THE UNIVERSITY OF WARWICK

WMG Service Systems Research Group
Working Paper Series

White Paper on Digital Dependency:

The Need for a Personal Data Exchange Infrastructure

Principal Author: Irene C L Ng

Contributing Authors:
Jonathan Holtby
Xiao Ma
Andrius Aucinas
Paul Tasker

ISSN: 2049-4297
Issue Number: 02/17

About WMG Service Systems Group

The Service Systems research group at WMG works in collaboration with large organisations such as GlaxoSmithKline, Rolls-Royce, BAE Systems, IBM, Ministry of Defence as well as with SMEs researching into value constellations, new business models and value-creating service systems 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 service systems 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.

Research Streams

The WMG Service Systems research group conducts research that is capable of solving real problems in practice, and also to create theoretical abstractions from or research that is relevant and applicable across sector and industry, so that the impact of our research is substantial.

The group currently conducts research under six broad themes:

- Contextualisation
- Dematerialisation
- Service Design
- Value and Business Models
- Visualisation
- Viable Service Systems and Transformation

WMG Service Systems Research Group Working Paper Series
Issue number: 02/17
ISSN: 2049-4297
March 2017

Service Systems Group
Warwick Manufacturing Group,
University of Warwick, Coventry CV4 7AL, UK.
E-mail: sswmg@warwick.ac.uk

Acknowledgement: The authors gratefully acknowledge the funding contribution of Research Council (UK) Digital Economy to the HAT project (http://hubofallthings.com) grant reference EP/K039911/1 which has contributed substantially to the research conducted in this paper.

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

Ng ICL, Holtby J, Ma X, Aucinas A, Tasker P (2017) White Paper on Digital Dependency: The need for a Personal Data Exchange infrastructure. HAT Data Exchange Ltd. WMG Service Systems Research Group Working Paper Series, paper number 02/17, ISSN 2049-4297. Available at: http://www.hatdex.org/digital-dependency/

HATDeX

White Paper on Digital Dependency:

The need for a Personal Data Exchange infrastructure

Digital Dependency

In an increasingly connected society, we are coming to rely on our Internet services just to function day-to-day. We book train tickets, navigate to meetings and social functions, maintain the climate and security of our homes, answer our doorbells, communicate with our friends and colleagues, and manage our finances, all online. Digital natives are becoming increasingly dependent upon connectivity – and this trend is clearly set to increase.

Benefits of a digitally connected economy

The benefits to society of a digitally connected economy are obvious. Sales data and contextual data over our purchasing and lifestyle habits provide consumers with value that minimises economic waste and maximises consumer value. The power and the potential of technology to make human beings more productive allows us to focus our resources on the activities and relationships that matter, ultimately improving health, wellbeing, and happiness. Societally, the exploitation of data can improve our quality of life and enhance decision- and policy-making at an organisational level. This allows for the intelligent allocation of activities, services and resources, and it helps us to increase the differentiation between signal and noise when identifying human preferences and desire. A digitally connected society is critical for the development of Smart Cities, the useful deployment of the Internet of Things (IoT), the growth of predictive analytics, and the future of the digital economy itself.

Of course, where connectivity brings convenience it brings challenges – especially in the area of personal data. Here, the challenge is about more than just a privacy concern. Every user signing on to a new Internet service creates for themselves a unique user account every time, surrendering some form of their personal data in the process. As a result, a few hundred, if not a few thousand equivalent accounts are created by each of us in our lifetimes. And the number of times we will do this each week, each month, or in our lifetime is increasing rapidly as the IoT expands.

This paper discusses the challenges of a connected digital economy from a personal data standpoint, and proposes a best practice of individually-owned, private data accounts in the form of ‘microserver’ containers, as a solution to the challenges we will face in the coming years and decades.

The societal and economic challenges of personal data accounts for Internet services

Having every individual create and own a few hundred Internet accounts over the course of their lifetime is a nightmare scenario. It's torture for the individual and a cyber security risk to society. At the same time, it is impossible to enforce personal data regulation. The entry cost for app makers is low enough that app builders can set up, obtain data for, and then tear down applications simply for the user information inside, and there is no market or regulatory vehicle to prevent them. Any regulation would also create an adverse selection problem, where legitimate applications bear the expense of their less cooperative competitors.

It is likely that in the long term, the market will emerge a personal data consolidator to address the consumer pain of having multiple online accounts, and the likely candidates for that are the incumbent players. Facebook, Amazon, Google, and Apple already have billions of individual consumer accounts, and there is a proliferation of sites that currently allow their users to login with those identities instead of creating a new one.

