*: Correcting the Market for Identity Data," WMG Service Systems Research Group Working Paper Series, (01/20), ISSN 2049–4297

WMG Service Systems Research Group Working Paper Series
Issue number: 01/20
ISSN: 2049-4297
January 2020

Homo Dataicus: Correcting the Market for Identity Data

Abstract
Effective digital identity systems offer great economic and civic potential. However, unlocking this potential requires dealing with social, behavioural, and structural challenges to efficient market formation. We propose that a marketplace for identity data can be more efficiently formed with an infrastructure that provides a more adequate representation of individuals online. This paper therefore introduces the ontological concept of *Homo Dataicus*: individuals as data subjects transformed by HAT Microservers, with the axiomatic computational capabilities to transact with their own data at scale. Adoption of this paradigm would lower the social risks of identity orientation, enable privacy preserving transactions by default and mitigate the risks of power imbalances in digital identity systems and markets.

Keywords: Identity, personal data, Hub-of-All-Things, Digital ID

Main text
Digital identity usage is essential in modern society because it creates greater inclusivity and helps with the formation of groups. Formal identity helps reduce fraud and protect citizen's rights, and digital forms of formal identity ensure real time access to resources, ease of use of online services and greater efficiency gains for both individuals and firms. The result of digital identity usage is often speedier access, better coordination, greater outreach and wider communication.

The economic value of digital identity is substantial. The McKinsey report, *Digital ID: a key to inclusive growth*, issued in April 2019[1] highlighted the potential economic value of good digital IDs for individuals and organisations, as well as the potential risks and challenges. The report found that the economic value potential for digital ID in emerging economies is the equivalent of 6 percent of GDP by 2030 and for mature economies, 3 percent of GDP by 2030. Identity is a new form of currency, where, increasingly, only those with identities are allowed to transact for resources (Birch, 2014). The World Bank estimates 1.8bn people on the planet have no legal form of identity, resulting in their exclusion from social, economic and democratic activities. Without identity, individuals are more likely to be exploited, trafficked or confined to a life of servitude (Dahan & Gelb, 2015).

Identity is a set of data attributes that work together to verify the identity of a person, legal entity or device. Verification of that identity is the task of a trusted organisation that holds the "source of truth". For example, a passport can verify the identity of a person because the issuer of passport, a national government, has created the "truth" document, and is seen as a trusted body for the source of that truth - that the person is a citizen of that state.

Digital identity traditionally comprises of 3 components. First, identification. . This involves enrollment and identity proofing; data attributes such as name, date of birth etc are used to bind real world identities to digital identities.Second, authentication and lifecycle management; this enables control and management of these digital identities. Third, authorisation and assertion; this binds digital identities to entitlements, and requires presentation of credentials before transactions can take place. (CITATION).

[1] *Digital identification: A key to inclusive growth*. April 2019, McKinsey Global Institute

Digital identity systems are central to identity and access management (IAM). There is much interest in IAMs in both research and practice. IAM is a framework of governance and technologies to ensure that only verified people are granted access to digital resources[2]. Identity management systems not only identify, but also authenticate and authorise the usage of such digital resources (Benantar, 2005). Such systems have proliferated in recent years. On the demand side, identity verification systems are called to tackle the challenge of accessing increasingly diverse digital resources from crucial services such as banking, to the mundane day to day of signing into apps and websites. Identity verification is now an essential component of transactions by online companies from eCommerce and financial institutions to gaming and dating. On the supply side, technologies from cloud computing to blockchains have resulted in the creation of more rigorous and complex systems to counter potential security threats and mitigate the risks associated with fraud[3].

While there is a need to verify identities to enable access to resources, the mere act of verification bestows upon both the verifier and the user of that verification the knowledge of those individuals, in terms of when they access the resources, what they access and the way they use it. Such knowledge gives greater power to verifiying authorities and systems using the verification; applied by big data companies, it has led strength to warnings about the power of 'surveillance capitalism' (Zuboff, 2019). This has brought the security focus of identity management into work on scrutiny, privacy and rights within the legal domain (e.g, Pounder, 2008). To mitigate some of the risks, a three part test (purpose, necessity and balance)[4] is often used to scrutinise the legitimate interest of identity requests (Ferretti, 2014). However, variability in practices underscores the challenges to this test. Checking that such requests are legitimate is a challenge to implement and enforce especially on a free Internet where anyone from anywhere can create an app or website, and where 'legitimate interest' has been criticised as too broad a basis for processing (Kamara & DeHert, 2018). The temptation to creep and to acquire more data and power from personal data is hard to resist, just as it is hard to police its usage and misuse (cf. Greenleaf, 2008).

