WORLD ECONOMIC FORUM

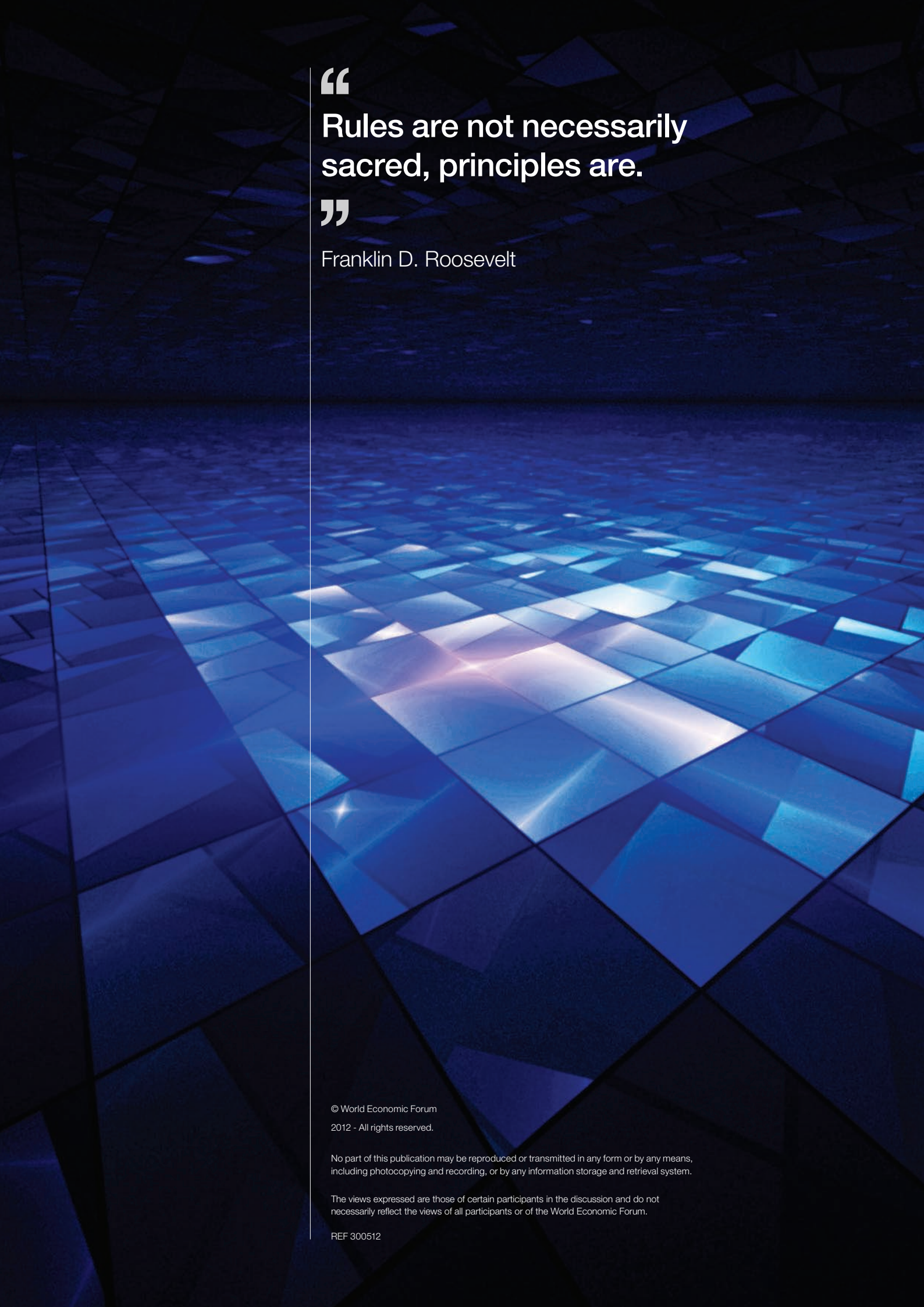
COMMITTED TO IMPROVING THE STATE OF THE WORLD

Industry Agenda

Rethinking Personal Data: Strengthening Trust

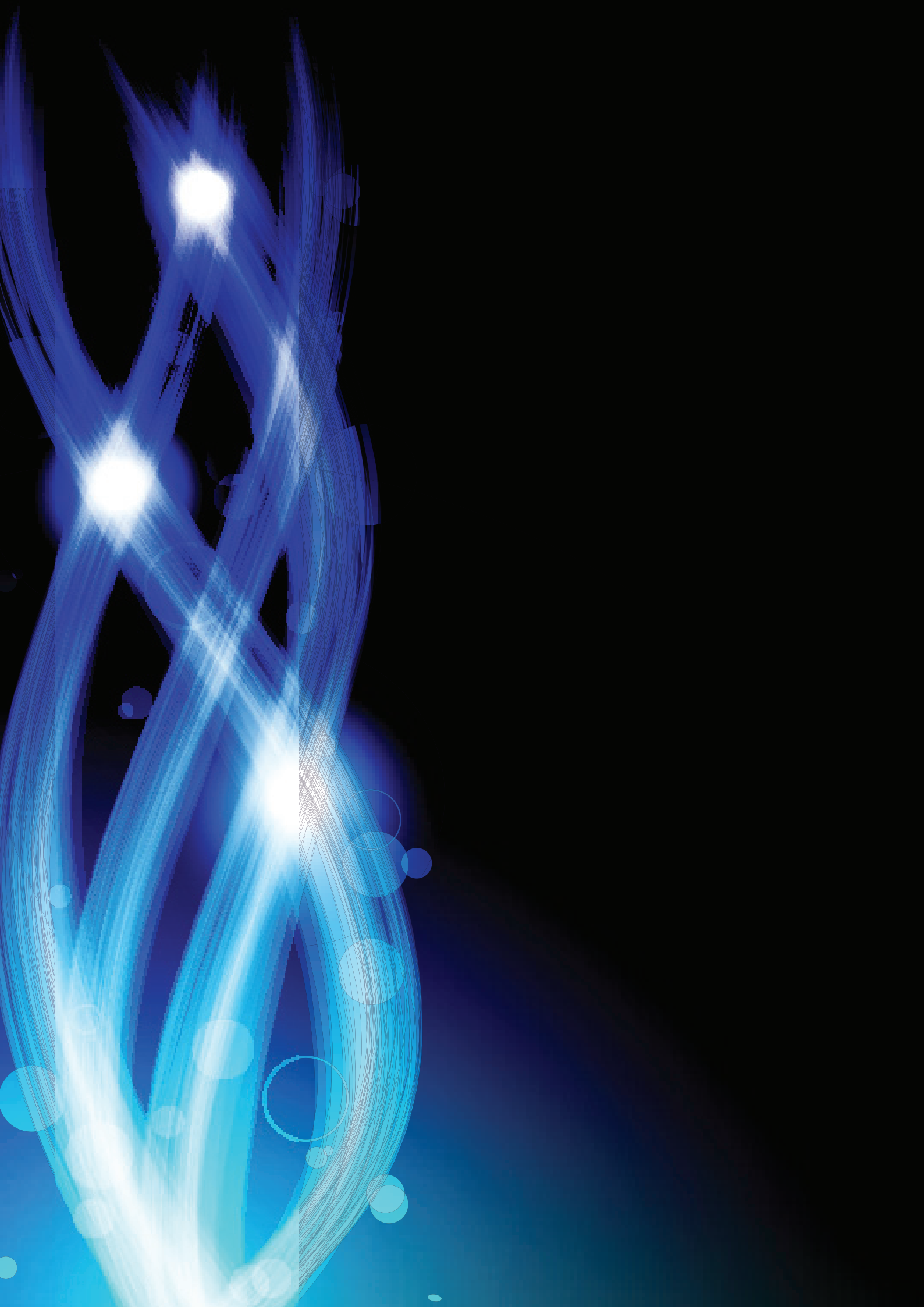
Prepared in collaboration with The Boston Consulting Group

May 2012



REF 300512

Contents



Executive Summary

The explosive growth in the quantity and quality of personal data has created a significant opportunity to generate new forms of economic and social value. Just as tradable assets like water and oil must flow to create value, so too must data. Instead of closing the taps or capping the wells, all actors can ensure that data flows in a measured way. But for data to flow well, it requires the same kinds of rules and frameworks that exist for other asset classes.

The reality has been quite different, however. High-profile data breaches and missteps involving personal data seem to be reported almost daily by the media. Tension has arisen between individual perceptions of harm and powerlessness versus organizational feelings of control and ownership. The result: a decline in trust among all stakeholders. Individuals are beginning to lose trust in how organizations and governments are using data about them, organizations are losing trust in their ability to secure data and leverage it to create value, and governments are seeking to strengthen trust to protect an individual's privacy. Yet, at the same time, consumers continue to share personal data and online retail continues to grow.

Among the three actors – individuals, organizations and governments – dialogue about personal data is currently anchored in fear, uncertainty and doubt. Together, these issues have the potential to undermine the economic and social wealth possible from this new asset. All stakeholders in the ecosystem face a challenge of unprecedented size, speed and complexity. Rules and norms change faster in a hyperconnected world and potentially outstrip the ability of traditional rule-setting approaches to keep pace. But, there is debate among different stakeholders and different regional jurisdictions on the best approach for establishing rules that allow data to flow in a trusted manner.

To restore trust, this report proposes three separate, but related questions, which need to be addressed by all stakeholders:

1. Protection and Security: How can personal data be protected and secured against intentional and unintentional security breach and misuse?

2. Rights and Responsibilities for Using Data: How can rights and responsibilities, and therefore appropriate permissions, be established for personal data to flow in ways that both respect its context and balance the interests of all stakeholders?

3. Accountability and Enforcement: How can organizations be held accountable for protecting, securing and using personal data, in accordance with the rights and established permissions for the trusted flow of data?

Answers to these questions will not be easy. Stakeholders have different cultural norms, timeframes for action and paths to a potential solution. Different regions are at different stages of this process of establishing a framework for dealing with personal data. However, the global nature of data flows suggests that leaders need to work together to achieve a coordinated yet decentralized approach to this challenge.

This report recommends that all the stakeholders take four main steps:

1. Engage in a structured, robust dialogue to restore trust in the personal data ecosystem. The debate needs to focus on achieving consensus on some of the key tensions, including securing and protecting data, developing accountability systems, and agreeing on rules for the trusted and permissioned flow of data for different contexts. Central to this dialogue is the inclusion of individuals, who play an increasingly important role as both data subjects and as data creators.

2. Develop and agree on principles to encourage the trusted flow of personal data. The simple slogan of "think globally, act locally" can help frame these principles (i.e. shared principles can help all the actors aim towards the same outcomes, even if their approaches for how to get there differ).

3. Develop new models of governance for collective action. Regulators, organizations and individuals can play complementary roles in establishing accountability systems, enforcement mechanisms, rights and permissions.

4. Establish "living labs". Given the complex social, commercial, technical and regulatory uncertainties and interdependencies, an environment which can provide stakeholders with the ability to test and learn in real time (and at scale) needs to be established. These labs can provide a safe context for more fully understanding the system dynamics and collectively identifying shared opportunities, risks and the means for effective collaboration.

Chapter 1: The Personal Data Landscape

Introduction

The digital world is awash in personal data.[1] Every day, people send 10 billion text messages, make 1 billion posts to a blog or social network and generate millions of entries into their growing electronic health records.

In addition, with approximately 6 billion mobile telephone subscriptions in the world, it is now increasingly possible to track the location of nearly every person on the planet, as well as their social connections and transactions.[2] And mobile phones are not the only devices recording data – Web applications, cars, retail point-of-sale machines, medical devices and more are generating unprecedented volumes of data as they embed themselves into our daily lives. Estimates are that by 2015, 1 trillion devices will be connected to the Internet.[3]

Companies and governments are using this ocean of Big Data to unleash powerful analytic capabilities. They are connecting data from different sources, finding patterns and generating new insights – all of which adds to the ever deepening pool of data.

In short, the growing quantity and quality of personal data creates enormous value for the global economy. It can help transform the lives of individuals, fuel innovation and growth, and help solve many of society's challenges. As the first-phase World Economic Forum report on personal data elaborated, personal data represents an emerging asset class, potentially every bit as valuable as other assets such as traded goods, gold or oil.

Historically, the strength of a major economy is tightly linked to its ability to move physical goods. The Silk Route, the Roman roads and the English fleet all served as the economic backbones connecting vast geographies. Even though it is a virtual good, data is no different. Data needs to move to create value. Data sitting alone on a server is like money hidden under a mattress. It is safe and secure, but largely stagnant and underutilized.

As an emerging asset class, personal data currently lacks the rules, norms and frameworks that exist for other assets. The lack of trading rules and policy frameworks for its movement have resulted in a deficit of trust among all stakeholders and could undermine the long-term potential value of personal data. Different jurisdictions are looking to tackle this deficit of trust through different approaches, ranging from fundamental rights-based approaches to harm-minimization approaches.

Several characteristics of personal data make establishing rules and frameworks uniquely challenging:

- The digital nature of personal data means it can be copied infinitely and distributed globally, thereby eliminating many of the trade barriers which exist for physical goods.

