mydex

Unlocking the value of our data

The individual as the point of integration

Contents

Executive summary

The value of our personal data is much higher than we have realised so far. It will form the backbone and the lifeblood of our use of the internet in years to come. But the way we currently deal with it is not working. The technology is here today, but the structure and the processes are not in place to reflect the opportunities in front of us. To realise the full value of our data, we can start with a few realistic, achievable changes that will begin to dissolve the gap between digital and real, and bring us into a new era of interaction between organisations and individuals.

The real value of personal data

Data is no longer just about profit, knowledge, advantage, 'number crunching' or predictions. Data is becoming the backbone of modern life. In almost every institution, transaction, interaction, exchange—data is involved at some point. Governments want to move to 'digital by default', the world's biggest companies base their success on the data they can collect and process.

The fastest growing startups and the most innovative ideas of the last 10 years have found their success in bringing a layer of information—a layer of data—to existing markets: AirBnB, Uber, eBay, Amazon's online marketplace. People of all ages and backgrounds have found their voice on the plethora of social media platforms—been discovered, become famous, become role models for a generation of young people.

Online communications tools like Slack, WhatsApp, Telegram, WeChat, Snapchat, to name but a few, have changed the way we do things at work, at home, and in between. An entirely new kind of workforce has emerged thanks to the ability to send and receive data in different ways—remote, 'always-on', efficient, flexible; an idealistic but not entirely inaccurate picture of the possibilities available to employers and employees today.

For many people, an online service would be the first port of call for a vast number of activities they undertake: working, chatting, planning, banking, shopping, traveling, sharing, learning...And yet some things are still remarkably difficult on the internet. Proving things about yourself is notoriously time consuming and frustrating. Security processes are convoluted—far too 'present' for the majority of people who are just trying to get something done. It should come as no surprise that the average user is often willing to 'sacrifice security' if it enables them to 'just do something'. To them, this is not what they're doing. They are not thinking in terms of sacrifices, personal data and potential repercussions. As you reach the sign up screen for the new online tool that your friend recommended to you because it's so useful, or the new app you download because of its tailored-to-you recommendations for events in the city you're visiting, why would you bother to enter your email address and think up a password when you can just log in with one of your pre-existing accounts on Google, Facebook or Twitter?

The problem has two principal layers:

1. Trust
2. Benefit

Trust in the eye of the beholder

The experience described above—using your social networking credentials to authenticate yourself on other services—has become known as 'social login'. As useful as this is, it serves a limited purpose. It speeds up the creation of an 'identity'—a digital credential—that you can use to access a service by reusing one you already have. This is, indeed, very convenient. With a minimal amount of effort, the new tools of the internet are at our fingertips.

But what about when we need to prove something about our 'real-life' selves? Using your Facebook credentials to access something online proves little more than the fact that you have a Facebook account.

Bridging the trust gap

As our 'real' and digital lives become ever tighter intertwined, it is not only brand new ideas that make it onto the internet. Processes we are used to doing in person are being replicated and, ideally, improved on the internet. 'Disruption' is the favourite buzzword of the online service industry, and not without reason. Monoliths and household names of almost any market are being challenged by the possibilities of information-technology and internet-based components. They are being challenged to change the way they operate—in the back-office, with their customers, partners, investors, and the wider world.

As the digital and the physical collide, and traditional brick-and-mortar is augmented or even replaced by online offerings (think banks with no branches, online stores, publishing, doctors, government departments, the list is endless), the essential thread, the vital, shared component required to make the transition, is *trust*. Can't prove to your bank that you are who you say you are? No online banking. Can't provide proof of address to your local parking authority? No parking permit. No way of proving your identity, income and energy consumption? No energy benefits. The more complex or sensitive the task, the more evidence is required about an individual to be able to provide them with the service that they require. The amount of effort is exponential, both for individuals and for the organisations providing them with a service.