This paper argues that the first component of digital identity, i.e. ***identification***, face systemic challenges that are social and behavioral at the micro level and structural at the meso level. We propose that the endowment of individuals with a personal data server, forming *Homo Dataicus* - the *Data Man* - would enable individuals to obtain a set of axiomatic capabilities to obscure the binding between real world and digital identities that can mitigate social risks, while enabling individuals to transact directly in the market, thus still achieving the two other components of digital identity, that of authentication/lifecycle management as well as authorisation/assertion. We propose that the ontological concept of *Homo Dataicus* could correct the distortions currently faced by the digital identity market.

Literature

Social and Structural challenges of digital identity usage

Identity is a complex phenomenon in the social science and humanities. Philosophers believe identity to be a relationship between mind and body[5]. In sociology, identity is socially constructed, a process of negotiation with oneself and with society through actions and roles on the meaning within an identity (Swann et al., 2009). In psychology, identity relates to

---

[2] *Introduction to identity and access management*, January 2018, National Cyber Security Centre
[3] *Blockchain and retail banking: making the connection*, June 2019, McKinsey
[4] *What is the 'legitimate interests' basis?* Information Commissioner's Office
[5] *Personal Identity*, September 2019, Stanford Encyclopedia of Philosophy

self-image, a self referential mental model, tied closely to self-esteem and individuality (Stets & Burke, 2014) .

It might seem that the issues in the technological, security and legal domains that seek to verify that people are who they say they are for access to resources have no relationship to issues of social, psychological and philosophical identity. Unfortunately, the two sides are implicitly linked.

For clarity, we define social and psychological identity issues to be "internal referencing" issues and the side of security, technological and legal identity issues to be "external referencing" issues.

Identity Orientation. External referencing issues implicitly assumes identity is stable, and therefore encourage a utilitarian view, i.e. a person gives his or her identity information to gain access to resources and the provider of the resources receives identity information as a validation of that right. Unfortunately, resources in interactions are far more complex than can be explained from a utilitarian approach. Self and identity are central to the understanding of interactions (Blumer, 1969; Goffman, 1959; Mead, 1934). A person may adopt one of three identity orientations in any digital ID exchange: personal, relational, or collective depending on how the person defines him/herself (Brewer & Gardner, 1996). What this implies is that adoption of a digital ID for verification depends on the person's perception of who is asking, how is it asked, and the context in which such verifications occur. In all cases, the outcome of such verification results in a practice of *self-evaluation (that could reinforce self-esteem or self-worth)*. When an academic logs into the University of Warwick and the system asks to verify who she is through her phone (2FA), it reinforces her identity as an academic, which would reinforce (or make her doubt) her self worth of such a role; when the same academic logs into a garden tags app after 6 months inactivity, the app where she keep the photos of her plants and garden, she again confirms (or doubts) her commitment to that role to herself. Digital ID requests sit within a constantly changing *frame of reference*, a filter used to process the information requested that relates to a person's self-concept, implicitly influencing *motivations* that would direct the person's behaviour online. As pointed out in a South Yorkshire Credit Union report on challenges in digital identity, "identity management" means different things to different people (Simpson & Lindsey, 2014). All this implies that the criteria for different forms of verification are closely related to a person's identity orientation. This, in turn, impacts on when a person feels "safe" to transact within the digital identity space, beyond merely technological security. An individual who signs up to help with an LGBTQ cause, and is issued a digital ID for access to community activities may not wish this to be known when he returns to his hometown. A "stable", utilitarian identity socially prescribed by the state or any organisation may be viewed negatively by the person. Worse, when a single identity is used across all contexts, it may create dissonance within the person's perception of his or her roles, as that identity becomes stripped of its contexts. Identity orientation risk is therefore a key risk that impacts on the usage of digital identities.

Power Relations. In addition, the usage of Digital ID requires verification from the entity that controls the source of truth. This can be a formal source of truth, such as citizenship, or a less critical but nonetheless still relevant source of truth, such as having run a marathon in Cambridge, or having completed a diploma course. Such sources of truth are diverse in type, size and quality of data. It could comprise of one data point that can be easily made transactional such as marital status, but could also comprise a larger set of relational data such as "love folk music" or "travel a lot", verified by data from the music the person has been listening to, or the person's location data. As the Internet moves to collecting more data about the person across multiple apps and websites, more "truth-verifying" organisations are emerging that will hold varying degrees of power over the person. The potential for

subjugation on those who are asked to identify themselves for access to such truths creates power inequalities between the person and the verifier of the source of truth, as well as the requester for that truth. The ability to persuade individuals to share digital identification needs to be understood within a framework of power relations. Power relation risk is therefore a second key risk in the usage of digital identities.