This 'market solution' and the alternative of having things stay the way they are in a 'personal data fragmentation' solution are both far from ideal. They are suboptimal for the economy, and undesirable for our society. As we will elaborate below, we run the risk of creating illegal data markets, deterring the creation of apps and services, allowing monopolisation in the digital economy, marginalising of entire populations of the market, crippling our efforts to deal with cyber security, and subjecting the digital self to costly dispersal and ageing.

Suboptimality to the digital economy

Three suboptimal outcomes are possible for the digital economy.

First, if we allow personal data fragmentation to continue as it is, illegal markets will form for personal data exchanges. Non-anonymised data is valuable, and the process of anonymisation to comply with data privacy regulation is costly. The opacity of selling data will make enforcement difficult, which will inevitably lead to a black market that would thrive from the flow of the most valuable data.

An additional malady of fragmentation is that even with the granting of individual consent (perhaps via API-to-API exchange) services with an appetite for more personal data need to invest heavily to build the multiple codesets that allow them access to corporate-owned information. This presents a deterrent to new services, and incentivises services to build on the platforms of incumbents,

where one set of APIs can bind multiple disparate datasets. Fragmentation in this way may inevitably result in centralisation.

The third suboptimality to the digital economy arises from that centralisation, as greater monopolistic power is bestowed upon the firm that acts as the personal data custodian. Internet customers of every online service could ultimately become locked to that single provider of personal identity, which could force the entire ecosystem to do business with a single monolith, or risk losing access to the information that makes their Internet services run.

Undesirability to society

Both personal data fragmentation and centralisation outcomes are also undesirable to society, for several reasons.

First, corporate data ownership may result in the loss of privacy and confidentiality that, to some users, could ultimately be so great that either a large segment of the market withdraws from the Internet entirely – and in so doing becomes marginalised – or an adverse market effect in which only the rich can afford to secure and encrypt their data could come to govern the personal data privacy landscape. Even before the market fractures, the integrity and the quality of the personal data in this economy will begin to suffer, as advocates subvert the system with fake accounts that are easy to generate but difficult to cull.

Second, the cyber security risk of having so many points of vulnerability on the Internet would become too costly for the state to manage and police. Securing one honeypot is hard against the state-of-the-art attacker. Securing an entire market of them is next to impossible.

Third, in a fragmented market, the digital assets of individuals could come to be dispersed across the Internet. As our digital society ages, a lack of consolidation and personal ownership rights will make it difficult, if not impossible to pass on or manage our Internet selves through our families. Our digital identities, trapped in their multiple accounts, may tragically come to disintegrate more quickly than do our physical ones, as we lose the skill to process and use the Internet service accounts that gave us access to our own information in the first place.

Ironically, the more that digital users are able to reap the benefits of a connected society, the more difficult it will be for them to consolidate their own data, and the faster it will be that the digital self fades into oblivion as a result.

The role of a private data microserver container

We propose that the above challenges can be addressed through the proliferation of private, standalone databases for personal data that can be owned, solely controlled, and used by individuals.

These privately owned databases have the potential to make individuals data controllers and processors, in the same way PCs liberated individuals from mainframes in the eighties. This can be accomplished through the use of containers, which encase various discrete components of application logic and require only minimal resources to do their job. Unlike virtual machines, containers do not need an operating system. Instead, they call for operating system resources via an Application Programming Interface (API). Containerising databases in this way can isolate them at the (micro) server level. The content within them can be encrypted and backed up regularly, and traditional direct database access can be replaced by server-level API calls. This isolation creates an added extra layer of security, localising the impact of any breach and mitigating the risk of *sysadmin*-granted unauthorised access. Through containerisation, modular and micro cloud services are beginning to supplant large cloud architectures due to their portability and scale.

We propose the use of such 'microserver' containers as private data accounts as a universal best practice. Whilst containers have been in the past used to create agile services, there is no reason why they cannot today be used to manage a database containing a user's personal data.

Advantages of microserver containers as private data accounts

The architecture of a containerised private database with microservices has a few advantages. First, a user's personal data sitting within his or her own dedicated database means that data at rest within that database can be legally owned. This is in contrast to case laws that currently imply data cannot be legally 'owned' due to the fact that they often have no fixed boundaries, and are mixed between the public, the corporation, and the person. Ambiguity like this creates massive economic challenges, as the uncertainty of data's boundaries, entities, and rights increase market transaction costs and limit the creation of new services. Left unchecked, this friction could result in full market failure.