- Data, unlike most tangible assets, is not consumed when used; it can be reused to generate value.

- Data grows ever more connected and valuable with use. Connecting two pieces of data creates another piece of data and with it new potential opportunities (as well as new potential harms).

- The role of the individual is changing. Individuals are no longer primarily passive data subjects. They are also increasingly the creators of data. In addition, personal data is intimately linked with an individual's background and identity, unlike interchangeable commodity goods.

[1] There are many definitions of personal data across different jurisdictions and even different sectors within the same jurisdiction. For the purposes of this report, personal data is used to refer to data and metadata relating to a specific, identified or identifiable person. This is similar to definitions used by both the proposed US Privacy Bill of Rights and the EU's Data Protection Regulation.

[2] "Key Global Telecom Indicators for the World Telecommunication Service Sector". International Telecommunications Union, http://www.itu.int/ITU-D/ict/statistics/at_glance/KeyTelecom.html.

[3] King, Rachel. "IBM panel discusses tackling big data storage as problem escalates". smartplanet, September 2011, http://www.smartplanet.com/blog/smart-takes/ibm-panel-discusses-tackling-big-data-storage-as-problem-escalates/19010.

All stakeholders in the personal data ecosystem face a challenge of unprecedented size, speed and complexity. Rules and norms change faster in a hyperconnected world and outstrip the ability of traditional rule-setting approaches to keep pace. Solutions that focus on isolated examples or one-size-fits-all approaches quickly grow outdated as social, commercial and regulatory contexts change. Enterprises, governments and individuals need to creatively collaborate to develop new rules and frameworks that are both robust enough to be enforceable, yet flexible enough to accommodate the world's accelerating and constant change.

The Opportunity

Personal data plays a vital role in countless facets of our everyday lives. Medical practitioners use health data to better diagnose illnesses, develop new cures for diseases and address public health issues. Individuals are using data about themselves and others to find more relevant information and services, coordinate actions and connect with people who share similar interests. Governments are using personal data to protect public safety, improve law enforcement and strengthen national security. And businesses are using a wide range of personal data to innovate, create efficiencies and design new products that stimulate economic growth.

Estimates are that the Internet economy amounted to US$ 2.3 trillion in value in 2010, or 4.1% of total GDP, within the G20 group of nations. Larger than the economies of Brazil or Italy, the Internet's economic value is expected to nearly double by 2016 to US$ 4.2 trillion.[4] But this growth could be severely constrained if the flow of personal data on which e-commerce and other economic activity depends becomes overly restricted.

Consider the many ways personal data can create economic and social value for governments, organizations and individuals:

- Responding to global challenges. Real-time personal data and social media can help to better understand and respond to global crises like disaster response, unemployment and food security. It represents an unprecedented opportunity to track the human impacts of crises as they unfold and to get real-time feedback on policy responses.[5] Or consider the case of Google Flu Trends, which uses individuals' ostensibly private flu-related search words and location data[6] to detect potential flu outbreaks in real time, as opposed to the weeks-old government data that currently exists.[7] Researchers have found that the data has a high correlation with upswings in emergency room activity, and it could provide the basis for early-warning systems to detect pandemics and save millions of lives.[8]

- Generating efficiencies. For centuries, increased access to information has created more efficient ways of doing business.[9] These days, organizations in every industry are using vast amounts of digital data to streamline their operations and boost overall productivity. For example, US$ 700 billion in health cost

[4] Dean, David, Sebastian DiGrande, Dominic Field and Paul Zwillenberg. "The Digital Manifesto: How Companies and Countries Can Win in the Digital Economy." The Boston Consulting Group. January 2012.

[5] UN Global Pulse, http://www.unglobalpulse.org/about-new.

[6] It is important to note that Google Flu Trends data can never be used to identify individual users because it relies on anonymized, aggregated counts of how often certain search queries occur each week.

[7] Google Flu Trends, http://www.google.org/flutrends/about/how.html.

[8] Dugas, A. F., Y. H. Hsieh, S. R. Levin, et al. "Google Flu Trends: Correlation with Emergency Department Influenza Rates and Crowding Metrics." Clinical Infectious Diseases, 2012. Described in *Science Daily*, http://www.sciencedaily.com/releases/2012/01/120109155511.htm.

[9] Gleick, James. *The Information: A History, a Theory, a Flood*. London: Fourth Estate, 2011.

savings in the US, or about 30% of total healthcare spending today, could result in large part through the increased flow of personal data. Improved information flow could reduce duplicative lab testing and imaging, fraud and inefficiencies, as well as lead to better care coordination and treatment.[10] In financial services, personal data is already being used to generate significant efficiencies through facilitating online commerce and payments as well as saving billions of dollars through fraud prevention.

- Making better predictions. Personal data is stimulating innovative new products tailored to and personalized for the specific needs of individuals. For example, Amazon's "Customers Who Bought This Also Bought" collaborative filtering tool suggests related items to buy that customers might not have discovered otherwise.[11] In financial services, tailored insurance products are being developed based on devices that track driving behaviour rather than just age, gender and neighbourhood. News and content websites can customize the articles each individual views based on their interests and preferences. In addition, organizations can use business intelligence derived from the aggregation of millions of individual consumer transactions to prepare for likely events. For example, before a hurricane strikes, Wal-Mart knows to stock its shelves with not only flashlights and batteries, but also with Pop-Tarts.[12]

- Democratizing access to information. Consumers benefit from "free" services like search engines, e-mail, news sites and social networks that previously either did not exist or have a significant monetary cost in other forms in the offline world. However, individuals are beginning to realize that targeted advertising, based on data about them and their online behaviour, [13] fuels most of these ostensibly free services either directly or indirectly. As is widely quoted online: "If you're not paying for something, you're not the customer; you're the product being sold."[14]

- Empowering individuals. Empowered consumers are taking greater control over the use of data created by and about them. Rather than individuals being passive agents, they are engaging with organizations in a collective dialogue.[15] As a result, people are starting to volunteer information that only they know. Such volunteered personal information (VPI) includes their updated contact details as they happen, reasons why they took an action or made a purchase, and future plans and preferences. Some observers have estimated that VPI could reach approximately US$ 32 billion in value by 2020 in the UK alone.[16] In addition, individuals are using the information they share about themselves, their beliefs and their preferences to connect like never before. In the past few years, social media tools built on the foundation of widely shared personal data have played a key role in bringing down governments.

The above list is far from comprehensive. Some of the most important value-creation opportunities from personal data remain as yet unknown. Most personal data still resides in silos separated by different technology standards and legal contracts. And a lack of an effective system of permissions prevents data from moving in a trusted and secure way to create value. Creating such systems will allow ever greater "data leverage". But such opportunities are not without potential harms and risks, which manifest themselves in a loss of trust among all stakeholders in how personal data is protected and used.

Given the variety of applications in which personal data can be leveraged, estimating the impact of the "economics of trust" is difficult to measure. Existing research by BCG on the Internet economy in the G20 has forecast that online retail will grow to US$ 2 trillion by 2016.[17] However, this estimate is influenced by consumer perception of trust in how personal data is used. Online retail could grow even faster to US$ 2.5 trillion by 2016 with enhanced trust or to only US$ 1.5 trillion if trust were to be eroded. Given that this US$ 1 trillion range is from just one small part of the broader personal data ecosystem, it provides an indication of the magnitude of the potential economic impact when other sectors (health, financial services, etc.) are considered – potentially in the tens of trillions of dollars.

The Evidence of a Decline in Trust

Ample evidence suggests that there is a decline in trust in the personal data ecosystem. Many of the existing regulatory, commercial and technical mechanisms for strengthening trust are no longer fit to do the job. The widespread loss of trust is unmistakable: security breaches, identity theft and fraud; concern from individuals and organizations about the accuracy and use of personal data; confusion from companies about what they can and cannot do; and increasing attention and sanctions from regulators.

Some recent events serve as leading indicators of the potential for future instability on a much more massive scale. In 2011, Sony revealed breaches in its online video game network that led to the theft of names, addresses and possibly credit card data belonging to more than 100 million accounts, in what was one of the largest-ever Internet security break-ins.[18] Experts said the breaches could cost Sony and credit card issuers US$ 1-2 billion.[19]

Data breaches are also growing more anarchic and motivated by state-sponsored political and anti-business sentiments. The loose-knit hacking movement known as Anonymous claimed to have stolen thousands of credit card numbers and other personal information belonging to clients of US-based security think tank Stratfor.[20]

[10] "Where Can $700 Billion in Waste Be Cut Annually from the U.S. Healthcare System?" Thomson Reuters, http://www.factsforhealthcare.com/whitepaper/HealthcareWaste.pdf, 2009.

[11] Tene, Omer and Jules Polonetsky. "Privacy in the Age of Big Data". *Stanford Law Review*, 2 February 2012. http://www.stanfordlawreview.org/online/privacy-paradox/big-data.

[12] "A Different Game". *The Economist*. 25 February 2010. http://www.economist.com/node/15557465.

[13] Most online advertising today relies on anonymous cookies and is linkable to a device rather than an individual.

[14] Lewis, Andrew. MetaFilter Weblog. 26 August 2010. http://www.metafilter.com/95152/Userdriven-discontent#3256046.

[15] Searles, Doc. *Cluetrain Manifesto*. http://www.cluetrain.com.

[16] Ctrl-Shift, http://www.ctrl-shift.co.uk/themes/volunteered_personal_information.

[17] For country-by-country forecasts see Dean, David, Sebastian DiGrande, Dominic Field, Andreas Lundmark, James O'Day, John Pineda and Paul Zwillenberg. "The Internet Economy in the G-20: The $4.2 Trillion Growth Opportunity." The Boston Consulting Group. March 2012.

[18] Baker, Liana B. and Jim Fincle. "Sony PlayStation Suffers Massive Data Breach". Reuters. 26 April 2011. http://www.reuters.com/article/2011/04/26/us-sony-stoldendata-idUSTRE73P6WB20110426; "Sony suffers second data breach with theft of 25m more user details". *The Guardian*, Technology Blog. http://www.guardian.co.uk/technology/blog/2011/may/03/sony-data-breach-online-entertainment.

[19] Miller, Mary Helen. "Sony data breach could be most expensive ever". *Christian Science Monitor*. 3 May 2011. http://www.csmonitor.com/Business/2011/0503/Sony-data-breach-could-be-most-expensive-ever.

[20] "Anonymous targets US security think tank Stratfor". Associated Press. 25 December 2011. http://www.guardian.co.uk/technology/2011/dec/25/anonymous-security-thinktank-stratfor

In addition, some data breaches happen accidentally or unintentionally; think of the employee who leaves an unencrypted laptop on a train containing thousands of data records. And breaches are not limited to the companies that originally collected the personal data, which adds to the loss of trust. In 2011, hackers compromised the database of Epsilon, a marketing company responsible for sending 40 billion marketing e-mails on behalf of 2,500 customers, including such companies as Best Buy, Disney and Chase.[21] Very few customers had ever heard of Epsilon, nor did they know it held data about them.

Stakeholders Have Different Perspectives and Concerns

Tension is rising. Individuals are growing concerned that companies and governments are not protecting data about them and that they are instead using it in ways not necessarily in their best interests. Many organizations are struggling to protect and secure the explosion of data they have access to, and they are unsure what they can and cannot do with it.

Governments are trying to strike a balance between protecting individuals and encouraging innovation and growth. Such uncertainty creates instability, which in turn manifests itself differently across three main groups of actors in the personal data ecosystem – individuals, government policy-makers and organizations.

Individuals
Surveys show that individuals are losing trust in how data about them is being collected, used, shared and combined by both organizations and governments. For example, according to European Justice Commissioner Viviane Reding, 72% of European citizens are concerned that their personal data may be misused, and they are particularly worried that companies may be passing on their data to other companies without their permission.[22] However, a disconnect exists between what people say and what they do; the world of personal data is no exception. While many people say they care about privacy, they also share information quite widely on social networks and elsewhere online.

A large part of concern stems from the fact that individuals often sign up to services not knowing how their data will be protected or if it will be shared. While legally they have given organizations their consent to the stated rules for usage, few individuals actually read these privacy policies or terms of services. Individuals therefore have little visibility into the practices of the organizations they are putting their trust in – until their data is breached or misused. As government's use of personal data grows, concern likewise grows over government protections of individual privacy.

Another concern of individuals is how to manage their online identity and the different aspects of their digital lives. Health data about an individual has a different impact when shared in a healthcare, work, family or social context. The lack of contextual control and permissions represents another cause for concern among individuals. At the moment, one of the few ways for individuals to keep different parts of their digital lives separate is to use different names and e-mail addresses for different contexts, to use pseudonyms, or to prevent their data being captured or linked to them in the first place.

Government Policy-makers
Policy-makers and regulators around the world share similar objectives – to stimulate innovation and economic growth while at the same time protecting individuals from harmful uses of personal data. Striking this balance lies at the heart of the tension facing the regulatory community. However, regulators are growing concerned about the loss of individuals' trust in how organizations are using and protecting data about them. Meanwhile, different regulators are taking different approaches to balancing these objectives, potentially adding to the instability as they work out how to protect individuals through new laws, policies and regulations. This ranges from privacy bills of rights and "do-not-track" options for consumers, to requirements that consumers be granted access to their personal data. A more comprehensive approach is needed that does not limit the solution to one sector or jurisdiction to allow for a global flow of data.