Privacy, Security and Value tradeoffs. Privacy and security are perhaps the most debated and discussed aspects of digital ID. However, the context and content of data surrounding privacy and security ***and when they matter*** is often not discussed. The 'privacy paradox' (Barth & de Jong, 2017) suggest that individuals say they care about privacy but are actually not willing to act on it. This argument mirrors the argument in service literature where certain attributes are hygiene factors as proposed by Hertzberg (1993). Privacy and security can hence arguably be perceived as a hygiene factor for online services. The reality, of course, is that they are not. This has much to do with the complex interaction between the content and context of identity data with privacy and security categories. For example, a profile photo on Facebook has a very different social importance as compared to a photo of the same person being drunk. This means that privacy and security optimisations must be able to handle diverse privacy-security-value trade-offs and these tradeoffs must be co-created with the individual. One cannot assume that the cyber security threat model for securing Instagram photos (which are often public) is the same as the person's credit card, nor can an organisation presuppose how individuals would assess the tradeoffs when often they only wish to get quicker service. Better tools to manage trade offs are necessary (Petkovic & Jonker, 2007). This lack of choice (and tools) is a third risk in the adoption of digital identities.

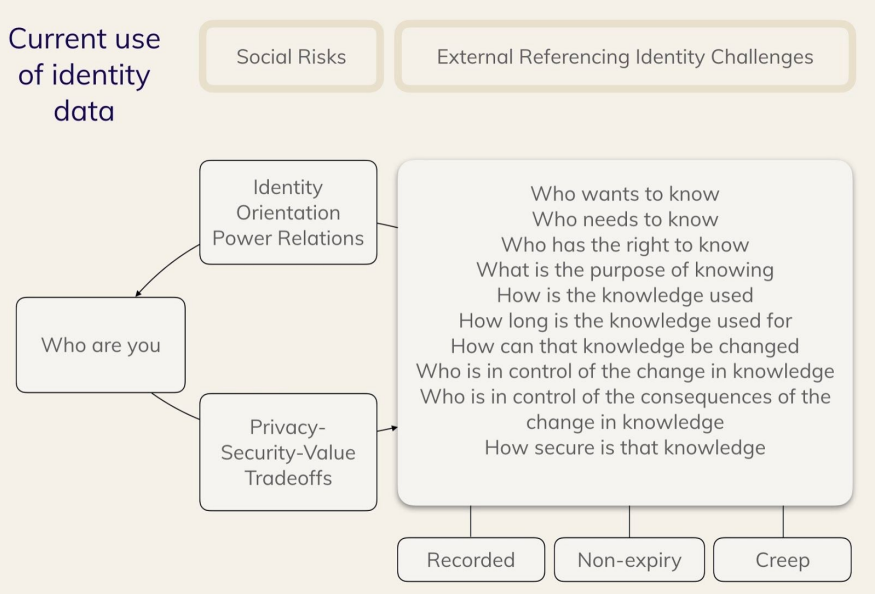

Figure 1: Social risks and Identity challenges

Structural (meso) issues. It is easy to evaluate micro level social and behavioral challenges and believe that meeting those challenges alone will be the key to individual adoption. Equally easy to fall into is to think of macro levers such as regulatory controls as a solution. Unfortunately, digital identity adoption is inconveniently a service ecosystem challenge that includes micro, meso and macro levels (Ng & Wakenshaw, 2018b). Meso level challenges are normally not well understood in systems work. The easiest way to think of meso challenges is to imagine a hotel. Human behaviours within the hotel are micro level challenges, the design of fixed structures of the hotel (rooms, lobby, kitchen, staircases, doors) are macro challenges and meso level is the design of soft furnishings; furniture, towels, toiletries, desks; these often drive templates of behaviours e.g. incentivising working outside the room by creating a co-working space in the lobby area. Micro, meso and macro levels of a system interact where changes at meso levels

can lead to micro or macro level changes. In digital identity, micro level social challenges must therefore also address meso and macro level challenges. For example, efforts within the technology and security domain often endeavour to strip context out of verification and authentication of a digital identity so that the solution can be much more scalable. Such a well intentioned market forming strategy at a meso level would paradoxically make identity orientation risks soar at the micro level leading to poor adoption. Also at the meso level, identity data for verification held by a central organisation is not easily scalable *from a social perspective*. At some point, even with a trusted body like a national government, power and trust issues will rear its head especially if everything is verified by a single universal identity controlled by one central body. Observers have raised fears that systems such as the 'Fatherland Card' of Venezuela can be used by governments as a tool to monitor citizen behaviour, and apply coercion via control of access to goods and services[6], and that they can be used to prevent access to goods and services by marginalised groups[7].