But while data may not be owned, data*bases* can be, and there is a corpus of case law to support this. Users who are the legal owners of their personal databases can therefore be afforded all of the property rights of the database,

reducing the above ambiguity, cost, and friction. The database can be treated as property – i.e. a good – that confers upon the individual its bundle of rights: the right to use the good; the right to earn income from the good; the right to transfer the good to others; and the right to enforce property rights over the good. More importantly, digital assets within the database can be managed and used, and be a part of the individual's estate, much as would the physical assets within a home.

Individuals as data controllers

Containing one individual's data within an entire database allows the individual themselves to be a data controller, and to some extent a data processor as well, operationalising the bundle of rights to which they are due. Individuals can exchange the personal data within their database for their own benefit, deriving income from it or transferring it for fun or service if they wish. The containers themselves help individuals do this, using standard APIs, whilst the individuals themselves stay in full control.

Transfers and exchanges set up by the data controller, with the accompanying reduction in ambiguity, means that bargaining solutions (trade) can be achieved. This allows a new primary market for personal data that is acted upon by individuals, rather than corporations.

Integration with Internet services

Two further benefits to containerised private databases follow. First, a database that is wholly owned by the individual can become an effective on-demand data supplier to firms building services that require personal information. Personal data form entry, personalised quotations, assessments, online identity verification, and user account creation all carry risks and costs for the corporation. An alternative that allows for the sharing of personal information, by users, from their own private data containers can save both businesses and individuals time, effort, risk, expense, and liability.

Widely adopted, the number of personal user accounts that sit within apps and services worldwide could be systemically reduced, eliminating pervasive cyber security vulnerability and greatly reducing the incentives for cyber attack. A penetration into one secure database container yields the perpetrator of that attack exactly one database, where in the current system a similar risk would yield up to billions of records of personal data instead.

Digital Personhood

Finally, a containerised private database that belongs to the individual can be consolidated across the vertical silos of industry. Such 'horizontalisation' of personal data can generate a far better usage of data-in-context for Internet users, helping to introduce new types of analytical tools. If the database schema is created as an open standard, a market for such services could emerge where users can buy intelligence and curation services, emerging new schemas for digital personhood, better curation of the digital person for wellbeing and health, and even the digital personification of the human, much like its genome equivalent.

Integration of a container as a private data account with Internet services would require data exchange standards and protocols to be established. By building these protocols on Internet open standards and ensuring they are transparent to all in the ecosystem, every stakeholder in that ecosystem will stand to benefit.

The need for a personal data exchange infrastructure

'*How do you sneak a table cloth onto a table full of food, crockery, and cutlery while people are still eating*?'

The proliferation of private data containers across the Internet for all digital users requires more than just creating the technology. The Internet economy is a complex network of technologies, markets, economic models, behaviours, and ecosystem design. To succeed, stages of innovation diffusion and the articulation of a clear value proposition to entrenched players must be considered. Every aspect of adoption, from marketing and branding to technology and integration design, must be considered at the ecosystem level.

The challenge of detaching a user from a service, and then re-integrating them into the same service as a new modular private data container, can be met by building the data exchange infrastructure on standard protocols. It is important to uphold familiar, tried-and-tested rules, while demonstrating that such an alternative can work seamlessly, and operate without compromising the performance of the service.

We therefore propose an exchange infrastructure that can facilitate:

- semi-private data exchanges between embedded individual private containers and their service providers;

- fully private data exchanges between individual private containers, the services they use on their own containers and the way they can organise and use the containers privately;
- public data exchanges in which data leaves their containers for the benefit of consenting parties.

Such an infrastructure incentivises a market in three places. First, as trust brokers for the provisioning of private data containers to other third-party firms and users. Second, as a facilitator of market exchange, an infrastructure of this sort could facilitate inter-organisational and organisational-individual exchange. And third, as a provider of tools, apps, and services for private data use where there is no access to user data (e.g. through the creation of bots and AI analysts), and finally, as enabling services for peer-to-peer data exchanges (e.g. food or car shares).

Where the need to integrate personal data containers into the established ecosystem is of initial concern, it is in fact the actors in the digital economy of the future that stand to gain the most from its proliferation. For these new entrants, the value proposition of a personal data ecosystem with a steady supply of diverse personal data provides a competitive advantage that can be used to challenge incumbents.