Identity crisis

As we've mentioned above, traditional models for accessing an online service involve setting up an account and remembering those credentials for future reference. Social login simplified this by allowing us to use one set of credentials to access multiple services, but it didn't solve

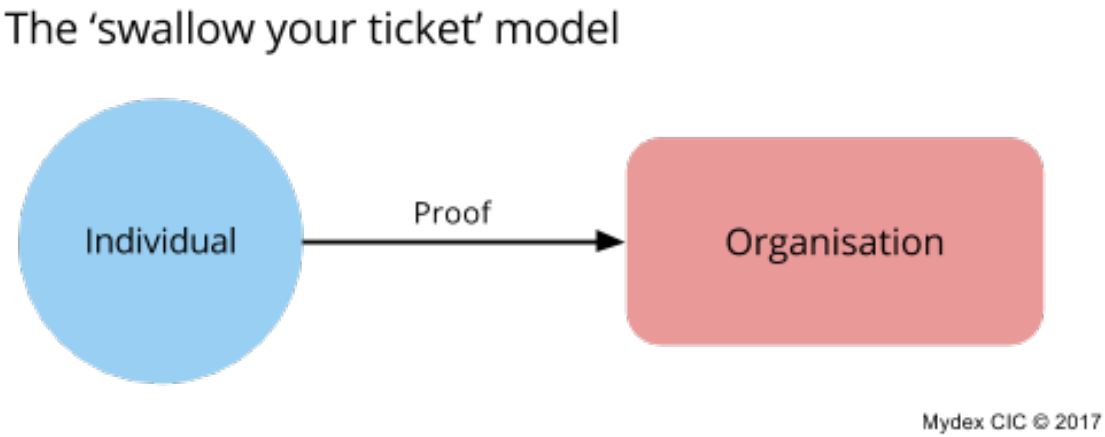
the issue of trust: how can I provide the necessary proof about myself and my life to be eligible to access this service? Many organisations use third-party verification services to, for example, scan and verify an official document as a one-time proof of identity. This might work in the context of your access to one service, but viewed in the context of your life as a whole, this is unsustainable for both you and the organisation serving you.

When you complete a traditional sign-up, you 'create an identity' within the environment of the service you want to access—let's call this Service A. There might be something else you need to get done online, related to a different part of your life; you might do this with Service B. On Service A, you have carved out your existence with the creation of your username and password, your 'access credentials'. You've done the same with Service B. To you, as an individual, Service A and Service B are both present and relevant to you. You are one person, and Service A and Service B are things that you use to get things done.

Conversely, to Service A and Service B, you are two people. Your account and credentials for Service A are not related to those for Service B. More critically, the evidence you provided to Service A to prove who you are did just that, and nothing more. You had to provide evidence again for Service B, they verify it via a third party or other means, and so it goes on. You now exist, verified, in two separate, unrelated online environments. Multiply this by the degree necessary to reflect the variety of your life, and you might suddenly find that, despite there only being one 'real' you, 'you' are represented hundreds of times across hundreds of services.

Not a one-way street

If we look at this state of affairs in terms of data flow, it helps to visualise the root of the problem. Each time you access a service, you provide the evidence, and this is verified by the organisation. What you get in return is access to the service. Crucially, however, all that effort in creating an account and proving who you are ultimately goes to waste.