Following the release of the National Strategy for Trusted Identities in Cyberspace in 2011, which focused on identity assurance to support online interactions, the US government has outlined a comprehensive online Privacy Bill of Rights that aims to give individuals more control over how personal information about them is used on the Internet.[23] The Privacy Bill of Rights includes the goals of individual control, transparency, respect for context, security, access and accuracy, focused collection and accountability. The government intends to work with stakeholders to translate the document into specific practices or codes of conduct, to develop legislation based on these rights and to provide the US Federal Trade Commission (FTC) with enforcement authority. The approach builds on existing US self-regulation practices in which the FTC steps in to enforce unfair or deceptive practices.

In Europe, on the other hand, the European Commission has approached the issue from the perspective of protecting fundamental rights, although it shares with the US the same common goal of providing individuals with greater control to restore trust. The proposed Data Protection Regulation issued in January 2012 includes a requirement that Internet companies obtain explicit consent from consumers about the use of their personal data and delete the data forever at the users' request or face the prospect of fines for failing to comply.[24] The rules would also extend the reach of European law to companies outside the EU that handle data relating to EU residents. This raises, however, jurisdictional questions given the global nature of data flows. While many observers have praised these efforts as an important step in giving individuals more control over data about them, some have raised questions about the possible unintended consequences on innovation and economic growth and the use of data for purposes such as health research.[25]

Many of these regulatory requirements have reasonable motives, but could also have unintended consequences that might undermine economic and social value. For example, the Indian government introduced rules in April 2011 it said would protect individuals,

[21] Lennon, Mike. "Massive Breach at Epsilon Compromises Customer Lists of Major Brands". *Security Week*, 2 April 2011. http://www.securityweek.com/massive-breach-epsilon-compromises-customer-lists-major-brands.

[22] Reding, Viviene. "The EU Data Protection Reform 2012: Making Europe the Standard Setter for Modern Data Protection Rules in the Digital Age." Speech at the Innovation Conference Digital, Life, Design, Munich, Germany, 22 January 2012. http://ec.europa.eu/commission_2010-2014/reding/pdf/speeches/s1226_en.pdf.

[23] "We Can't Wait: Obama Administration Unveils Blueprint for a 'Privacy Bill of Rights' to Protect Consumers Online". Office of the Press Secretary, US White House, 23 February 2012, http://www.whitehouse.gov/the-press-office/2012/02/23/we-can-t-wait-obama-administration-unveils-blueprint-privacy-bill-rights.

[24] Sengupta, Somini. "Europe Weighs Tough Law on Online Privacy". *The New York Times*. 23 January 2012. http://www.nytimes.com/2012/01/24/technology/europe-weighs-a-tough-law-on-online-privacy-and-user-data.html?_r=1&pagewanted=al.l.

[25] For example, see the US government's position on the draft EC regulation available at http://www.edri.org/files/US_lobbying16012012_0000.pdf; http://blogs.law.harvard.edu/infolaw/2012/01/25/more-crap-from-the-e-u; http://www.pcworld.com/businesscenter/article/251573/proposed_eu_data_laws_under_fire_from_both_sides.html; http://www.marketingweek.co.uk/news/facebook-data-laws-could-stifle-innovation/4000557.article.

including requiring companies to take consent in writing from individuals about the use of the sensitive personal information they collect.[26] The requirement could have significantly constrained the US$ 75 billion Indian outsourcing industry, which employs 2.5 million people.[27] In August 2011, the Indian Ministry of Communications and IT issued a clarification effectively exempting outsourcers from the new law and calling into question the law itself.[28]

The role of regulators is likely to be very different when it comes to ensuring accountability for organizational stewardship of data as opposed to the setting of rules for what can and cannot be done with personal data. In the latter areas, the personal data ecosystem is increasingly too complex, fast-moving and global for traditional regulatory mechanisms to be effective. Some observers say approaches that treat governance as an afterthought and economic externality (through regulatory oversight and mechanisms such as notification and consent, hearings, rulings, legal challenges, injunctions and warrants) create huge costs, uncertain liabilities and problematic social acceptance.[29]

[26] Wugmeister, Miriam and Cynthia Rich. "India's New Privacy Regulations". Morrison & Foerster Client Alert. http://www.mofo.com/files/Uploads/Images/110504-Indias-New-Privacy-Regulations.pdf.

[27] "India's share in global outsourcing market rises to 55 pct in 2010". *International Business Times*, 3 February 2011. http://www.ibtimes.com/articles/108475/20110203/nasscom-global-outsourcing-share-india-it-bpo-sector-revenue-growth-of-it-bpo-sector.htm.

[28] Ribeiro, John. "India Exempts Outsourcers From New Privacy Rules". IDG News, 24 August 2011. http://www.pcworld.com/businesscenter/article/238706/india_exempts_outsourcers_from_new_privacy_rules.html.

[29] Clippinger, John Henry. "Design for a Trusted Data Platform and a Data-driven Bank: Overview and Next Steps". ID Cubed, January 2012.

Organizations

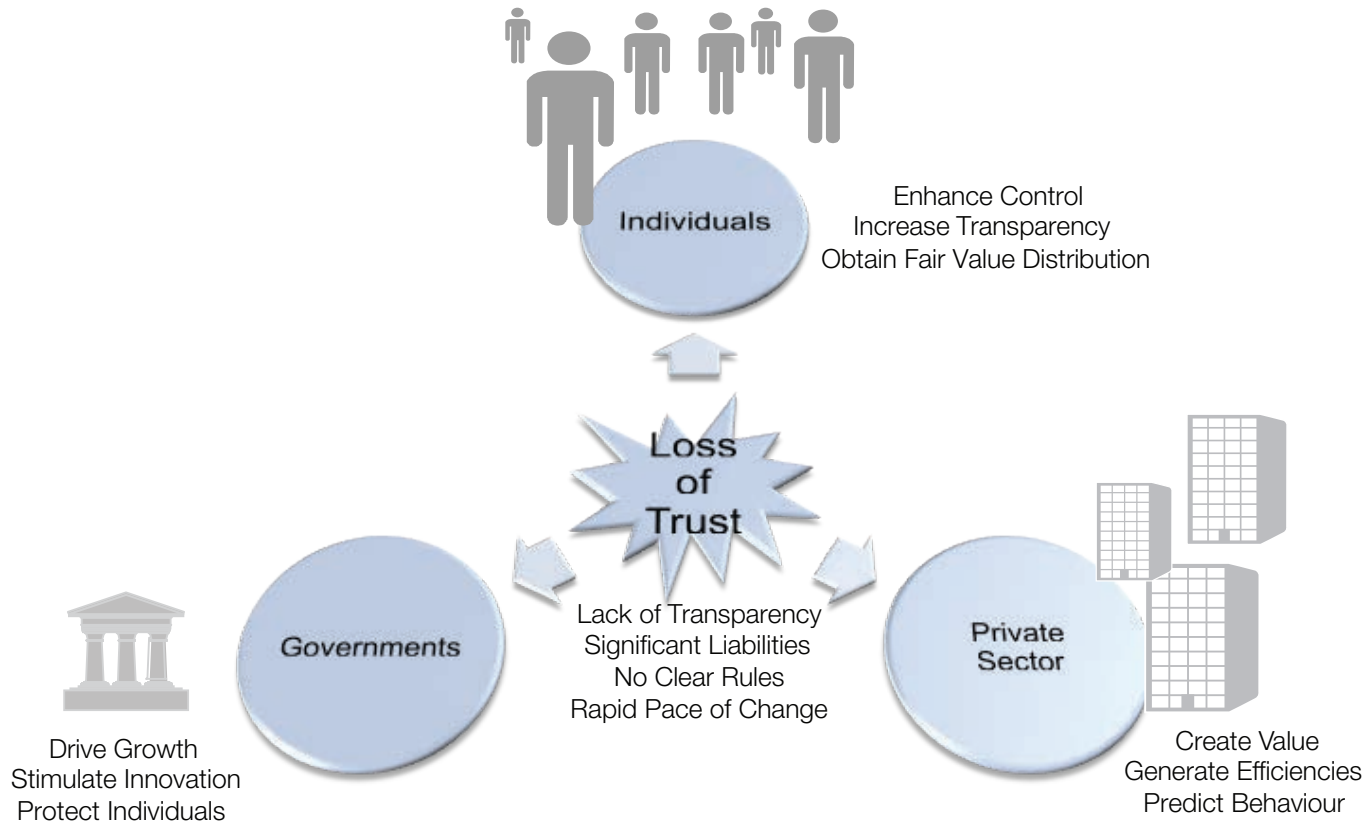
Commercial enterprises, not-for-profits and governments that have access to personal data are responding in unique ways to the explosion of information created by and about individuals. Some are unsure of the rules of the game and are concerned about legal liabilities and the negative brand impact of being seen as unfairly exploiting personal data given the heightened media attention on privacy and data breaches. As a result, some organizations are currently restricting the use of the personal data they hold and are underinvesting in the systems that can help to generate value from this data.

A number of private organizations, particularly those in the telecom sector, face significant legal and long-standing regulatory constraints on how they can use personally identifiable information. This lies in contrast with many of their competitors, which are using personal data much more freely to generate value. Others are innovating and developing new business models that offer individuals the tools to receive some form of control or even payment for the use of data about them.

The relationship individuals hold with Internet giants such as Google and Facebook was tested in 2011 when the FTC announced that it planned to heavily monitor the companies during the next 20 years. The settlements stemmed from charges related to the unauthorized sharing with third parties of personal data, failure to delete personal data from deactivated accounts and engaging in deceptive tactics that violated established privacy policies.[30]

[30] "Facebook and privacy: Walking the tightrope". *The Economist*. 29 November 2011. http://www.economist.com/blogs/babbage/2011/11/facebook-and-privacy; "FTC Charges Deceptive Privacy Practices in Google's Rollout of Its Buzz Social Network". Federal Trade Commission, http://www.ftc.gov/opa/2011/03/google.shtm.

Exhibit 1: A lack of trust has the potential to pull the ecosystem apart



In the meantime, some industry sectors are coming together to help establish best practices and norms to guide behaviour. In the marketing world, the Digital Advertising Alliance, which represents more than 90% of online advertisers, has established self-regulatory principles for online behavioural advertising.[31] The guidelines aim to give individuals a better understanding of and greater control over ads that are customized based on their online behaviour. They also establish specific limitations for the use of this data in cases when the potential harm is considered highest, such as employment, credit and healthcare eligibility information.[32] In addition, the Network Advertising Initiative has established monitoring and accountability mechanisms to encourage compliance with these rules.[33]

Similar efforts can be found in the mobile space with the recently announced GSM Association privacy principles.[34] In the online identity arena, the Open Identity Exchange was formed as a not-for-profit to support the creation of private sector-led legal rules and policy development and the creation of an open market for related identity and privacy products and services relating to online data and identity challenges.[35]

Governments are also increasingly using personal data for law enforcement and national security, including the monitoring of SMS messages, blogs, social network posts or geolocation data to fight criminal and terrorist activity, which is raising significant surveillance and privacy concerns. A 2011 report by the Brookings Institute noted that rapidly declining storage costs make it technologically and financially feasible for authoritarian governments "to record nearly everything that is said or done within their borders – every phone conversation, electronic message, social media interaction, the movements of nearly every person and vehicle, and video from every street corner".[36]

Governments are trying to improve service delivery and achieve significant cost savings by leveraging personal data. For example, a recent World Economic Forum report estimated the cost savings to emerging market governments could range up to US$ 100 billion per year with increased use of mobile financial services and the ability to utilize personal data more efficiently.[37]

Re-establishing Trust in a Complex Ecosystem

The existing dialogue about personal data is currently anchored in fear, uncertainty and doubt. This fear is potentially made worse by the increasingly shortened cycle between the discovery of an event or vulnerability and widespread media coverage. A researcher discovers that a website or mobile app has used data improperly, the mainstream press picks up the story, setting off a popular firestorm, which creates pressure on politicians to react.

While exposing company missteps is clearly important, the focus on fear and concern could result in a reactive and one-sided response. Unintended consequences could reduce the opportunities for value creation. It seems to be an intractable problem: how to create the rules and tools so that all stakeholders can capture the value from data in a trusted way.