Network effects and B2B channels

To ensure the ubiquity of private data containers, and to maintain an unencumbered system, multiple channels can be empowered to provision private data accounts. Firms can be free to adopt and provision accounts themselves, or use third-party service providers who emerge to perform this task. The same also goes for any individuals who come directly into the ecosystem.

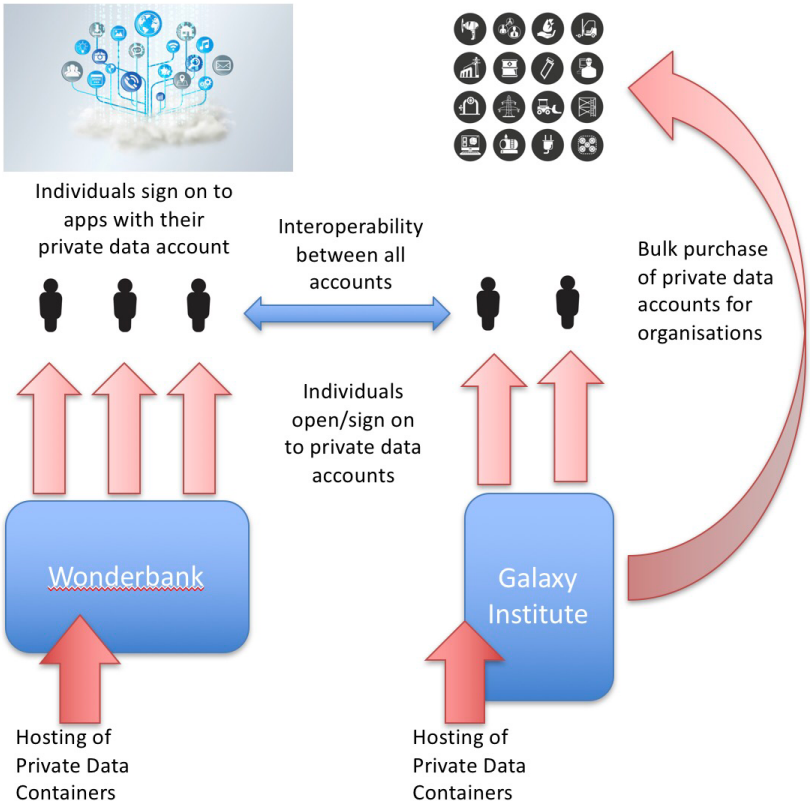

Figure 1: Trust brokering with private data accounts

In this respect, the provisioning of private data containers is analogous to the provisioning of an email account or a bank account, creating an implicit market for trust brokering. Large organisations could step forward to provision private data container accounts for the consumer market in much the same way they do bank accounts, or email accounts (when they come from a trusted provider). A business-to-business (B2B) channel could also come to exist, if organisations who do not want to risk holding onto personal data commission private data containers in bulk for their users. Organisations' HR departments may choose to provision such accounts for their employees. By having large organisations provision private data accounts for individuals, and establishing identity over those accounts with their own branding, they could even use them to cultivate competitive differentiation against their rivals.

Small companies in the IoT sector, for example, could integrate with private data accounts to mitigate the risks inherent to personal data security and this would spur greater growth and innovation in the sector with a reduced cost of secure personal data containment.

In addition, as each ecosystem member increases the supply of personal data into the private account ecosystem, every *other* ecosystem member gains the opportunity to request that same data back from mutual customers to improve the value proposition for their own services. Were a single market-leading social firm, financial services firm, cyber security firm, IoT firm, and health services firm to be simultaneously offering a single customer set services within the ecosystem, a near-comprehensive dataset over the individual consumer might be obtained, all with the customer's permission. These firms would find a significant competitive advantage within this environment.

Personal data regulation that is coming into force around the world is beginning to compel firms to give individuals the right to access their own data, which may expedite this process and cause leading firms to overcome their initial reluctance to embrace alternative data models. Others that are still reluctant to outsource personal data storage to individual private containers in their entirety, can still, as before, opt to hold onto some or all of their customers' personal data as data controllers in their own right. The private data account simply allows individuals themselves to do the same. The marginal costs of data duplication are negligible – though the additional security risk would still affect the firm.

Table 1: User, industry, and society benefits, presupposing the management and coordination of a private data container is no more onerous than managing and coordinating apps on a smartphone.