To mitigate the social risks outlined above, trust and familiarity are often used as market structures. Trust enables individuals to alleviate their cognitive load. As long as individuals trust an app or an organisation, they can think less, worry less and do more. It is therefore useful for coordinating the usage of digital identities (Yan & Holtmanns, 2008). However, since the risks of digital ID usage are not uniformly held across the Internet, and since individuals have selective retention, distortion and attention, trust is easily appropriated by organisations if they already hold a relationship with individuals and also easily used as a crutch by individuals if they don't want to have any cognitive load (Metzger & Flanagan, 2013). This creates market distortion in favour of incumbents which then leads to high switching costs.

The changing landscape of digital identity and the rise of economic identity

Real world personal identifying information such as name, address, passport numbers and email addresses are convenient attributes for Digital ID verification because these attributes are often stable, more digitally accessible, less time dependent and can be generated in multiple ways. However, it poses substantial social risks. Usage of a stable identity can be rejected by individuals due to fears of profiling and a rigid identity orientation. Verifiers of such identities also hold power over individuals, leading to fear of potential misuse. Meso risks are also compounded because the qualifications to transact are too rigid and narrow, and markets fail to form as they cannot tolerate the heterogeneity of trust, accuracy, and contexts.

For the longest time, a person's "real" identity was not a necessity for many systems, including the most critical ones. For example, with banking, it was not necessary for the bank to know who an individual was, but only that the individual was the same person that started the bank account. For many years, banking has operated on the notion that the customers actual name and identity was irrelevant, only that they had the right to open a bank account and that whenever they accessed it, it was the same person, or with the authority of the same person. However, regulation has since changed in many countries as governments crack down on money laundering and fraud, requiring banks to conduct more stringent "Know-your-customer" (KYC) checks and imposing "Common Reporting Standards" (CRS) on banks across the globe. To open a bank account now, individuals and organisations not only have to verify identity, they have to verify their source of funds and source of wealth.

---

[6] *How ZTE helps Venezuela create China-style social control*, 14 November 2018, Reuters
[7] *Rethinking Kenya's Digital ID System*, 19 December 2019, Open Society Justice Initiative

For individuals, giving identity data is giving the truth of who they are to third parties whom they think should know it because it gives them as individuals access to resources controlled by these third parties. It gives power to those who **verify that truth**, and those who are **users of that truth as gatekeepers to resource access**. Historically, the most common institutions that would verify that truth have been states. While the state has been able to verify identities for many years, the wider usage of digital identity to realise its value across the economy has had limited success. Across the world, even secure and privacy preserving digital identity initiatives have been a challenge to implement. In Nigeria, adoption of digital IDs stalled in 2017 due to the challenge of integrating uses across government departments[8]. In the UK, the government Verify programme[9] is proving to be limited in adoption. In Kenya, tribal politics and the initial absence of data protection laws led to heavy resistance[10] for the government's digital identity program. At the same time, issues of implementation and design can also be devastating, as demonstrated with India's massive digital identity program, Aadhaar. Authentication problems can prevent citizens from receiving food rations, leading to alleged cases of starvation[11]; poor linkage to bank accounts can mean wages and benefits are left unpaid, or directed to the wrong individual.[12]

Meanwhile, on the Internet, online digital identity usage have seen substantial growth, as more gateways for resource access are now online and as more resources become digitised. Daily lives are now lived with mobile communications, digital maps, online payments and banking, all resources that often require verification and authentication of identity.

Economic Identity: The proliferation of identity verification over the last decade has brought about a new form of identity which we term as "economic identity". Online economic identity is a derived form of an individual's identity on the Internet. It goes beyond knowing who the person is, to what the individual does, how he behaves and what he prefers in terms of the economic activity that the person can generate online. A similar concept is what organisations often refer to as "customer 360 view", a comprehensive profile of an individual for Customer Relationship Management purposes (Chen & Popovich, 2003). However, economic identity goes further than mere profiling. It is a set of attributes of a person that would result in the generation of economic activity, be it buying something online, reading news or sharing videos. For example, the attribute of "wine-lover" would result in the person read about wines, be interested in wine news and buy wine online. In a similar manner, "stressed working executive" attribute would be interested in calming apps or sleep therapy. Such attributes, rather than exhibiting a stable, utilitarian identity profile, economic identity recognises the dynamic aspect of individuals in terms of their constantly changing preferences, priorities and activities, and acknowledges that economic activities arise from such contextual circumstances. Economic identity is therefore more of a situational-dominant or contextual-dominant archetype rather than a psychological or sociological profile (cc. Ng, 2017). The Internet today enable websites and applications to amass petabytes of personal data recording what individuals do from browsing, searching, buying and posting on social media. Powerful analytics and algorithms can now be used to string together the personal data arising from disparate activities of individuals across their multiple devices. The economic identity of a person can now be inferred in real time and marketed to through ads, news and other media.