[31] http://www.aboutads.info/obaprinciples

[32] http://www.aboutads.info/msdprinciples

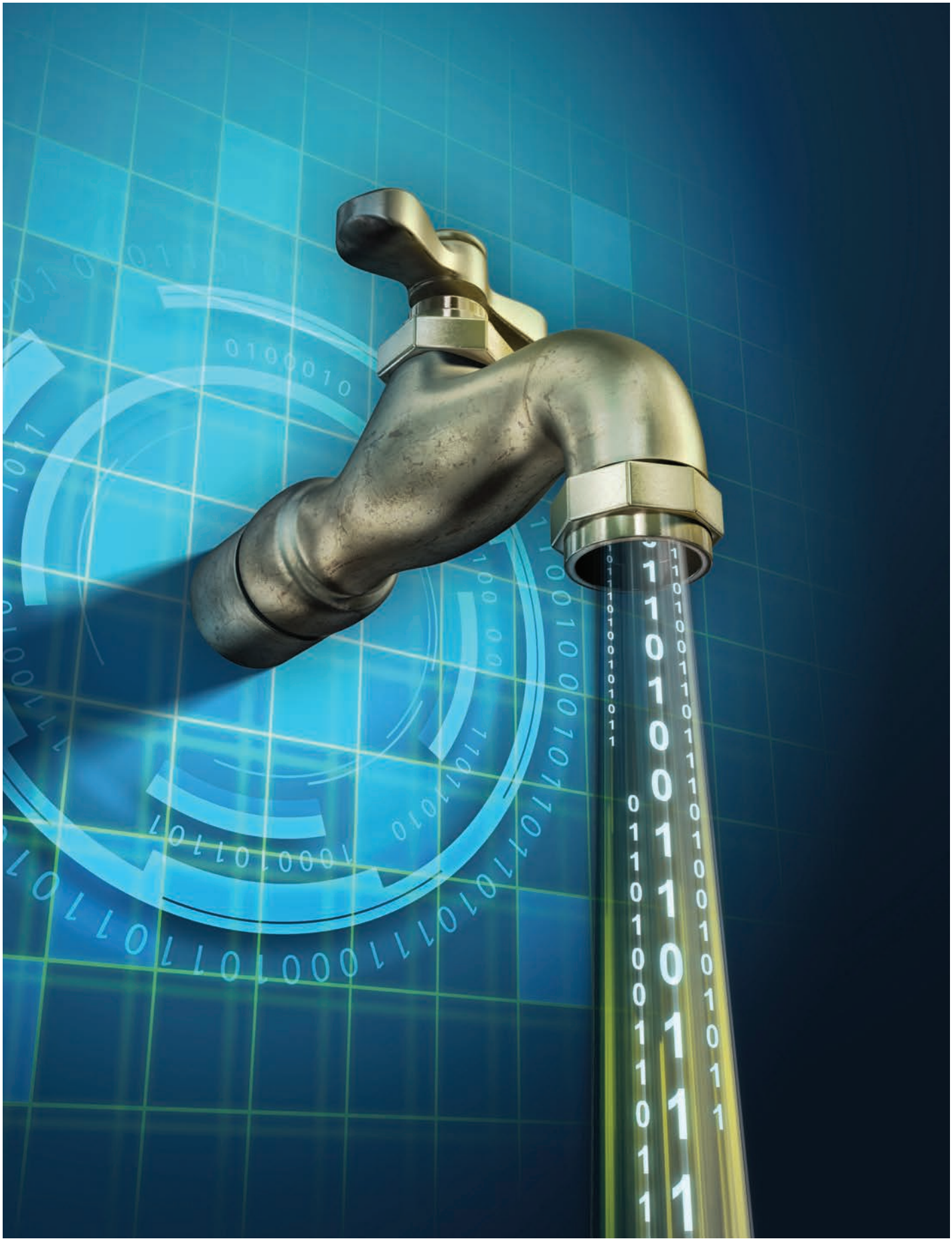
[33] http://www.networkadvertising.org/pdfs/NAI_2011_Compliance_Release.pdf

[34] http://www.gsma.com/mobile-privacy-principles

[35] http://openidentityexchange.org

[36] Villasenor, John, "Recording Everything: Digital Storage as an Enabler of Authoritarian Governments", Brookings Institute, 2011.

[37] *Galvanizing Support: The Role of Government in Advancing Mobile Financial Services*. World Economic Forum, March 2012.



Chapter 2: An Approach for Effective Dialogue

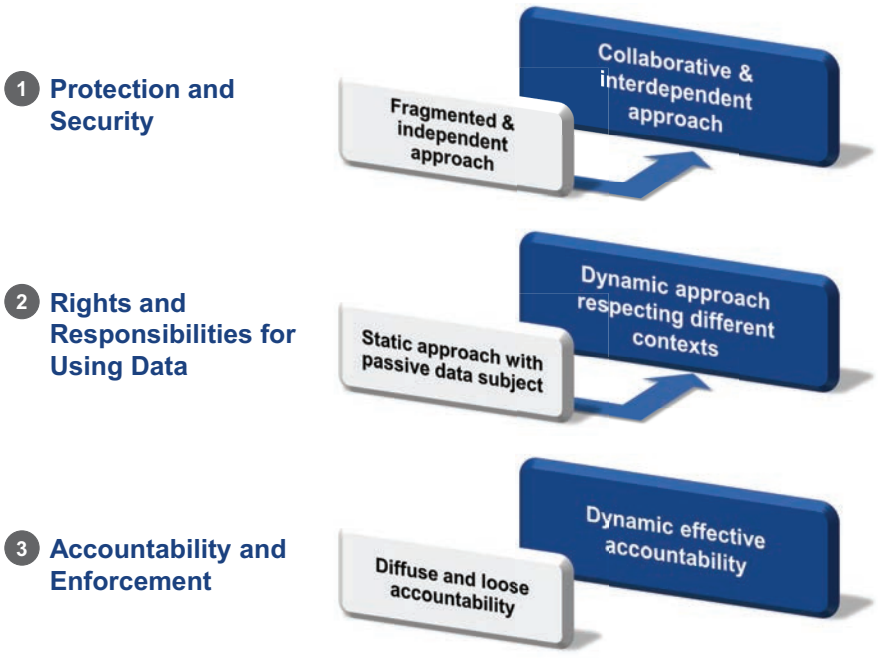
This chapter aims to address some of the key points of tension within the personal data ecosystem and to outline an approach that enables a richer understanding of the diverse issues and perspectives. The intent is to lay the foundation for a more focused and constructive dialogue that can address both the need to protect and secure personal data and to encourage its trusted flow to unlock value.

The chapter examines the competing ideas about control and harms, as exemplified in the current debates about data ownership and privacy. It argues that each of these competing ideas, as currently framed, is blocking the evolution of the debate.

One way of thinking about the issues incorporates a contextual mindset based on the collection, nature and use of data. To take a leap from the present debate to the future dialogue, all actors in the ecosystem need to focus on three separate but related questions (see Exhibit 2).

1. Protection and Security: How can personal data be protected and secured against intentional and unintentional security breach and misuse?
2. Rights and Responsibilities for Using Data: How can rights and responsibilities, and appropriate permissions, be established in order for personal data to flow in ways that both ensures the integrity of its context and balances the interests of relevant stakeholders?
3. Accountability and Enforcement: How can organizations be held accountable for protecting, securing and using personal data, in accordance with the rights and established permissions for the trusted flow of data?

Exhibit 2: Key issues in dialogue for achieving balance



Existing Barriers to Effective Dialogue

To get at the heart of the issues, the dialogue needs to unpack a number of key concepts in the current debate. They centre around two competing notions: one of ownership and control versus one of privacy and perceived harms.

A core point of confusion involves "what's mine, what's yours and what's ours?" Underlying this tension are competing notions of control. Another sticking point in the dialogue is the term *privacy*, which has competing definitions and an ambiguous taxonomy of harms, which are all interpreted differently.

The Problem of Ownership
The debate over who owns personal data has proven to be complex and a key source of tension. It is an emotionally charged debate in which stakeholders have radically different and valid points of view. The extremism stems partly from competing notions of control.

Some individuals believe that if a piece of personal data can relate to them, they should have some control over how it is used. At the other end of the spectrum, many organizations – particularly those in the private sector – consider the data they have captured and created about individuals as theirs. They feel data is an asset reflecting the investment of significant resources, and they expect to maximize its return.

Another way of thinking about it is that while personal data may be about an individual, it is generally created through the interactions of multiple parties. The actors involved, therefore, have valid rights and responsibilities to the data and may require different permissions to exercise those rights. These rights are therefore generally shared rather than exclusive. They are shared because rights arise in a social context and are realized only through the recognition by other parties.

In this light, the widely debated question of "who owns data" frames the issue as a binary "either/or" choice. Ownership is a complicated legal and social construct that does not necessarily grant exclusive rights. Even when an individual or organization is considered to "own" personal data, they most often do not have complete control over it.

An analogous situation exists with personal property. Homeowners cannot do whatever they want with their home. The mortgage holder, the government and local neighbours all have some say in what can and cannot be done with the building. Likewise, musical artists may nominally "own" their music, but other parties including publishers, distributors and record labels share in the rights over how it is used. Multiple parties share in how value is extracted and created at different stages over time.

Additionally, the concept of "ownership" is complicated by the digital nature of personal data. Unlike physical goods, anyone can make unlimited copies of personal data and distribute them globally. They cannot curtail the ability of others to uniquely innovate and create new value from the same piece of data. Intangible digital goods operate under the laws of abundance, not scarcity.

A focus on the construct of ownership, therefore, may limit the potential to move the actors in the ecosystem onto a more stable footing that rebuilds trust. Regardless of who "owns" the data about an individual, many stakeholders – including but not limited to individuals – may have a valid claim in how the data is used and value is extracted and distributed.

Privacy: Linking Identity to Data
It is impossible to read the news without stumbling across the subject of privacy: Google's changes to its online privacy policy, the US government's Privacy Bill of Rights, the United Kingdom's phone-hacking scandal, etc.

The danger in the current debate is that the word privacy can be used to suit nearly any purpose. It is a broad concept charged with a great deal of definitional, legal and moral ambiguity. It is also a construct anchored in the past, when siloed, mainframe computing architectures drove policies and regulations. It hinges on the general premise that those who create and hold data – primarily businesses and organizations – are structurally separate from individuals, the subject of that data. That world has changed.

In general, the confusion and tension surrounding the issue of privacy arise from multiple directions:

- The semantics of privacy: Privacy conveys a variety of overlapping harms, including for example the appropriation of a person's picture or name for commercial advantage, surveillance over individual affairs and public disclosure of private facts.

- Macro approaches to privacy: Jurisdictions, countries and cultures take different approaches to address the identified harms without any coordinated global policy approach.

- Micro approaches to privacy: Individuals display a range of inconsistent behaviours driven by individual choice and economic rationales, often saying one thing and doing another.

In its current state, the dialogue about privacy too often focuses on extremes or fatalism,[38] such as the notion that "privacy is dead". While the intricacies of the privacy debate are beyond the scope of this report, unpacking the concept at a high level remains a key discussion to hold. With a richer understanding of the dynamics shaping the current debate, shared opportunities, risks and means for collaborative responses can emerge.

A contextual view may also avoid the danger warned of by Jeff Jarvis in his book Public Parts: "With all this talk of privacy, privacy, privacy, we might well end up with more protection than ever – perhaps too much and lose opportunities to make connections in this age of links."[39]

[38] Acquisti, Allessandro, http://blog.ericgoldman.org/archives/2011/12/economics_of_pr.htm.

[39] Jarvis, Jeff. *Public Parts*. New York: Simon & Schuster, 2011.

A Perspective for Achieving Balance

With tensions around ownership and privacy in mind, an opportunity exists to focus the dialogue on personal data with a greater level of precision and a richer sense of the diversity of perspectives. The following approach can provide a means for achieving a balanced ecosystem. Three separate but related components can serve to guide the discussion.

1. Protection and Security

"*There are only two types of companies: those that have been hacked and those that will be.*"[40]

A distributed system governs data, with multiple parties involved in storage and management. Many of the companies that individuals think they are providing their data to actually have a host of back-office providers involved in storing and maintaining that data – providing the data centres, mirroring the information and other services. Experts have estimated that someone who gives a company his or her name and address could easily generate in excess of 1,000 copies of that information within 12 months.[41]

An interdependent, adaptive, chaotic system of data management and storage requires a collaborative approach to protection. This approach applies not only to personal data, but also to protecting all data in a hyperconnected world. No single organization can be "cyber resilient" by itself. Still, improvements to data security and protection within one organization can contribute greatly to overall global cyber resilience.[42]

Data encryption and security technologies are fundamental to an ecosystem that protects personal, corporate and governmental repositories of personal data. These technologies enable personal data to be safely stored and exchanged. A number of crucial technical and legal questions relate to the encryption of personal data, including: What level of encryption will be used for various types of personal data? Who will hold the keys to encrypted personal data? What legal protocols are needed for decryption?

Data can flow and create value only with robust technologies in place to deal with three security elements: *anonymization* (removing personally identifiable information that could be used to link data to an identified or identifiable person); *personal aggregation* (combining related personal data sets into larger aggregations of data); and *personal data parameterization* (converting data into a set of parameters that capture valuable insights without including information about a person or their digital identity). As with any global networked information technology, standards and best practices will prove essential to ensuring consistent interoperation and scale.

Along with systemically accounting for the technical side, organizations need to focus on the human element, as human error accounts for a high number of security breaches. Best practices involve limiting access to only those people who have a vital need to the data and limiting subsequent loss through the implementation of technical solutions, such as the effective encryption, compartmentalization or masking of data.