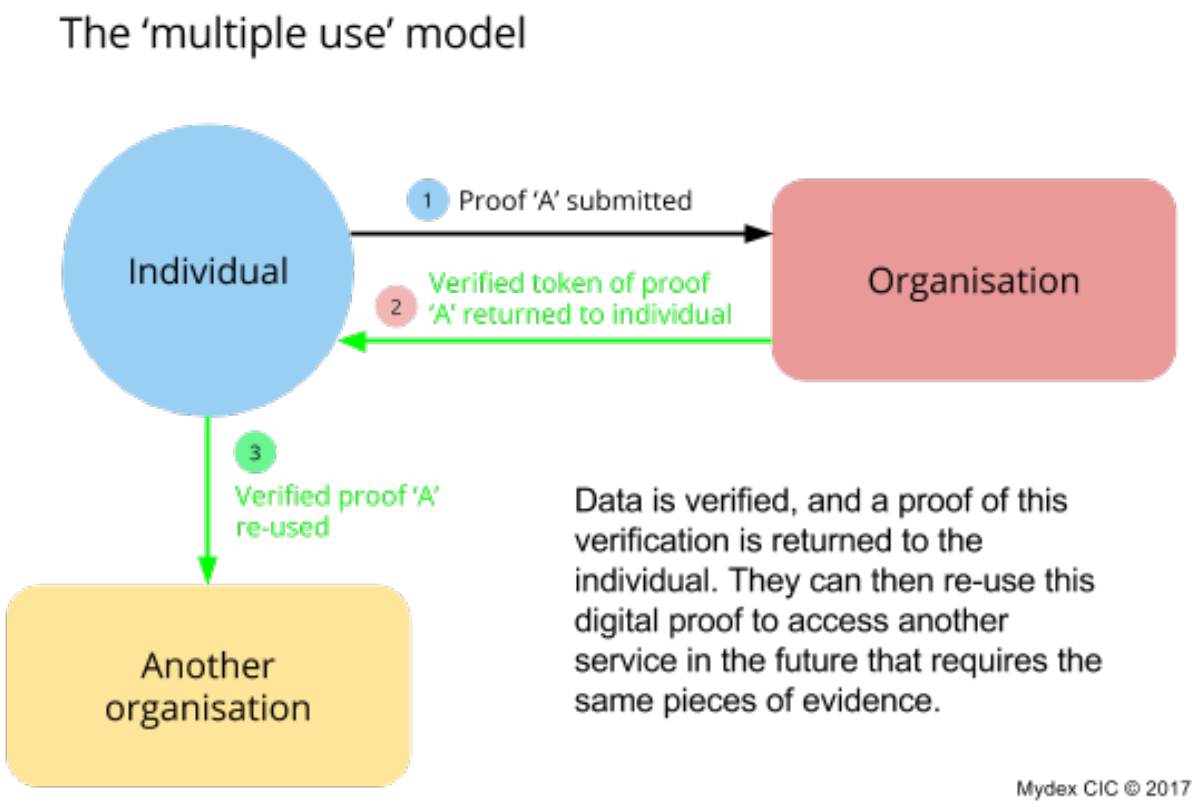


Think of ticket barriers for trains in the UK: the gate opens when you insert your ticket (your 'proof'). If you've bought a single-journey ticket, when you reach your destination, the gate will open, but 'swallow' and keep your ticket. In this case, it makes sense, as the ticket is no longer valid. However, when you verify your identity online, this is usually using documents that do not expire the moment after you've used them. That would be like having to apply for a new passport every time you travel abroad, only for the border staff to keep it when you hand it over at passport control. This, of course, does not happen. You apply for a passport once, and you can use it many times in many situations to prove who you are.

In a digital context, this same concept is of great value to both individuals and organisations. Instead of the information flowing one way (you provide the evidence, the evidence is verified, and then discarded), it can flow two ways. When you provide proof about an aspect of your life, and that proof is verified, you receive a 'token' that says this occurred and that what you claimed about yourself is true. You keep this token, and you can present it at another time when you need to prove the same or a similar thing about yourself, in another

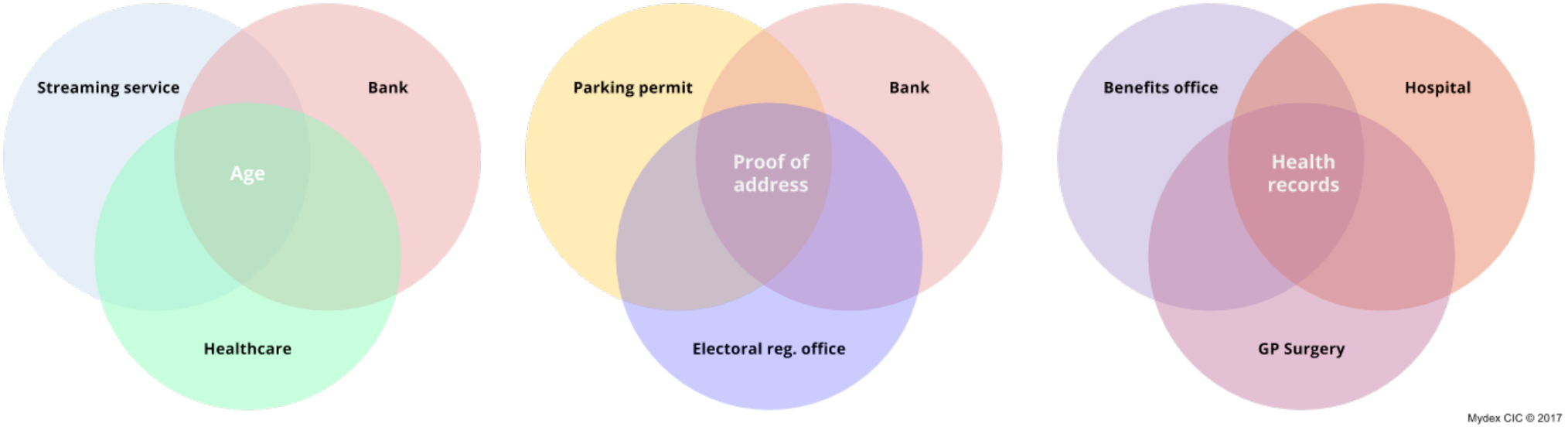
context. Your ticket doesn't get swallowed, you get to keep it and use it on another train.