Benefits to users	Benefits to organisations	Benefits to society
Greater control over personal data usage	Ability to request access to more personal data for new services at reduced development costs.	Lower cyber-security risks
More private	Lower risks and costs of personal data containment (assuming no duplication on the firm-side)	Better representation of individuals in the digital economy
Ability to use and re-use personal data for themselves	Champions digital empowerment and control	Enables peer to peer services without third party involvement

Ability to buy services to, organise and manage digital selves and personal effectiveness	Creates trust with customers	Enables individuals to engage with public services more seamlessly through data sharing
Ability make more informed decision based on historical and on-demand personal data	Create direct customer relationships rather than be dependent on third party supplier of personal data	Better operationalisation of a consent based digital economy of personal data
Able to buy intelligence services for computation and recall	Better quality of data as individuals are stakeholder of data quality	Creates a disincentive for secondary (and/or illegal) personal data markets since there is a primary market for personal data
Ability to share data for insights, recommendations	Increase supply of personal data resource without high costs	Efficient way for government and organizations to consult citizens e.g. polls, opinions, surveys
Ability to share data for discounts and personalised products/services	Ability to access personal data services (e.g. anonymisation, blockchains) through the ecosystem without having to develop themselves	Benefit from scale effects when introducing new data services e.g. ledgering / blockchains

Growing and regulating the ecosystem

The enabling technology of this form of platform ecosystem must be open-sourced, so that new potential entrants to the economy don't suffer from costs of being locked in, deterring their participation in the first place. Yet, there is still a role for the market – as the technology develops, first movers can charge rents on services to be commercially profitable to just the extent that it deters new entrants from building their own. When the technology is scaled, micro transaction fees for such services could create enough profitability to power the whole ecosystem without taxing any one party. At scale, network effects could

render a large portion of data exchange to be free, issued as a public good, while commercial exchanges benefit from scale economies and network effects.

Ultimately, this would create a self-regulating, self-reinforcing ecosystem that avoids price-gouging, while providing enough rents to fuel the ecosystem, attracting a variety of funding over the course of the ecosystem evolution that can include private and community investments, private equity, venture capital and public offerings.

Such a system would still require some oversight. We propose that a member-owned regulatory body collectively decide on interoperability standards, certification of new membership, and compliance. Within such a framework, a rating system can be set to define baseline adherence, while also allowing different degrees of privacy, confidentiality, security, and trust (PCST) levels to exist. This would give a great degree of freedom to members needing to make the difficult trade-offs between cost and compliance, and spurring innovation and growth within a peer-defined regulatory environment.

Summary

Our increasingly digitally dependent society is yielding tremendous value – in the data-driven empowerment of commerce, the growth of IoT, and an exploding digital economy. But its rapid development is also showing cracks as personal data accounts for Internet services fragment across the Internet or consolidate under market-dominant monoliths. But the risks are great as we are threatened by the possible creation of shadow economies, market division, systemic vulnerability, and the erosion of the digital self.

These challenges could be addressed by the proliferation of private data accounts in the form of containerised personal databases that can be owned, solely controlled, and used by individuals, synchronised with a vast number of Internet services. Deployment of this technology, with an exchange infrastructure that can facilitate the private, semi-private, and public transaction of personal data would generate more effective and efficient societal outcomes and shift the digital economy towards a new paradigm.

Principal Author:

Irene C L Ng, Chairman and Chief Economist, HAT Data Exchange Ltd

Contributing Authors:

Jonathan Holtby, community manager, HATDeX
Xiao Ma, Chief Technology Officer, HATDeX
Andrius Aucinas, Head if Engineering, HATDeX
Paul Tasker, Chief Executive, HATDeX

About HATDeX:

HATDeX is HAT Data Exchange, the operator of the Hub-of-All-Things (HAT) ecosystem, a personal data exchange ecosystem of organisations and individual private data containers initiated through a £1.2m Research Councils project with 6 universities. The HAT enabling technology includes various data infrastructure exchange services designed and built to ease the integration of private data containers with Internet applications. Internet services can now use their own consumer brands to provision private data containers for their customers, powered by the HAT as the enabling technology. The HAT, its schema, logic and database is fully open source, as is the HAT dashboard, Rumpel (on the web) and Rumpel Lite iOS app. HATDeX is regulated by the HAT Community Foundation, a members organisation tasked to regulate and certify ecosystem players, as well as innovate and grow the ecosystem.

Partners:

For more information on HATs, please visit https://hubofallthings.com