With a large number of unmanageable passwords to remember, many people simply reuse the same codes for all of their accounts, making it that much easier for identity thieves. The US National Strategy for Trusted Identities in Cyberspace (NSTIC) is one such effort to address the issue. It facilitates the creation of a vibrant marketplace that allows people to prove their identity through trusted credentials.[43]

To ensure trust, all organizations that collect, store and access personal data have an obligation to put in place effective and appropriate security in line with the sensitivity of the data and its vulnerability to theft or misuse. The aim is to protect data from being breached, whether that happens through theft, "hactivism" or accidental release.

2. Rights and Responsibilities for Using Data

Tensions can arise between different stakeholders in the way in which perceptions of control over the use of personal data compete with the potential for harm that these uses of data can create. The tension between control and harms come to life when examined through the lens of the different data and its uses.

Key to resolving this tension between control and harms, and avoiding an unmanaged situation where anyone can do anything with data, is the need to establish a more managed system for establishing who has what rights and responsibilities for using data and how to ensure appropriate permissions for usage. There are different views and approaches for the best way of achieving this.

One approach is a rules-based mechanism embedded in legislation that defines explicitly the rights and responsibilities of different stakeholders for all types of data and uses – for example, the newly proposed EC Data Protection Regulation strengthens the rights of individuals by proposing a right to be forgotten and a right of data portability, among others. At the same time, it also establishes responsibilities for organizations holding personal data, requiring them to notify serious data breaches within 24 hours, and the responsibility to ensure that whenever consent is required for data processing that it is given explicitly through an opt-in mechanism rather than being assumed. It also harmonizes data protection rules across the EU, making it easier for international businesses to understand their obligations.

Such an approach aims to give an individual more control over data about them, provides clarity to all stakeholders on their rights and responsibilities and sets out clearly enforceable rules. It aims to provide stability to businesses and individuals on the rules for usage and is anchored in a foundation of fundamental rights, which has worked to protect individuals' privacy.

However, the complicated and ever-changing nature of a hyperconnected world has potentially outstripped the ability of traditional rule-setting legislative approaches to keep pace. This raises a number of questions. Is it possible to develop legislation to establish rights and responsibilities appropriate for all possible current uses and types of personal data? Can a legislative approach be adaptable enough to respond to possible future new uses of data that are not yet even understood? What will the costs of such regulation be and how will it impact innovation?

Another approach to establishing rights and responsibilities for usage of personal data is an adaptable, flexible approach built on context. Context could be the key to striking a balance between

[40] Speech by Robert Mueller, Director, Federal Bureau of Investigation (FBI), 1 March 2012, http://www.fbi.gov/news/speeches/combating-threats-in-the-cyber-world-outsmarting-terrorists-hackers-and-spies.

[41] http://healthprivacy.blogspot.com/2011/03/jeff-jonas-how-many-copies-of-your-data.html

[42] See the World Economic Forum's Partnership for Cyber Resilience for more details on the interdependent nature of cyber security and resilience at http://www.weforum.org/cyber.

[43] http://www.nist.gov/nstic

ensuring a robust establishment of rights and responsibilities and ensuring they are flexible enough to adapt to the changing world. This would mean considering different collection methods, types and uses of data in establishing rights and responsibilities of different stakeholders.

Collection of Personal Data
Personal data can be volunteered, observed or inferred. Volunteered data comes directly from the individual – photos, blogs, tweets, videos, comments, "likes", e-mail messages and so on. Observed data is created as a result of a transaction between an individual and an organization – location data from a mobile phone, credit card transactions, purchase history at a retailer, etc. Inferred data, also called derived data, is the output of data analysis, combination or mining, and it includes credit scores, predictions of preferences and purchase intent. If volunteered data feels like an intimate social gathering, observed data can feel like the paparazzi snapping photos, while inferred data can feel more like an all-knowing Big Brother watching the security camera.

With each step along the spectrum – from volunteered to observed and finally to inferred data – organizations tend to feel an increased sense of ownership and control, particularly as the time, energy and financial resources devoted to creating it increases. There are very few incentives for organizations to share this data either with individual or with competitors. But at the same time, the perceived privacy harms increase as individuals lose a sense of control, with perceived damages growing as data moves from volunteered to observed to inferred data. The more distant data gets from the awareness of an individual and the more intimate and predictive it becomes, the greater the sense of unease and suspicion. This loss of control and sense of intrusion could lead to widespread disaffection and abandonment from the system altogether.

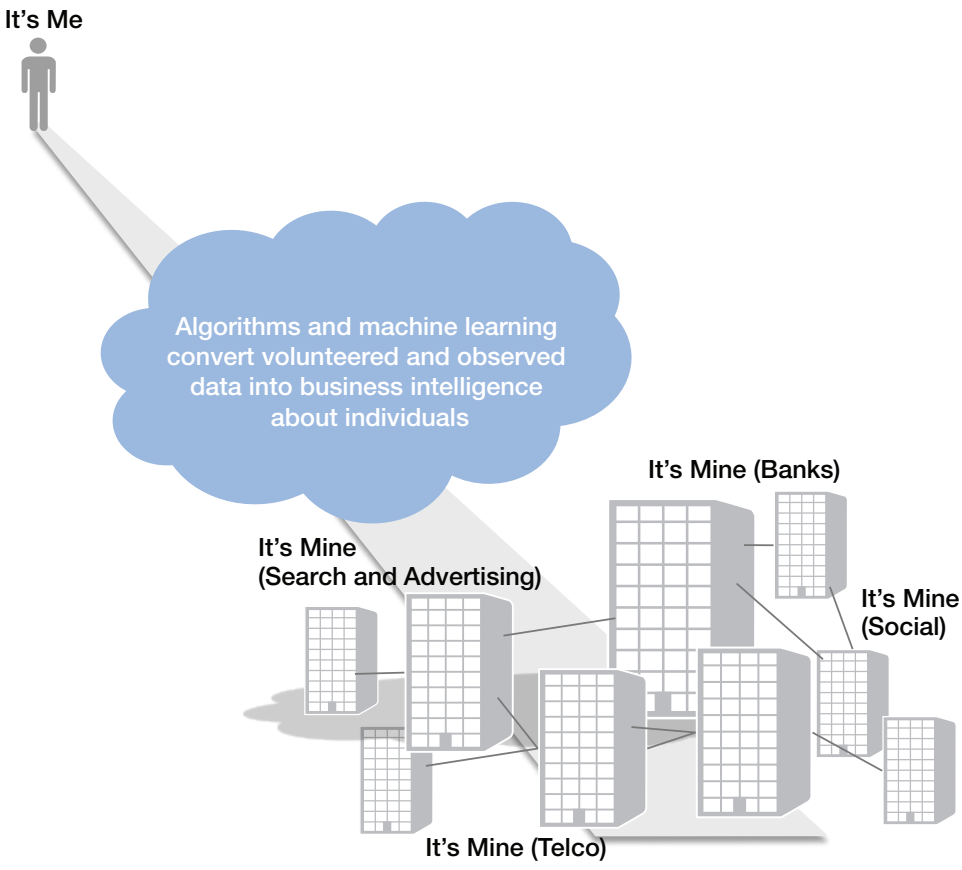
The tension between the individual's perception of harm and their sense of control versus organizations' increasing sense of proprietary ownership and value potential, is a key point of tension in the personal data ecosystem. It leads to an increasing sense of mistrust, lack of accountability and suspicion towards organizations about how they are collecting and analysing data.

Types of Personal Data
In addition to the different ways in which data can be collected, individuals consider some types of data more private than others depending on how closely it relates to them. For example, health and financial data tend to be seen as more sensitive and in some jurisdictions are subject to more stringent restrictions. These perceptions can help gauge how extreme the consumer backlash might be if usage permissions are not set appropriately.

But such perceptions are not universal. In the US, consumers think of personal financial information as highly sensitive, while in Norway and Finland tax and personal income data are published openly for all citizens.[44] And it is certainly not uniform at a country level. People have unique life circumstances that make certain kinds of data more or less sensitive for them.

[44] Jarvis, Jeff. *Public Parts*. New York: Simon & Schuster, 2011.

Exhibit 3: The tension of control over "inferred" personal data



In large part, the sensitivity of data can be linked to its predictive power. For example, consider the current debate around the increasing capture and use of geolocation data through devices like mobile phones. An individual's location patterns, coupled with those of other individuals, can be tapped to better understand and quantify human behaviour at a scale never dreamed of before. As researchers at MIT[45] and elsewhere have shown, when combined with a few other pieces of data, location data can be used to uniquely identify individuals and reveal their behaviour without an individual's knowledge.

Each type of data requires a balancing between the rights and responsibilities of the different stakeholders, and that balance is likely to vary with context. A flexible approach to setting permissions would aim to reflect the types of data; address the sense of ownership, control and harms that depend on how the data is captured; and balance the competing rights and claims of stakeholders. After all, just because an organization has access to data does not mean that it can use it however it wants without appropriate permissions, especially when they use it in ways that are inconsistent with individuals' interests.

Use of Personal Data

In thinking about an approach that is based on context, rights and responsibilities depend not just on the data itself, but also on what it is being used for. For example, mobile phone providers may have the right to track a user's location and call patterns to determine the cell tower that routes their calls and to deliver better service. But other uses are possible. Can a mobile phone provider use the same data to market products that are more suited to customer needs, such as proposing international prepaid packages for individuals based on the countries where they travel? Can it go one step further and sell the information to third parties like hotels? Can it share who you have been sitting next to at the bar, their credit rating in relation to yours and if you are both stumbling as you walk to your car? In some jurisdictions it is not clear whether such uses, though possible, are allowable while in other jurisdictions such usages would not be possible without explicit consent of the individual.

Alternatively, consider an individual's digital health record. A record can be used in many ways – by a doctor to treat a patient, by a hospital in combination with all other patient records to assess the performance of doctors, by researchers to develop new treatments and by insurance companies to ensure payment to health providers. As the usage changes, the rights and responsibilities of the different stakeholders change and perhaps, therefore, the permissions required should change.

Organizations that break the context of how data is used – for instance, using data in ways that diverge from stated intents and unstated expectations – undermine the trust of individuals and create a backlash of suspicion and doubt. Furthermore, signing a long and detailed policy when registering for a service does not change individuals' expectations about how the data will be used. That agreement is unlikely to reduce the perception of a breach of trust when the original context is broken and an unexpected usage takes place.

Many uses of data can be difficult to anticipate at the time of initial collection and some of them may generate significant value for society. For example, an analysis of clinical and cost data from Kaiser Permanente about the adverse effects of the arthritis drug Vioxx led to the withdrawal of a drug linked to 27,000 cardiac arrests. Other examples of secondary uses of data that create

[45] de Montjoye, Yves-Alexandre. Hidalgo, C., Verleysen, M. and Blondel, V. "Spotted in the Crowd: Mobility Data as a Digital Fingerprint", MIT Media Lab Working Paper, 2012.

The Spectrum of Data

At the moment, no real mechanisms allow individuals to share data, but also restrict its use to the context in which it was volunteered, observed or even inferred. The lack of context in the process represents a key point of tension, but also contains a potential solution.

Volunteered data typically involves a strong sense of ownership and rights to individual control – *my* photos, *my* blog posts, *my* tweets, *my* e-mails – regardless of who legally owns the data. Yet individuals have often given their implicit consent for basic levels of use by opting into the services and then giving their explicit consent to certain uses of the data through the terms of service or privacy policies. But these uses often represent a point of tension. Few individuals read such policies. Consent is often a tick-box exercise that individuals race through when they sign up for a service. Some observers argue that such consent for "volunteered" data actually has been coerced or given under duress – provide data or receive no service.

Observed data, on the other hand, shifts the sense of ownership and control to the organization which captured it. Individuals may not even know or fully comprehend how much observed data is being captured about them, or be aware of all its uses. Yet if individuals knew more, they might demand greater control from organizations.

Inferred data, which involves information computationally derived from all the data volunteered and observed, shifts the implicit sense of control even further away from the individual. Organizations generally see the analytics and insights derived from inferred data as a proprietary asset. They have few incentives to openly share these insights. Although the inferred insights might be direct and intimately tied to an individual (or a group of individuals), individuals' sense of direct control and awareness often remains limited. Inferred data has predictive capabilities that have become concentrated in the hands of a few large companies, the "6,000-foot giants" of personal data, to paraphrase physicist Albert-Lázló Barabasi, giving validity to the concerns of social control and surveillance from a few "really, really big brothers".[46] Interventions related to these risks and concerns need to be balanced and thoughtfully understood.

[46] Barabási, Albert-László, *Linked: How Everything Is Connected to Everything Else and What it Means for Business, Science, and Everyday Life*. Cambridge, Massachusetts: Perseus Publishing, 2002.

value for society include electricity smart grids, traffic management and control systems, and retail-store layout optimization.[47] Using personal data for the common good represents a context just as unique as some less desirable uses. The issue remains how to establish mechanisms that allow such good uses while preventing less desirable ones.

Each new level of use represents a point of "data leverage". That is, each use offers an opportunity to pause and reflect about whether any concerns require certain permissions to address stakeholders' needs.

A Balance of Rights and Responsibilities
Given the spate of breaches and misuses of data and the absence of the individual in the discussion, much of the current public and regulatory debate around personal data has centred on protecting the rights of individuals to privacy. One way is to focus the dialogue on balancing the rights and responsibilities of *all* stakeholders and establishing appropriate permissions that allow the positive possibilities while preventing the negative things from happening.