Once again, let's imagine this scenario extrapolated to a realistic extent. There might be quite a large number of services you want to access that require the same proof. Suddenly, you are able to provide them with that proof with little-to-no effort, and they are able to accept it with next-to-no cost, as it has already been verified. Instead of cloning yourself and walking into bank branches, doctors' surgeries and border control with the same paper evidence that gets shredded each time, the single, 'real-life' you is allowing multiple organisations to view a single piece of relevant proof.



When a claim you make about yourself hasn't yet been verified, this will still have to happen, but once done one time, you keep this evidence and use it with other services that require proof about that thing—and so it continues. The more verified proof you collect, the more areas you have 'covered', until you have a collection of evidence that

Examples of overlapping proofs

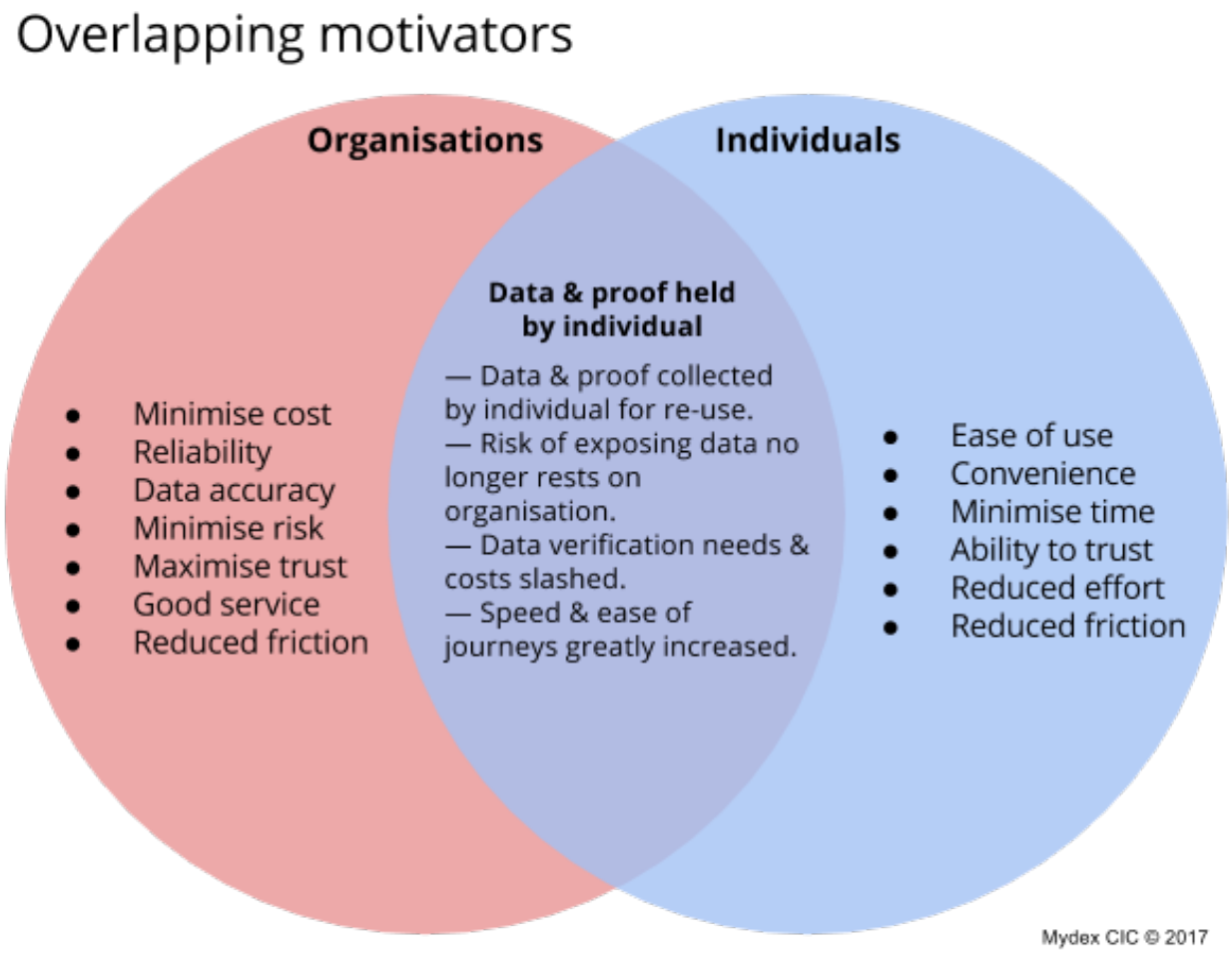


you can provide as and where necessary to get on with what it was you originally wanted to do.

More than just ownership

Let's not forget: we started all this talking about *convenience*. It's what drives so many of our choices, such as choosing to use social login, paying a little more to get something straight away, to save time, or to run an errand.

The case for giving data back to the individual can seem like a quest for ownership, motivated only by a sense of injustice. While the injustice is very real, and despite a rising tide of resistance to unauthorised mass-collection of data, as well as a greater public understanding that many of the 'free' services they are using are, in fact, subsidised by their own personal data, the most powerful behavioural motivator is still convenience and reward.



The result of this motivation is a rare overlap between the goals of businesses and 'consumers' (individuals). Individuals just want to get things done. The less friction they encounter in a journey, the better. They will, more often than not, take the path of least resistance, despite the potential costs to them. Organisations' traditional motivations still hold sway in the realm of digital services: minimise cost, increase reliability and accuracy, minimise complaints, maximise trust, and provide the best service possible within these margins.

Pressure from above and below

As if the common-sense, least effort approach wasn't enough, organisations providing services online are now being squeezed both from the top down and the bottom up. Key court battles and investigations into marketing, market competition and surveillance have given