The new "governments" that verify economic identities are now the tech giants and data brokers, managing online authentication and flows of information. Many of these organisations now have more insight into the lives of netizens, more than states have ever had on their

---

[8] *The state of identification systems in Africa - a synthesis of country assessments*, 1 April 2017, World Bank
[9] *Investigation into Verify*, 5 March 2019, National Audit Office, UK Cabinet Office
[10] *Kenya's plan to store it's citizens DNA is facing massive resistance* , 21 February 2019, Quartz Africa
[11] *Aadhaar Failures: A Tragedy of Errors*, 5 April 2019, Economic and Political Weekly
[12] *How a glitch in India's biometric welfare system can be lethal*, 16 October 2019, Guardian

citizens in the past. Even though these organisations may not be the source of truth on identity the way the state is, they have grown powerful as they hold much of the personal data, and are able to verify the economic identity of a person - that he is the owner of a unique email address, that he owns a unique mobile phone number, that he generates tremendous economic activity online through his preferences, activities and priorities. In many cases, the individual's actual identity is almost irrelevant, as long as the economic identity can be ascertained. Paradoxically, the individual can be completely exposed, while being completely private; not because you can't infer a person's identity from the data (studies have shown you almost certainly can), but because the real person may no longer be relevant, a phenomenon often referred to in IS research as ontological reversal (Baskerville, et. al., 2019).

Crucially, the personal data collected on behaviours can now become the new sources of truth for the economic identity, creating a reinforcing loop. Car data can verify driving behaviour, TV data can verify viewing habits. The personal data collected from these behaviors become verifiers of economic identities in the same way a passport is a verifier of a real citizenship identity.

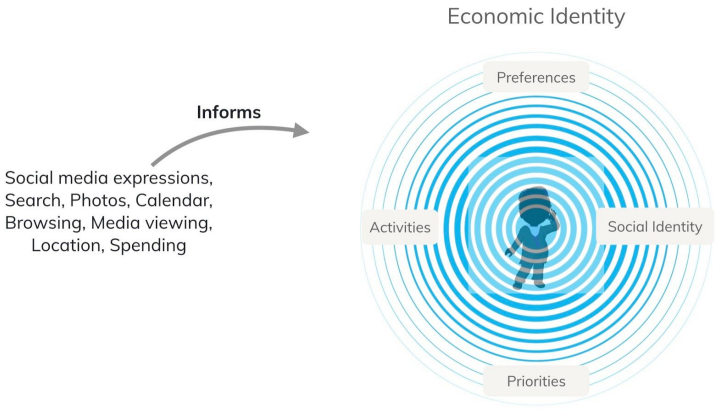

Figure 2: Digital behaviors inform Economic Identity

Personal data as sources of truths of online economic identity confer great power to the organisations that are able to verify its authenticity, as well as organisations that are able to use it. It is, in essence, an economic asset much the same as any currency.

If held by one organisation, such power would clearly result in a monopoly. Yet, if it is held by the state, it would be analogous to central planning, a form of data communism. If we seek **neither** to fall into a central planning data economy **nor** a monopoly, how can we enable markets to function more effectively for identity data?

We argue that the current online market for digital ID is distorted because individuals are currently not able to transact their identity data on the Internet directly on demand and at scale. Almost all verification and authentication of an individual's identity data online depends on third party providers acting as the "agent" for individuals. This poses a twofold problem. First, we are already seeing signs of market failure where identity data verification is monopolised by a few large players while also creating externalities in the form of privacy losses. Second, a third party agent representing individuals would result in typical principal-agent problems, where the agent (for example the verifier), who may not have a large stake in the outcome of the verification, could have interests that are aligned elsewhere e.g. to collect more personal data for other uses.

Correcting the market for identity data requires a market redesign, as defined by Kominers et. al. (2017),

"Market design seeks to translate economic theory and analysis into practical solutions to real-world problems. By redesigning both the rules that guide market transactions and the infrastructure that enables those transactions to take place, market designers can address a broad range of market failures."

We propose that a marketplace for identity data can be more efficiently formed with an infrastructure that provides a more adequate representation of individuals online.

Homo Dataicus

In much the same way that *Homo sociologicus* is the portrayal of a person's capacity or role as a social being and a member of a society and *Homo economicus* is the portrayal of the human person as rational, self-interested, and in the pursuit of optimal means-end decisions, we define *Homo Dataicus*, *the Data man*, as individuals with axiomatic computational capabilities to transact on the Internet with their own data on demand and at scale.