Establishing a sustainable personal data ecosystem will require balancing competing rights or claims of various stakeholders, the potential risks of harm to other stakeholders, and the value created and captured. By its nature, such a balance needs to be contextual.

For example, the fear that digital medical data will be used by employers or insurance companies to discriminate against individuals is a serious and valid concern. However, this needs to be balanced with the value that the data creates for individuals in terms of better treatment, the value for society in terms of better research and cures, and the value for governments and other healthcare providers in terms of reduced costs. Or consider the multiple rights and responsibilities that must be balanced in the case of privacy such as the individual's right to privacy, the press's right to freedom of expression, the government's responsibility to provide national security, and the rights of the public to protection against infectious diseases.

An important element concerns the linkage of data to an individual's identity. Personal data can be thought of like the layers of an onion: some core uses are directly linked an individual, farther from the centre are the layers of use linked to someone as a member of a group or community, and towards the outside are layers of use linked to someone as a member of a sovereign state. The closer the usage is to an individual's identity, the greater the value but also the potential risk of harms to that individual.

Until recently, anonymous data had all personal components of the data deleted or changed and thus it was seen to be unconnected to a person's identity. The many collective benefits that can be created from anonymous or aggregated data explain why traditionally most regulators subject personally identifiable information (PII) to much stricter rules than anonymous or aggregated data. But technology is breaking down the barriers that once enabled anonymity (See sidebar on the Breakdown of Anonymity). The breakdown is calling into question the rights and responsibilities of all stakeholders in using anonymous data and forcing a reconsideration of the level of permissions required for such use.

The Breakdown of Anonymity

Traditionally, organizations have used a variety of techniques to de-identify data and create value for society while protecting an individual's privacy, including medical research into the effectiveness of medicines. Removing the identifiers created a set of anonymous data that was not subject to the same rules as personal data, as an individual could not be identified from it. But technology is breaking down the barrier of anonymity. Someone's identity can be re-established, known as de-anonymization, by cross-referencing a set of related data.

The "triple identifier" of birthday, gender and zip code is all that someone needs to uniquely identify at least 87% of US citizens in publicly available databases. Those three pieces of data can be found through cross-referencing widely available data sets, as Netflix and AOL learned the hard way when "anonymized" data they released for research was used to re-identify individual users.

Given these changes, one approach to ensure adequate protection is to continue to debate the boundaries of personally identifiable data, anonymous, pseudo-anonymous or aggregated data. The boundaries between these categories are likely to continue to shift with improvements in data mining and analytic capabilities.

Some observers have argued that this means the death of anonymous data and that all data is effectively personally identifiable and should be treated as such.[48] The proposed European Commission's Data Protection Regulation expands the definition of personal data to include anything that "directly or indirectly" is "reasonably likely to be used" to identify a person, including an "identification number, location data and online identifier". The regulation, however, maintains that the principles of data protection should not apply to data rendered anonymous in such a way that the data subject is no longer identifiable.

Others argue that the distinction is being used by businesses to create incredibly detailed profiles of individuals, which while anonymous, are so closely tied to the individual that they should be treated as if they are. Other observers urge caution, arguing that this would curtail many of the beneficial uses of anonymous data with minimal gains in privacy.

An alternative approach is to shift to a new paradigm of protection based on context. Rather than focussing only on the data itself, a contextual approach considers different uses of data and how linked the usage is to an individual. Data that is linked to an individual but used at an aggregated level or in an anonymous way is different to uses that specifically identify the individual.

The solution cannot yet be discerned, but without a solution, the value of anonymous data may be lost to all.

[47] Tene, Omer and Jules Polonetsky. "Privacy in the Age of Big Data". *Stanford Law Review*, 2 February 2012. http://www.stanfordlawreview.org/online/privacy-paradox/big-data.

[48] Ohm, Paul. "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization". University of Colorado Law School, 17 August 2009, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450006.

3. Accountability and Enforcement

In a world of systematic data security breakdowns, accountability and enforcement are diffuse at best. Currently, the system has an "oops" accountability and enforcement mechanism – misuse data or allow it to be breached, beg forgiveness from consumers and settle with regulators over the consequences.

Trust is critical to the functioning of any sustainable, networked system. But trust is an ambiguous term open to many interpretations. As many observers have pointed out, trust is impossible without accountability from organizations that collect, secure and use personal data.[49] In addition, accountability requires enforcement mechanisms, if it is to be more than a hollow organizational promise. There is a need to avoid accountability concepts that sound good politically but are not enforceable in practice. The issue then becomes how to build accountability and enforcement mechanisms that address the unique requirements of the hyperconnected world.

In effect, accountability and enforcement mechanisms are needed to ensure two actions:

1. Organizations protect and secure data
2. All stakeholders use personal data in accordance with the rights and established permissions for the trusted flow of data

The specific approaches are likely to be different for these two actions. For instance, accountability for protecting and securing data is less contextual – all organizations, regardless of the types of data they hold and how they plan to use it, have an obligation to protect and secure the data to avoid it being compromised or misused. Stakeholders involved in creating the accountability or enforcement mechanism for stewarding data against breach are likely to be different from those involved in stewarding the context and usage rights to ensure data flows. For example, governments and regulators potentially have a clear and active role in setting the right incentives to encourage all organizations (regardless of context) to take actions that prevent data breaches and misuse, whether through fines or other mechanisms.

On the other hand, organizations potentially can play a much more active role in defining the standards of accountability for using personal data in accordance with rights and responsibilities and established permissions. Enforcement of those standards likely requires a much more dynamic and flexible approach, involving different stakeholders depending on the context.

To make these two actions the norm, a number of possible accountability and enforcement options present themselves, outlined below.

Voluntary principles to guide behaviour. Organizations can commit themselves publicly to act based on a set of principles, even if they lack a formal sanctioning ability. The recently announced Partnership for Cyber Resilience from the World Economic Forum lays out principles to align individual and organizational action to secure data for the collective good. The partnership represents a flexible approach that empowers individual organizations to take responsibility for protecting data against breach without unduly limiting the manner in which they pursue that goal. The Fair Information Practice Principles (FIPPs),[50] although written for the pre-Internet era and therefore seen by many as outdated and flawed,[51] have long been the basis for guidelines and model codes for organizations.

Public reporting of data stewardship. Once standards have been agreed to, governments, companies and non-profits can be audited against them, with public disclosure requirements that require top leadership sign off. Companies in various jurisdictions, for example, already face requirements to review and sign off on audits of financial statements, environmental compliance and even actions taken to prevent potential risks like Y2K. Other industry approaches, some of which are similar to the ISO9000 quality standards, might provide useful examples for how all parties in the data supply chain can be confident that they can share data with others safely and securely – if they can move from a "check the boxes" certification exercise to one that truly leads to better outcomes for consumers.

Financial or criminal penalties. At the other end of the spectrum, governments and regulators can impose penalties for not stewarding data appropriately. The recently announced European Commission Data Protection Regulation, for example, has focused on sanctions for not adequately securing data, and it includes fines of up to €1 million, or 2% of global annual profits, for a company found to be in breach of the requirements.[52] The recently concluded FTC investigations into both Facebook and Google are examples of regulators stepping in to hold organizations to account for stewarding data in accordance with context and usage rights the companies had already established.

Crowdsourced solutions. A major challenge in setting rights and responsibilities will be to create mechanisms that more effectively include the 7 billion individuals worldwide. Norms of behaviour are needed much like the norms that cause most people to pick up a wallet on the street and turn it in to the police station. With the right norms and mechanisms, the world's population will have enough of a stake in the process to serve as a global digital neighbourhood watch, so to speak, to police the use of personal data.

Market mechanisms. As individuals make choices and express preferences for solutions to their data-handling concerns, the market can help achieve accountability. Today's privacy, security and liability problems will fuel the solutions of tomorrow's services and products.

Part of the complexity in establishing accountability and enforcement arises from the lack of understanding that individuals generally have about how data that is collected on them is used and transferred. The Personal Data Ecosystem Consortium has conducted a value mapping exercise that tries to diagram the complicated flow of data (see Appendix I).

At the same time, it is worth comparing the situation with equally complex ecosystems in which individuals do have a high degree of trust and accountability. With credit cards, for instance, cardholder liability is capped in most jurisdictions to a small amount (for example US$ 50 in the US if certain conventions are observed). Individuals are assured that they will not suffer a major loss through fraudulent use of their card. Consumers therefore have far less interest in the intimate workings of the system than they otherwise might. As long as they will not be harmed, they engage. For personal data, energy might be applied towards similar harm-mitigation strategies to encourage accountability and trust.

[49] Hamlin, Kaliya. Notice of Inquiry Response to NSTIC. http://www.identitywoman.net/wp-content/uploads/2011/07/NSTIC-NOI-Kaliya1.pdf.

[50] For more information on the history and different guises of FIPPS, see http://itlaw.wikia.com/wiki/Fair_Information_Practice_Principles.

[51] See, for example "The Failure of Fair Information Practice Principles" by Fred Cate, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1156972.

[52] European Commission of Justice, http://ec.europa.eu/justice/data-protection/index_en.htm.

Chapter 3: Setting Permissions Based on Rights and Responsibilities

Essential to ensuring a trusted flow of personal data is the need to establish rights and responsibilities for the use of data that respects context and balances the interests of all stakeholders. Rights and responsibilities are defined and made real through what are known in the world of personal data as "permissions".

Permissions are different from rights or claims. They are the licence or authority to do something. *My landlord may have the right to enter my apartment, but only after obtaining my permission*. Traditionally permissions in the personal data ecosystem have been static and often obtained at the time of collection of the data. But given the current climate, permissions need to be dynamic and need to vary based on context.

But exactly how should permissions for the use of data be set in such a diverse, decentralized and dynamic ecosystem? No single answer exists to the question. The answers are likely to involve, in equal parts, technical, regulatory, market-based and social mechanisms.

This section highlights the existing ways in which most permissions are set, some of the problems with these approaches and efforts to improve them. In addition, the report explores some emerging alternative ways of establishing permissions to use data, including efforts to allow individuals to play a more active role in setting permissions. Each approach illustrates the complexities of establishing rights and responsibilities for usage. Each also underlines the centrality of a structured dialogue in making progress to answer such an important open question.

Current Approaches to Permissions

At the moment, data permissions are built on traditional models of notice and consent in which an organization seeks permission and notifies a passive data subject how their data will be used. That model is now breaking down as individuals become more active agents and creators of data.

Typically, individuals must wade through long privacy policies that notify them and seek their consent in legal terms. But they leave open the question of whether that consent is "informed" in any real sense. These policies often serve more as liability disclaimers for businesses than as privacy assurances for consumers.[53]

Efforts are under way to find more effective means for organizations to notify and obtain consent from individuals without overwhelming them. Such approaches require flexibility depending on the context, but they also depend on simplicity to ensure they do not undermine the online experience. Clear and conspicuous notice and explicit consent may be required to ensure an individual's trust in certain circumstances. But in other circumstances, those same requirements might restrict the flow of data and the creation of value.

The notice and consent model was modelled on traditional contract frameworks, and it found expression in the Fair Information Practice Principles that arose in the more leisurely data collection context of the early 1970s.[54] Questions linger about whether this model can ever be sufficiently improved so as to effectively balance rights and permissions for usage in a massively networked information environment.

[53] Tene, Omer and Jules Polonetsky. "Privacy in the Age of Big Data". *Stanford Law Review*, 2 February 2012. http://www.stanfordlawreview.org/online/privacy-paradox/big-data.

[54] For example, the Fair Information Practice Principles, Federal Trade Commission, http://www.ftc.gov/reports/privacy3/fairinfo.shtm.

Notice

In essence, this permission concerns the information that is provided to an individual about how others are expected to use the collected data about them. It is intimately linked with the notion of consent, for without effective notice it is not possible to truly achieve consent. The current approach usually involves individual notification about possible uses of data at the time of collection or signing up for a service. But to make effective choices, a prerequisite is that individuals must actually read and understand the documents, which they mostly do not. In addition, many of the possible uses of data are not known at the time of collection, which calls into question the effectiveness of a static notice approach, even if it is clear and easily understood.

More information does not necessarily mean more effective notification. Diminishing returns are clearly now in effect. Studies show privacy policies are hard to read, read infrequently and do not support rational decision-making. In fact, if Americans were to read online privacy policies word for word, the value of time lost is estimated be about US$ 781 billion annually.[55] To address these issues, organizations need to understand the context. Depending on the situation, notice can range from implicit to explicit (though different jurisdictions allow different levels of notification):

Implicit notice: In contexts in which the usage of data is clear and obvious, the notice required may be more implicit. For example, implicit in collecting a customer's address when ordering a book from Amazon is that the information will be used to deliver a package. Sometimes, disclosure of certain facts may not be useful or valuable to the individual, such as a detailed notice of all the steps that credit card companies take to protect against fraud. Other times, organizations may anonymize data to better understand their own customers without sharing it with others. They may feel that the use is already implied in the customer relationship.

Explicit notice: In many cases, explicit notice is required to ensure trust (and meet legal requirements). But, as noted above, the current model of notice is failing to adequately build trust as notices tend to be long and are rarely read. All organizations can ensure that the needed disclosures are as short and simple as possible so that individuals can make informed choices rather than skipping over the document.