rise to significant changes in regulation, most famously the upcoming General Data Protection Regulation (GDPR), coming into obligatory force in May 2018, which enforces significant changes to consent processes, data usage and data portability in the European Union. This has sent incumbents and startups alike scrambling for solutions to their processing of personal data, getting consent from their customers or users, and challenges popular business models centering on the mass collection and subsequent sale of personal data to third parties.

Since the infamous Snowden leaks of 2013, public discourse around personal data, its use, abuse and regulation has burgeoned enough to enter regular, mainstream media coverage. Cyber-attacks are an increasingly popular method of theft, political disruption or other motivated aggression. Consumer trust in brands in the context of data protection is at an all-time low, and platforms for consumers to make their voices heard are growing. Being the one big company in sector 'x' that hasn't suffered a recent data breach is becoming a valuable differentiator, and apps' inexplicably invasive data access requirements are being shunned in favour of privacy-protecting equivalents (secure messaging service Telegram added 40 million users in under a year, to reach 100 million, despite the existence of alternative such as WhatApp, which has garnered scrutiny and a shift in user trust since being acquired by Facebook in 2014).

From the top (regulators) and the bottom (consumers), then, we are seeing a demand for protection of personal data. Mix this with the ever-increasing demand for convenience, and options become limited for companies intent on processing and storing their users' personal data themselves. Ideas posited long ago that anticipated a shift in the relationship between organisation and individual, including Doc Searls' Project VRM (2006) and Alan Mitchell's book *Right Side Up* (2001), are gaining traction as the market is seemingly catching up to these concepts as the solution to two large and very real difficulties: giving individuals a good experience when interacting with online services, and minimising the risks of dealing with personal data on the internet.

From static to portable

An early noticeable impact of the shift away from the silo-based infrastructure of the Web 2.0 era was the reaction of incumbents such as Google and Amazon. Amazon—as discussed in a 2015 paper by the Boston Consulting Group and reviewed by Mydex—

broke their services down into components that allowed them to provide infrastructure for a vast range of online ventures and platforms, including some of their competitors (Netflix, for example, is hosted on Amazon Web Services, despite Amazon's Prime Video service competing for market share in the streaming industry).