The transformation of individuals to *Homo Dataicus* can be achieved when individuals possess their own personal data server, for example, the HAT Microserver™ [13], an open sourced technology borne out of more than £3m, and 6 U.K. universities' projects (Ng, 2018a). The HAT Microserver, unlike a personal data store hosted by third parties, is a fundamental infrastructural capability that transforms individuals into *Homo Dataicus*, enabling individuals to transact with their own data with a *set of axiomatic computational capabilities: owning, storing, sharing and acquiring data (different sources of truths) into their servers; entering into contracts directly with the data; and executing on the transaction with their own servers in a way that is private, secure, legally binding, on demand and scalable.* Such a representation may seem similar to self sovereign identity. However the purpose of *Homo Dataicus* goes beyond self sovereignty. This is the Internet after all, where supremacy is measured by computational abilities. *Homo Dataicus* is an ontological construct of a data subject with scale and computational capabilities, bringing individuals to the same level of capability as that of corporations and their own servers on the Internet. It seeks to achieve adequate representation for individuals in an ontologically reversed world.

As long as *Homo Dataicus*, or *H-D* for short, is endowed with the axiomatic capabilities previously stated, they can acquire their own HAT Microservers from the market supplied by third party providers, including servers with greater security, privacy and other service enhancements e.g. to obtain private AI. For *H-D* to transact with their own digital ID, however, their servers must exhibit 4 necessary attributes of a good digital ID: **verification**, **uniqueness**, **consented sharing** and **privacy preserving** (McKinsey, 2019). We elaborate on *H-D* and their transacting capabilities below.

Verification for Purpose. In their report[14] McKinsey concluded that Digital ID must be verified and authenticated to a high degree of assurance and meets government and private-sector standards across all usage digitally. This makes an assumption that all verifications and authentications are for use cases that are of a critical nature. Online usage of identity data shows that the market requires assurances and standards across a spectrum of trust, accuracy and contexts. New and innovative applications often start with low level of assurances and non standard identity verifications. Holding a rigid position on assurances and standards threatens innovation and creativity. What is needed is a system that can incorporate very high standards and assurances while being tolerant of low level ones.

---

[13] https://www.hubofallthings.com

[14] *Digital identification: A key to inclusive growth*, April 2019, McKinsey Global Institute

Homo Dataicus can create more efficient markets by taking a ***verification-for-purpose*** approach with their data servers. Verification-for-purpose (VfP) looks at the purpose for verification and enable the *H-D* to give only the minimum necessary data to provide the verification necessary for assurance. By taking a VfP approach, the market can transact across a spectrum of purposes, and with a myriad of supply data, even without the need for personal identifying information.

Since *H-D* themselves can be their own verifiers through the data they acquire into their own servers, a market for various sources of truths can exist. For example, an education certificate can be acquired by a *H-D* and as long as the data is tamper proof in its journey into the data server, within the data server and out of the data server to whomever individuals wish to give the verification, the truth can be verified without the need to give away the data or even other sensitive data such as personal identifying information. Such a verification, tied to the purpose of verification, then provides a safe and secure verification process that can potentially mitigate social risks of identity, privacy and power relations.

At the meso level, a variety of transactions from the most mundane to the most trusted verifications can occur with a *H-D*. *H-D* can verify that they like folk music from their Spotify listening data for the purpose of taking a survey on music; or that they have attended college, or even verify the keys of their crypto currencies. The assets themselves within their data servers could be securitised, leading to more innovative digital markets. Heterogeneity of accuracy and privacy can exist and thus create better allocation of personal data across the market in a way that is privacy preserving. Data such as Fitbit steps can be used for the purpose of games; location data can be shared for 15 minutes to get lunch recommendations, and so one. Sharing economic identity and behavioral data that fit the purpose and create benefits for both individuals and organisation also reduces the probability of function creep, and if the identity data is queried and not downloaded, security risks could be minimised. The diagram below illustrates our point.

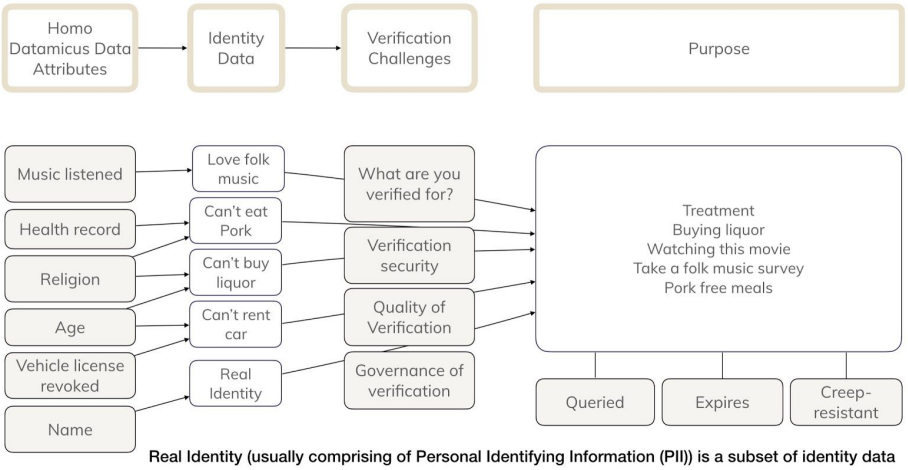

Figure 3: Verification of Purpose by *Homo Dataicus*

Uniqueness for purpose. A good digital ID ensures that "an individual has only one identity within a system, and every system identity corresponds to only one individual" (McKinsey, 2019). The danger of uniqueness, of course, is that it is a slippery slope towards a stable and socially prescribed identity by the power that verifies that uniqueness, potentially leading to creep and misuse.