Several examples serve as evidence that effective notification can be achieved through simplified approaches. For example, the US Department of Health and Human Services' PHR Model Privacy Notice is a model of effective notice. It guides the release and security policies for personal health records, answering just six simple but central questions.[56]

Some apps on the iPhone and other smartphones offer another example. When someone downloads an app, a highly visible, impossible-to-miss message appears that asks if the person wants the app to either push notices to them or use their location. PrivacyChoice has developed a system to score websites on a scale of 0 to 100, based on how a site collects and uses personal data. This could give users effective notification in an easy to compare way and allow them to make smarter decisions about which sites to trust.[57]

More innovation is needed to ensure that individuals are effectively notified, when needed, of the different potential uses of their data. This will in turn ensure that any required consent is truly informed.

Consent

The nature of consent (opt-in, opt-out, implied consent) for different uses of data is the matter of significant public debate with strongly held and varying views by different stakeholders and policy-makers. For example, the new EC Data Protection Regulation now clearly defines consent as always needing to be explicit and opt-in for all contexts. The US Federal Trade Commission has noted that to simplify choice, companies should not have to seek consent for certain commonly accepted practices. It said it was "reasonable" for companies to engage in certain practices without consumer consent – namely, product and service fulfilment, internal operations such as improving services offered, fraud prevention, legal compliance and first-party marketing.[58]

To strike a balance between all stakeholders and between the needs for individual control and trusted data flow, one approach is to consider that individual consent requirements need to vary based on the context. Most current approaches to consent are constructed as binary decisions. The ecosystem would benefit from a more dynamic approach to consent that allows for more than just an on-or-off switch. Currently, arriving at consent varies from the need for an individual to explicitly opt in to data usage, to opt out, and in certain contexts, to have no right to opt out, for example, certain government requests for data.

[55] McDonald, Aleecia M., Cranor, Lorrie Faith, "The Cost of Reading Privacy Policies". *I/S: A Journal of Law and Policy for the Information Society*, 2008. http://lorrie.cranor.org/pubs/readingPolicyCost-authorDraft.pdf.

[56] PHR Model Privacy Notice, The Office of the National Coordinator for Health Information Technology, US Department of Health and Human Services, http://healthit.hhs.gov/portal/server.pt?open=512&mode=2&objID=3770.

[57] Vega, Tanzina. "A New Tool in Protecting Online Privacy". *The New York Times*, 12 February 2012, http://mediadecoder.blogs.nytimes.com/2012/02/12/a-new-tool-in-protecting-online-privacy.

[58] US Federal Trade Commission, http://www.ftc.gov/opa/2010/12/privacyreport.shtm.

Opt in: Many observers see this form of consent as a solution to the need for individual control, since it gives the individual the right to decide how data about them is used. It also provides a clear policy which can be easily understood by individuals. But a uniform approach that requires individuals to opt in to all uses of data could have several unintended consequences:

- *For individuals.* Requiring opt-in consent for every data use could overwhelm individuals and lead to a lack of consumer focus on the important setting of permissions. The IAB Europe industry coalition set up a fictitious European news site (www.cookiedemosite.eu) to illustrate how the Dutch version of the European E-Privacy Directive would degrade a consumer's Web surfing experience. Intrusive pop-up windows hit the screen, and unless the user gives consent to accept numerous advertising cookies, the site's content vanishes from view.

- *For businesses.* Opt-in models may undermine the delivery of free services online or make advertising less effective. The digital economy depends to a large degree on data rather than on direct payment to make money. Many online business models also rely on individuals being able to try out a service and then to decide on the basis of experience whether they want to continue. Opt-in consent could stifle innovation and consumer choice.

- *For the public good.* Many uses of data that do not affect an individual directly, such as medical research or public health tracking, would potentially be at risk if individuals were required to opt into every use (See sidebar on Rebuilding Shared Commons).

Opt out: To balance the rights and responsibilities of stakeholders, an opt-out approach to consent can often allow individuals control over how they respond to different options, without overwhelming them with requests, and still ensure the public value from shared data.

"Do not track" browser options are offering an increasingly common way to opt out in an aggregate way without undermining the web browsing experience. In February 2012, a coalition of major Internet companies agreed to support a do-not-track button, to be embedded in most Web browsers, that prevents the use of people's Web browsing habits to customize ads.[59] The companies have promised not to use browsing data for employment, credit, healthcare or insurance purposes. The data can still be used for market research, product development and law enforcement.

However, such a non-contextual approach sets the valve for data flow to either the "on" or the "off" position. Additional controls are needed so that the somewhat blunt instrument of a binary "do-not-track" mechanism can evolve into a more nuanced relationship that simultaneously protects privacy and adds value.

Rebuilding a Shared Commons

Many uses of data create value for society as a whole rather than a particular individual. Nearly every recent public policy debate has benefited from the mass dissemination of anonymous data about individuals – in other words, from a so-called "shared data commons". This shared data commons contains all the diffuse collections of information made broadly available for research, including tax returns, medical records, government censuses and rewards cards. But as technology breaks down the barrier of anonymity, some observers have called for requiring individuals to always consent to the sharing of such anonymized data.

However, such an approach could lead to "the tragedy of data commons".[60] The tragedy comes from the fact that, if allowed, individuals have an incentive to remove their data from the commons to avoid the risk of their identity being re-established from anonymous data, no matter how small the risk. These individuals receive the indirect benefits of health and policy research, but the collective benefits disappear when data subjects opt out of use.

Technical solutions could be one part of addressing this challenge.[61] But another solution rests on the balancing of the rights and responsibilities of different stakeholders discussed in Chapter 2. Just because a researcher, company or government has the right to use anonymized data for a given use (e.g. health research) does not give them the right to re-identify individuals from the data just because they can do so. With rights come responsibilities.

As author Jeff Jarvis writes, just because someone can use fertilizer for bomb making does not mean the world should stop selling fertilizer. Instead, governments make bomb-making illegal.[62] So it goes with anonymous data – just because re-identification of an individual can be done does not mean it should be.

Society may determine that individual control to prevent any possible breach of privacy trumps the value of the shared data commons, as some jurisdictions seem to be embracing. But that decision needs to be a matter of open debate and conscious choice. Web pioneer Tim Berners-Lee has said that after eliminating data that holds personal information, an untold wealth of knowledge still exists to be found in what remains. The world should not lose the opportunity it affords. Mining that data may set off the next gold rush.[63]

[59] Angwin, Julia Angwin. "Web Firms to Adopt 'No Track' Button". *The Wall Street Journal*, 23 February 2012. http://online.wsj.com/article/SB10001424052970203960804577239774264364692.html.

[60] Yakowitz, Jane. "Tragedy of the Data Commons". *Harvard Journal of Law and Technology*, Vol. 25, 2011. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1789749.

[61] For some technical approaches, visit: http://blogs.bluekai.com/2011/10/removing-the-you-in-online-targeting.

[62] Jarvis, Jeff. *Public Parts*. New York: Simon & Schuster, 2011.

[63] Jarvis, Jeff. *Public Parts*. New York: Simon & Schuster, 2011.

Disposal of Data

Disposing of data immediately after it is used for its original purpose might make sense in some contexts. But in others, such as in the treatment of patients, disposal might seriously undermine opportunities for research that extends the lives of future patients. As with other rules, setting permissions for data deletion needs to balance the rights of all parties to the data, including the individual. For example, the rules for disposal are likely to differ for volunteered, observed and inferred data, given the different rights and responsibilities of stakeholders to such data.

In the current environment, increasing pressure has come for individuals to have a right to delete data about them, as exemplified in the revised EC Data Protection Regulation enshrining the "right to be forgotten." This movement reflects several breakdowns: the lack of trust individuals have about how both companies and governments might use their data; the failure of the current notice and consent model to give the individual effective control mechanisms for how data can and cannot be used; and the lack of accountability by organizations for securing data that has resulted in numerous breaches. Under such circumstances, it is only natural that many people see data deletion as the only tool an individual has to combat misuse.

As yet, it is unclear the feasibility of effectively deleting data in a world in which data can be copied, transferred and moved instantly. In addition, if data really is a highly valuable asset in the information age, then deleting it would suggest a reduction in value for all. Establishing accountability and enforcement mechanisms for both securing data and mechanisms to provide individuals trust that data will not be misused could potentially deal more appropriately with the underlying issue.

Alternative Ways of Setting Permissions

Future solutions to the issue of setting permissions are likely to come from a combination of the technical, policy, legal and business worlds. Turning to technical solutions, advancements in the field will likely depend on some form of digital rights management (DRM). The approach would set clear rules at the point of creation or collection about how data could be used; the rules would "travel with" the data wherever it goes. Such a system has appeal because it features a built-in accountability mechanism – the technical rules would help ensure the data could only be used in specified ways. However, it also potentially limits the use of data beyond the originally specified purpose. These other uses could create value if permissions balance the interests of all stakeholders.

Regarding policy solutions, some options currently being discussed include allowing individuals to set permission levels based on classes of contexts, similar to the way anti-virus software has settings that range from restrictive to open. These options allow individuals to make decisions depending on the context without having to make an individual decision for each website they visit or every transaction they conduct.

An example of a relatively simple legal approach includes Creative Commons, an organization which has created a scalable platform that encourages sharing of creative works for others to build upon legally. It has created an easy-to-understand, one-page explanation, with associated visual symbols, so that people can decide which rights they reserve and which rights they waive. A similarly simple system for notice and consent, or other future structures of rights communication and negotiation, might allow individuals to set permissions for usage in an agile, context-specific way.[64]

A Copy of Data

As a first step in helping individuals exercise active control over data, some observers have pushed for individuals to receive a copy of all the data an organization holds about them (See sidebar "A Copy of Data about You"). The policy allows an individual to understand what data is collected and held about them and to then use that copy of data as they wish. In some ways, this could be seen as an extreme form of notice. While it does not establish mechanisms to set permissions, it does allow individuals to exert limited rights over how a copy of personal data about them gets used.

Efforts could help enable consumers to become "curators of their own data assets", according to UK analyst group Ctrl-Shift, which is working with the government of the United Kingdom on its "midata" initiative that aims to provide individuals a copy of data about them. Individuals could assemble all the data about themselves from multiple sources – banks, retailers, mobile phone providers, energy companies, governments – into one trusted repository that people control and use in ways that both offer new benefits and allow them to compare their activities with others.[65] Such data copies could help establish rights among different parties as a stepping stone toward creating true shared rights and responsibilities, while continuing to allow organizations that hold data to use it in a permissioned way.

Individuals in Control

Finally, the emergence of "personal data lockers" – among many such terms for the services – represent a market mechanism to set the rules around permissions for collection, usage and sharing of data in different contexts and for different types and sensitivities of data. At their simplest, personal data lockers allow an individual to securely store and aggregate data about them from multiple sources and to set permissions for others to access the information in a controlled way.

In addition to lockers, the services have been called many things, including stores, vaults and wallets, as well as vendor relationships, personal identities and personal information management services.[66] Regardless of the names, these digital services have one thing in common – they aim to give individuals greater control over how data about them is used.

The vision is a new model of data management. Rather than organizations collecting, storing, analysing and targeting individuals based on personal data, individuals would be better able to manage and control how the information is used and shared, and they in turn could share more directly in the data's value through such benefits as product discounts and personalized offers.

A growing number of start-ups now believe they can profit from these services, including Personal.com, Singly, Connect.Me, DropBox and Reputation.com in the US; Mydex in the United

[64] http://www.w3.org/2005/Security/usability-ws/papers/26-idcommons/

[65] "The New Personal Data Landscape". Ctrl-Shift, 22 November 2011. http://ctrl-shift.co.uk/about_us/news/2011/11/22/the-new-personal-data-landscape.

[66] Project VRM, http://cyber.law.harvard.edu/projectvrm/Main_Page; Personal Data Ecosystem Consortium, http://personaldataecosystem.org; Fatemeh Khatibloo. "Are You Ready for a World of Consumer-Managed Data?" *Forbes*, http://www.forbes.com/sites/forrester/2011/10/03/are-you-ready-for-a-world-of-consumer-managed-data; "The New Personal Data Landscape". Ctrl-Shift, http://ctrl-shift.co.uk/about_us/news/2011/11/22/the-new-personal-data-landscape.

Kingdom; Qiy in the Netherlands; TrustFabric in South Africa; and MyInfoSafe in New Zealand.[67] Big corporations are also looking to expand into this space. In addition, services such as Enliken allow individuals to "donate" their online clickstreams and associated revenue to the charity of their choice. These actors act as a trusted data agent for an individual's data, aiming to create leverage from data much like a bank creates leverage from deposits.

Data lockers have corollaries in the healthcare and financial industries. Microsoft HealthVault allows individuals to collect, store and share information about their health, while financial aggregation services such as Mint.com allow people to gather and share their financial information in exchange for money management advice and targeted offers. While such lockers are still in their infancy, many groups are trying to shape the structure, trust frameworks and networks that would facilitate such a transformation of the personal data ecosystem, including the Personal Data Ecosystem Consortium, Connect Me, ID3 and many others.[68]

A number of challenges exist for the development of such an ecosystem at scale. It can be challenging to get individuals to take proactive control of how their data is managed, rather than relying on the default settings. And the effort required to collect all their data in such a locker can be prohibitive, although this can in large part be overcome by getting a copy of all the data other organizations already have about an individual. Also, the services do not address permissions for all the additional data about an individual that others in the ecosystem control.