As the BCG point out, massive corporations recognised that problem-solving was becoming a sort of 'mix and match' activity: instead of building an all-encompassing solution for a technological problem, small components—really good at a particular small set of things—are joined together via interoperable APIs to give the individual the end result they are looking for. Enabled by the advent of a reasonable level of consensus in terms of open standards and web protocols, 'integrations' are now a common selling point for many platforms ("we have integrations with all your favourite services including…"). But, as we concluded in our review of the Boston Consulting Group's paper, the individual has to be able to be agile and flexible as well, and this requires them to have the means to do so.

Dissolving the asymptote

Many service diagrams would show you in great detail how all the components of a modern platform link together to process data and serve up the result that someone is expecting to see. The individual using the platform comes in from the top, above the user-experience layer (what you see when you open an app or website), inputs their data, and gets something in return. Very few, however, place that individual under the visual layer, in the midst of the clever joined-up services, APIs and data processing.

But consider the following assumptions, drawn from what has been discussed thus far:

1. In order for people to be able to access certain services online, they need to be able to prove things about themselves.

2. In order to minimise the cost of proving these things, and in order to maximise the efficiency on the sides of both the individual and the organisation, the individual needs to easily be able to collect and reuse this proof.

3. The risk and regulatory difficulty of exposing large sets of personal data to the internet as a means of making them useful is becoming prohibitive for organisations on a number of fronts (financial, customer loyalty, threat vector).

4. In order for agile, interoperable and joined-up services to be of benefit to individuals, they need to work with personal data.

Taking these four assumptions into account, consider this hypothesis: ***the individual and their personal data has to be involved as a component in the process of providing services online, be integrated into the flow of information and the structure of the service, not just as a start or end point 'looking in from the outside'.***

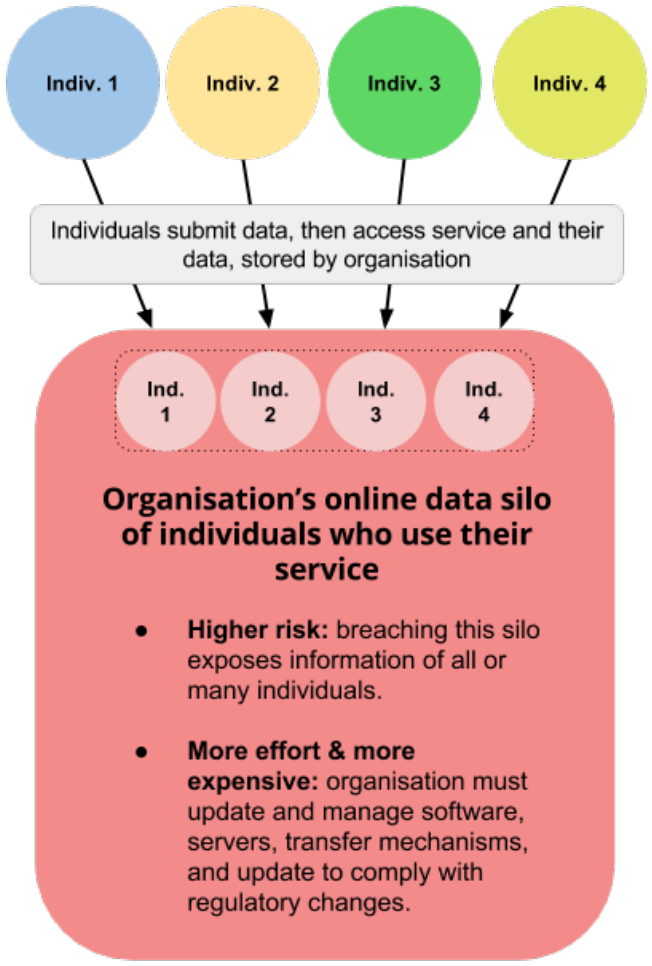
Not only this, but one could go as far as to say that the individual must be the *centre* of these processes, or what we have come to describe as the 'point of integration'. In order for a group of components, a group of organisations, or a group of platforms to access the same data—to connect with the single, 'real' individual who is trying to access their services—it is necessary for the individual themselves to be the centrepoint from which the crucial personal data is accessed, updated, collected, co-curated. This data, a potentially very diverse range of information, verified from previous interactions, can be presented as and when it is necessary to do so, to the relevant parties, as the key piece in the puzzle of online service provision.