Following the same principles of verification for purpose, *H-D* can authenticate themselves for uniqueness for purpose. Personal Data Servers consist of multiple folders, each of which can

be unique to different systems online and inaccessible to one another. A British citizenship folder must therefore hold a unique identity that cannot be similar to anyone else, and its use is for a purpose that is transacted by the H-D and the requester. Similarly, a Spotify folder must hold a unique identity in the Spotify system and could be used to participate in music surveys. The *Homo Dataicus* therefore benefit from being universally unique to themselves but have the freedom to only transact on uniqueness for a purpose.

Created and consented for sharing. All third party digital IDs would require individuals to be aware of how an ID is used and what data is captured. The capabilities of *Homo Dataicus* as a legal entity for transacting at scale would require the same, except that the market now opens for new services to create new sources of truth that can be acquired by the *H-D*, to be used as a Digital ID, for H-D to transact on. Such sources of truth can, in themselves, be third party ID providers who could have acquired new technologies such as blockchains and crypto protocols to create immutable sources of truths. *Homo Dataicus* is a new entity that can now potentially use these sources of truths to transact with other websites and applications or with one another. On the demand side, while the *H-D* may now be legally able to transact at scale, H-D can transact for other services in the market, perhaps with a robust governance system to support the *H-D* on what good transactions look like, to ensure transactions are transparent and in the interest of the *H-D* (see, for example, Dataswift services[15] with HAT Microservers).

Privacy preserving by sharing in context. *H-D* is private by default. However, similar to practices on sharing, transacting control often needs guidance, in terms of the governance of what good transactions look like. *H-D* can have different transacting options put forward by the market. For example, transacting on data based on duration and purpose reduces risk of exposure while enabling *H-D* to obtain good services. Transacting on identity data based on context is another market mechanism where only the data that is needed for a particular context is shared and verified. For example, verifying a person's address can be a query by a requester to issue an invoice and once the invoice is sent, the address can remain with *H-D* server and not stored by the application, to be accessed again when needed. Contextual sharing therefore disincentivises data hoarding and sharing of large and lumpy data as only what is necessary is shared; it reduces cognitive load on individuals as they only need to agree to the contract based on a simple explanation of context.

We argue that the axiomatic capabilities of *Homo Dataicus* can fundamentally correct the dysfunction of the market for identity data. The marketplace with *Homo Dataicus* can achieve *thickness* (Niederle et al., 2008), a condition in market design where a sufficient number of participants can come together to transact. This is because *Homo Dataicus* holds far more identity data that could include the acquisition of other sources of truths to transact on, for example behavioral data. The tolerance of heterogeneity in projecting unique personas for transacting with other systems while retaining universal uniqueness is also a key *thickness* factor for a marketplace of identity data to form on the demand side. By enabling transactions by *H-D* across a spectrum of data, accuracies, and contexts through multiple personas, potential market failure and *congestion* (for example when participants only wait for high quality verification), can be avoided. With privacy preserving, contextual transactions of more granular data, *Homo Dataicus* is able to mitigate social risks and enable *safety* in transactions (Vulkan et al., 2013), where it is more optimal to transact within the marketplace than outside.

Conclusion

[15] https://dataswift.io

This paper argues that the market for identity data is distorted and in need of correction. It introduces the ontological concept of *Homo Dataicus*, individuals as data subjects transformed by HAT Microservers, with the axiomatic computational capabilities to transact with their own data at scale. As *Homo Dataicus*, other capabilities can now be subscribed from the market, from security enhancements to contracting guidance and AI tools, including subscribing to the servers themselves. *Homo Dataicus* could correct the market for digital ID by lowering the social risks of identity orientation, enabling privacy preserving transactions by default and mitigating the risks of power imbalances. *Homo Dataicus* can form new markets in the digital economy (Ng, 2014), in particular for other data where more innovative data attributes of economic identities can be generated and shared for mutual gains (Ng, 2018c).