Impact

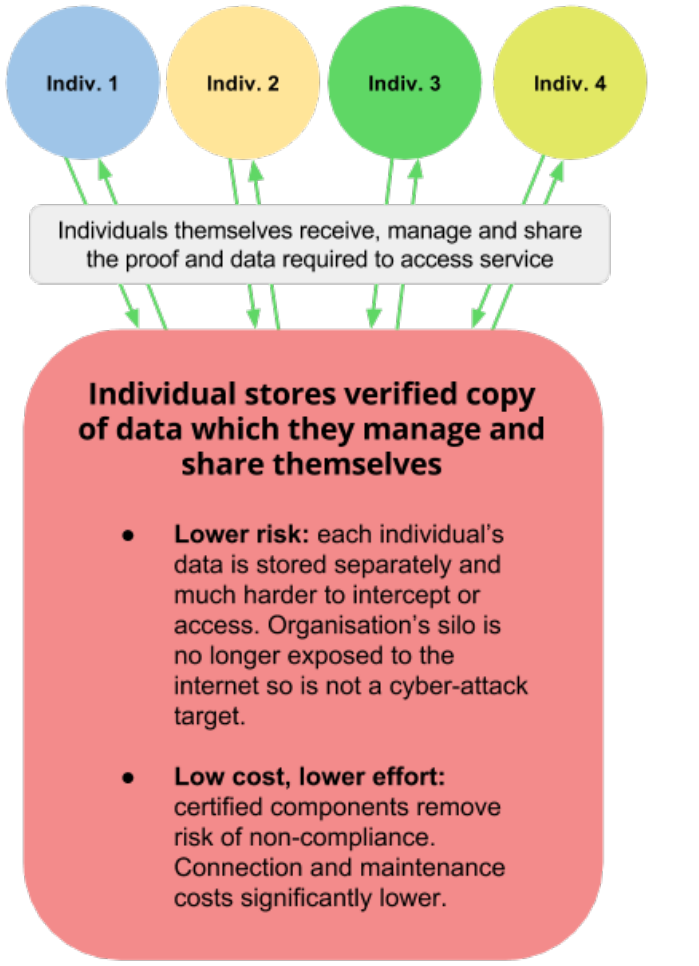
Managed by the individual and kept under their lock and key, there are a variety of short and long-term impacts for organisations. In the short term, 'offloading' the transfer and protection of the individual's data to the individual themselves (supported by the appropriate tools) absolves them of the demands to handle data in accordance with fast-approaching and stringent regulation. The organisation's internal data silos no longer need to be exposed to the internet.

It also reduces the threat vectors present for a cyber attack, handing the responsibility of protecting personal data, and the correct implementation of consent models, to services that are specifically designed to deal with such matters. By their very nature more likely to be successful in doing so, in that their entire purpose is to be the components for consent, security and all things personal-data related.

Personal data stored by organisation



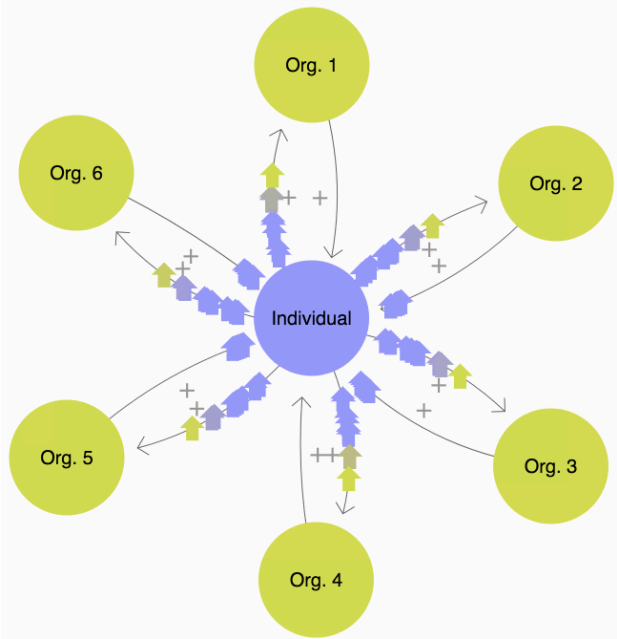
Personal data stored by individual



The long-term impact is diverse. The more verified data an individual can accumulate, i.e. the more organisations giving a verified piece of proof back to the individual, the likelier it is that the individual will already have the necessary piece of verified information to provide the proof required to access another service. In other words, the more verified data an individual can accumulate, the better it is both for organisations and for themselves. How does it benefit organisations? There are a number of costly snags that can be reduced or eliminated once an individual has even a few verified attributes which they look after:

1. Verification — the most obvious candidate, verification costs are high. Being able to receive a verified token of information that is common to many services would remove this cost.
2. Data accuracy — sending things to the wrong place and not knowing if they got there is an expensive business. A recent Royal Mail report estimates that it is costing UK organisations 5.9% of their annual revenue.
3. Fines — most organisations would rather not be fined €20m or 4% of their annual turnover for misuse of personal data.

If that's all a bit of a mouthful, the depiction below shows, at a surface level, the 'virtuous circle' that is created by these proofs and data flowing in both directions—to and from the individual and the organisations they connect with.



Simplicity

The complexity of the tasks we need to complete may increase as technology helps us do more in less time. But what about the complexity of the data itself? Many types of data and metadata can be generated as the result of the use of a service, but how complex and varied are the datasets that

we need to prove things about ourselves? In a large number of cases, we are asked time and again to show evidence for the same things. Things that are not highly complex and obscure digitally-generated datasets, but basic facts about our lives.

What happens now?

It would be unrealistic to expect this shift in the way data is exchanged and distributed to happen overnight. That is not to say that the technology doesn't exist. We at Mydex CIC have created the Mydex Platform, which has been live, certified and operational since 2007. With each passing year, the discussion, demand and developing legislation reflects the shift in perspective towards one that involves the individual as a component in the ways described above.

However, the requirements of the GDPR do herald the need for acceleration in certain key areas of personal data and information exchange. The enforcement for organisations to make individuals' data portable is a good example. This requirement makes a compelling case for organisations to—at the very least—give their customers, citizens, users a verified copy of their data. In the first instance, this could allow individuals to co-curate their data with the organisation they are using, keeping it more up-to-date and more 'complete'—something we know is of great value to many organisations (see 'Impact' above).

Individuals could also reuse this data to prove things about themselves, as discussed, or to switch to another service provider without losing the benefits of the data accumulated over the course of their involvement with that organisation. This is a more realistic short-term reality, and one that we welcome as a step towards the massive benefits that this change in information flow could bring.

Not just for us

It is important to remember that the need to prove things about yourself on the internet impacts the whole of society, not just the privileged. Solutions for personal data cannot just be aimed at the tech-savvy or the early-adopter. If data is the backbone of the emerging digital foundations with which we enhance and develop our real lives, and if personal data is the missing piece of this puzzle, this affects everyone. Imagine the person leaving prison who has no official documents to their name and struggles to get a foothold without evidence about their life. Or the elderly resident who is entitled to energy benefits but doesn't have the means to gather all the necessary physical evidence and travel in person to

the office responsible for making sure these are in place.

As Jos Creese said in his speech at the BCS President's Dinner in 2016, the IT industry is no longer just about computers and clever programming. The computing world is—and will continue to have—a vital and lasting impact on society and on the world. Considering personal data, trust and computing infrastructure in the context of society is a responsibility for those working to create the components that drive this innovation onwards.

About the author

Jack Mitchell

Jack, Head of Communications at Mydex CIC, has worked with Mydex since 2013 and been a keen observer of the shifts described in this paper. He has written extensively about the development of the digital ecosystem, the place of personal data within that development, and lectured on digital identity.

Email: jack@mydex.org

About Mydex CIC

Mydex is a Community Interest Company and social enterprise solving trust, identity and data management. The Mydex Platform provides components for personal data management, consent management and identity management. Organisations can use these to become regulation compliant, easily launch new services or incorporate them into existing processes. Only the individual can see the data held in their PDS, and when they share it, it is visible only to them and the third party to whom they have given access.

Mydex is an ISO27001 certified company for information security management, a FairData company and a certified G-Cloud supplier.

Web: mydex.org
Blog: medium.com/mydex
Twitter: twitter.com/mydexcic
Documentation: dev.mydex.org