The ability for *Homo Dataicus* to transact with identity data directly and at scale would result in better coordination and allocation of data resources, in a similar way that real markets would create better allocation across consumers and organisations. Such coordination could evolve to be a reinforcing loop, generating matches that can cause spontaneous and ongoing coordination of scalable economic activities between *Homo Dataicus* and organisations on the Internet.

References

Birch, D. (2014). Identity is the new money, London Publishing Partnership.

Dahan, M. & Gelb, A. (2015). The role of identification in the post-2015 development agenda. World Bank Working Paper 2015.

Benantar, M. (2005). Access control systems: security, identity management and trust models. Springer Science & Business Media.

Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. Profile Books.

Pounder, C. N. M. (2008). Nine principles for assessing whether privacy is protected in a surveillance society. Identity in the information society, 1(1), 1-22.

Ferretti, F. (2014). Data protection and the legitimate interest of data controllers: much ado about nothing or the winter of rights?. Common Market Law Review, 51(3), 843-868.

Kamara, I and De Hert, P, (2018), Understanding the Balancing Act Behind the Legitimate Interest of the Controller Ground: A Pragmatic Approach. Brussels Privacy Hub, Vol. 4, No. 12, August 2018.

Greenleaf, G. (2008). Function creep–Defined and still dangerous in Australia's revised ID Card Bill. Computer Law & Security Review, 24(1), 56-65

Swann Jr., W.B et al., (2009) Identity negotiation at work, Research in Organizational Behavior (2009),

Stets, J & Burke, P. (2014), Self Esteem and Identities, Sociological Perspectives, Volume 57(4) 409–433

Blumer, H. (1969). Symbolic Interactionism: Perspective and Method. Prentice-Hall

Goffman, E. (1959), The Presentation of Self in Everyday Life. Pelican Books.

Mead, G. (1934), Mind Self and Society from the Standpoint of a Social Behaviorist. University of Chicago Press

Brewer, M. B., & Gardner, W. (1996). Who is this “we”? Levels of collective identity and self-regulations. Journal of Personality and Social Psychology, 71, 89-93.

Simpson, Gary., & Lindley, Emma. (2014) Investigating Challenges in Digital Identity: Digital Identity Inclusion and Uptake, The Open Identity Exchange

Barth, S., & de Jong M. (2017). The privacy paradox: Investigating discrepancies between expressed privacy concerns and actual online behaviour. Telematics and Informatics 34 (2017) 1038-1058

Hertzberg, F., et al. (1993) Motivation to Work, Routledge. 3rd Edition.

Petkovic, M., & Jonker, W. (Eds.). (2007). Security, Privacy, and Trust in Modern Data Management. Springer Science & Business Media.

Ng, Irene C.L. And Susan Wakenshaw, (2018b) “Service Ecosystems: A timely worldview for a connected, digital and data-driven economy” in Handbook of Service Dominant Logic, Robert Lusch and Stephen Vargo Eds, Sage

Yan, Z., & Holtmanns, S. (2008). Trust modeling and management: from social trust to digital trust. In Computer security, privacy and politics: current issues, challenges and solutions (pp. 290-323). IGI Global.

Metzger, M. J., & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. Journal of Pragmatics, 59, 210-220.

Chen, I. J., & Popovich, K. (2003). Understanding customer relationship management (CRM) People, process and technology. Business process management journal, 9(5), 672-688.

Ng, Irene C.L. and Susan Y.L Wakenshaw (2017), “Internet-of Things: Review and Research Directions”, International Journal of Research in Marketing, vol.1, iss. 34, no.1, pp3-21,

Baskerville, R. L., Myers, M. D., & Yoo, Y. (2019). Digital First: The Ontological Reversal and New Challenges for IS Research.

Kominers, Scott Duke, Alexander Teytelboym, Vincent P Crawford, (2017) An invitation to market design, Oxford Review of Economic Policy, Volume 33, Issue 4, Winter 2017, Pages 541–571

Ng, Irene C.L. (2018a) “HAT data ownership model: first party IPR for individuals”, in *Data ownership, rights and controls: Reaching a common understanding: Discussions at a British Academy, Royal Society and techUK seminar* , 3 October 2018,

Niederle M, Roth AE, Sonmez. T (2008) Matching and Market Design. Durlauf SN, Blume LE. The New Palgrave Dictionary of Economics. 2nd Edition.

Vulkan N, Roth AE, Neeman Z (Eds) (2013) The Handbook of Market Design. (Oxford University Press, Oxford)

Ng, Irene C.L. (2014) Creating New Markets in the Digital Economy: Value and Worth, Cambridge University Press, Cambridge, ISBN No. 9781107049352

Ng, Irene C. L. (2018c) The market for person-controlled personal data with the Hub-of-all-Things (HAT). Working Paper. Coventry: Warwick Manufacturing Group. WMG Service Systems Research Group Working Paper Series (01/18). (Unpublished)