At such an early stage, it remains to be seen whether such a personally controlled data ecosystem will fundamentally transform the current data management model in marketing and advertising, or whether enough individuals will actively set data permissions for marketing purposes.

[67] For a full list, see Project VRM, http://cyber.law.harvard.edu/projectvrm/Main_Page.

[68] For more information, please see http://personaldataecosystem.org/2011/11/3589, which lists most of the organizations working in this space.

A Copy of Data about You

The debate about whether individuals have a right to get a copy of data organizations hold about them received a jolt during 2011 when 24-year-old Austrian law student Max Schrems requested and received a copy of all the data Facebook held about him. The resulting package totalled 1,222 pages and included everyone he had ever "friended" and "de-friended", every event he had been invited to and every private message he had received. It even included deleted data. This single request has grown into a grassroots movement, Europe versus Facebook, which so far has resulted in 40,000 people demanding a summary of all the personal data Facebook holds about them.[69]

Other initiatives have taken a more affirmative approach. The United Kingdom's "midata" initiative encourages organizations holding large amounts of customer data to release the data back to customers in a form that they can use for their own purposes.[70] So far, 26 major organizations have been working with the UK government to make this aim a reality. For example, the information could help customers get the best deal on their current mobile phone contract using their historical calling patterns from their mobile provider.

One of the first initiatives to give individuals a copy of their data was the Blue Button initiative.[71] Launched in August 2010 by the US Department of Veterans Affairs, the effort gives veterans the ability to download a copy of their health data. The Medicare programme and Department of Defense have since adopted it for its beneficiaries and the US government is expanding it to other sectors such as education.[72] The system has become so popular that insurance companies and retail outlets have also pledged to adopt a download capability for their customers.

On the regulatory front, the recently announced European Commission Data Protection Regulation proposes individuals should have "a right to obtain a copy of the stored data from the controller and the freedom to move it from one service provider to another, without hindrance".[73]

Initiatives and rules that aim to give individuals a copy of data about them have the potential to increase transparency and build trust. They also have the potential to allow individuals to combine the data with information from other sources and set permissions about how others can use data.

However, such initiatives are not without concerns. For example, creating many insecure copies of personally sensitive data raises potential security issues and presents an attractive target for theft. Questions have also been raised about the incentives that organizations would have to invest in the technology required to collect, combine and analyse personal data if they are then required to hand over this data to individuals, who might share it with competitors.

In addition, there is a question about the scope of such efforts – do individuals really want all the detailed operational data companies capture to deliver services to their customers? And how much would the cost be for companies to compile and collate this data across many legacy systems?

[69] Europe Versus Facebook website, http://europe-v-facebook.org/EN/en.html; O'Brien, Kevin K.. "Austrian Law Student Faces Down Facebook". *The New York Times*, 5 February 2012, http://www.nytimes.com/2012/02/06/technology/06iht-rawdata06.html?pagewanted=all.

[70] "The midata vision of consumer empowerment". UK Department for Business Innovation and Skills. 3 November 2011, http://www.bis.gov.uk/news/topstories/2011/Nov/midata.

[71] Blue Button, http://bluebuttondata.org/about.php.

[72] "Unlocking the Power of Education Data for All Americans". Office of Science and Technology Policy – White House, 19 January 2012, http://www.whitehouse.gov/sites/default/files/microsites/ostp/ed_data_commitments_1-19-12.pdf.

[73] "Safeguarding Privacy in a Connected World: A European Data Protection Framework for the 21st Century". European Commission, 25 January 2012, http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_9_en.pdf.

Chapter 4: Recommendations for Action

The personal data ecosystem is at a unique moment of time. Personal data represents an emerging asset class that has the potential to be every bit as valuable as gold or oil. Yet the lack of rules and tools to guide the trusted flow of data has created significant fears and controversies about security, privacy and misuse of data.

Stakeholders face enormous pressure to "do something" to fix these problems. At the same time, solutions need to work for *all* stakeholders in a coherent and comprehensive global way. They also need to reflect the unique characteristics of personal data as an asset class that increases in value with use, that can be copied infinitely and distributed globally, and that intimately affects 7 billion individual agents in the personal data ecosystem.

Given the magnitude of the opportunity and task, leaders from across government, industry and society need to work together to achieve an inclusive approach in the following ways:

1. Engage in a structured, robust dialogue to restore trust in the personal data ecosystem

Dialogue must take place among all the stakeholders, including the individual, in order to resolve a host of underlying tensions. At issue: How can the different stakeholders restore trust and create value?

In terms of securing personal data, this dialogue needs to be linked closely with ongoing efforts to address cyber security. The interdependent nature of data management and storage requires a collaborative approach to protection – the system is only as strong as its weakest link. The dialogue, therefore, needs to focus on establishing a joint commitment to global cyber resilience to guide individual and organizational actions that improve security and protection.[74]

In terms of setting permissions, the dialogue needs to reflect the fact that the flow of data across traditional boundaries creates value. This dialogue also needs to unpack several of the key points of tension currently affecting how different stakeholders approach permissions, including ownership and privacy. It also needs to address the potential misalignment of incentives between different stakeholders in the collection and use of personal data. But the dialogue also needs to be decentralized, reflecting the hundreds of separate, use-specific contexts involving stakeholders with different cultural norms, as well as different time frames for action and different paths to a potential solution.

If rights and responsibilities depend on context, and given the speed of change and global nature of data flows, it is difficult for any given regulator to set the rules for every possible usage of data. One challenge will be to determine how to involve the individual in this dialogue. It will be essential to tap into citizen deliberation methods that can effectively catalyse input and engagement.[75] The marketplace, if appropriately structured, can also be the home for such a 7-billion person dialogue.

By contrast, the dialogue around ensuring accountability and enforcement for securing data is likely to be largely context independent. This requires a multi-industry, global discussion to ensure all organizations, regardless of sector or intended use, to create trust so that data will be safe from breach and misuse. A clear need exists for all stakeholders to use data in accordance with the agreed rights and responsibilities for the trusted flow of data. In establishing such mechanisms, government has a clear role as the regulator to ensure that these rules have teeth, to provide incentives for stakeholders to act properly, and to ensure that enforcement mechanisms are as consistent and interoperable as possible across jurisdictions.

[74] See World Economic Forum's Partnership for Cyber Resilience for more details on the interdependent nature of cyber security and resilience at http://www.weforum.org/cyber.

[75] For some suggestions about how this might be feasible, see http://www.identitywoman.net/insight-for-governance.

2. Develop and agree to principles that encourage the trusted flow of personal data

A common set of voluntary principles for leadership can guide the actions of individuals, organizations and governments. These principles should focus on the goal of encouraging the trusted flow of personal data to create value for all stakeholders.

These principles are necessarily distinct from the many industry or geographically-specific guidelines that currently occupy the debate. These principles should be global in scope, but also applicable across sectors and focused beyond merely minimizing data collection, storage and usage of data to protect privacy. The principles need to be built on the understanding that to create value, data needs to move, and for data to move, it requires the trust of all stakeholders. These principles need to address the key issues of trust, transparency, control and value.

These principles can be designed for all organizations, regardless of industry, sector, jurisdiction, geography or level of current personal data usage. They need to cover the protection and security of data against breach and misuse, as well as the setting of permissions based on rights and responsibilities and context. Through coordination around common principles, everyone can aim at the same outcomes even if they take differing approaches to get there, including government regulations, self-regulation, technical solutions and market mechanisms.

Voluntary principles can play a big role in the creation of enforcement mechanisms for networked information systems. No one mechanism can be applied at the exclusion of others, but principles frequently form the basis of private sector-led "self-regulatory" systems capable of crossing the multiple jurisdictions involved. The pathway that starts with voluntary principles frequently grows into processes to implement those principles in specific contexts and then into a virtuous cycle of rules improvement, refinement and enforcement.

These principles can be supported by guidelines that organizations can refer to when working out implementation in differing contexts. These guidelines need to reflect that specific standards, processes and legal requirements will vary over time and by industry, situation and jurisdiction. Such mechanisms sit at the centre of every commercial commodities and securities market, and they are used by the American Bar Association's self-regulation of lawyers, the American Medical Association's self-regulation of doctors and in a host of other settings.

The principles themselves need to be built through the collective dialogue discussed in the first recommendation; no one actor can enunciate the principles alone. However, based on the discussion in this report, a number of areas need to be present in any set of principles guiding the development of the personal data ecosystem. Key issue areas to cover could therefore include:

- Accountability: Organizations need to be held accountable for appropriate security mechanisms designed to prevent theft and unauthorized access of personal data, as well as for using data in a way that is consistent with agreed upon rules and permissions. They need to have the benefit of "safe harbour" treatment and insulation from open-ended liability, when they can demonstrate compliance with objectively testable rules that hold them to account.

- Enforcement: Mechanisms need to be established to ensure organizations are held accountable for these obligations through a combination of incentives, and where appropriate financial and other penalties, in addition to legislative, regulatory, judicial or other enforcement mechanisms.

- Data permissions: Permissions for usage need to be flexible and dynamic to reflect the necessary context and to enable value-creating uses while weeding out harmful uses. Permissions also need to reflect that many stakeholders – including but not limited to individuals – have certain rights to use data.

- Balanced stakeholder roles: Principles need to reflect the importance of rights and responsibilities for the usage of personal data and strike a balance between the different stakeholders – the individual, the organization and society. They also need to reflect the changing role of the individual from a passive data subject to an active stakeholder and creator of data. One perspective that is gathering momentum, though is far from universally accepted, is that a new balance needs to be struck that features the individual at the centre of the flow of personal data with other stakeholders adapting to positions of interacting with people in a much more consensual, fulfilling way.

- Anonymity and identity: The principles need to reflect the importance of individuals being able to engage in activities online in an anonymous way while at the same time establishing mechanisms for individuals to effectively authenticate their identity in different contexts so as to facilitate trust and commerce online.

- Shared data commons: The principles should reflect and preserve the value to society from the sharing and analysis of anonymized data sets as a collective resource.

3. Establish new models of governance for collective action

All the stakeholders in the ecosystem face a challenge of unprecedented size, speed and complexity. Rules and norms change faster in a hyperconnected world than traditional rule-setting approaches can keep up with. Solutions that focus on specific contexts will grow outdated rapidly as those contexts change. A fundamentally new approach to governance will be required, one that can create rules that are both robust enough to be enforceable and flexible enough to accommodate contextual differences.

As it evolves, the hyperconnected personal data ecosystem suggests different roles for stakeholders to build accountability systems and establish norms and rules for the flow of data within these new and evolving governance models:

Policy-makers and regulators should avoid the pressure to put forward one-size-fits-all solutions that may unintentionally lock down the flow of data rather than encourage the trusted sharing and use of data to create value. They need to work with businesses, civil society and individuals to ensure mechanisms are established for the trusted flow of data. In that way, they should guide the development of mechanisms that encourage the positive uses of personal data while preventing the bad, thereby increasing trust among all stakeholders. In addition, they also need to focus on building the conditions that can best foster trust and innovation in the personal data ecosystem. They need to focus on defining accountability systems with real teeth that hold organizations responsible for appropriately securing personal data to prevent breach and misuse, as well as for ensuring all stakeholders use data in ways that are consistent with agreed upon rights and responsibilities.

Organizations, whether they are commercial businesses, not-for-profits or governments, need to take responsibility at the most senior leadership level. The focus needs to be on increasing accountability for stewarding data against breach and misuse, as well as for stewarding the use of personal data in accordance with agreed upon rules and permissions. They also need to encourage other organizations in their supply chains to similarly ensure that data can flow in a trusted way under the guidance of globally agreed principles. Organizations need to embrace the notion of rights and responsibilities for the trusted flow of data. They also need to improve their own engagement with individuals – communicating more effectively about how data is being used, seeking consent in truly informed ways, and innovating to create new business models that create permissioned flows of personal data.

Individuals need to become more empowered actors in the ecosystem and take more control over how data about them is used. They can also reflect on the collective value that is being generated from their personal data as they make decisions about how much to share. They need to engage more proactively with organizations that hold data about them and demand more accountability about how data is used.

4. Establish living labs

Given the speed of change and the decentralized nature of the personal data ecosystem, stakeholders would benefit from mechanisms for sharing and coordinating findings from the many real-world pilots aiming to strike a balance between protecting individual privacy and unlocking value through the trusted flow of personal data.

The contextual approach to establishing rights and responsibilities will not be developed in one step. The only way stakeholders will establish effective approaches is through trials that figure out what works and what does not in the real world. For instance, a need exists for experts to come together to help eliminate existing technical choke points as they scale personal data solutions, such as establishing sample terms of service for commercial entities that could be collectively adopted.

In addition, learning labs can connect relevant industry practitioners and regulators, deploy concepts at national scale, and create collaborative real-time feedback loops that reflect technological change and the real life behaviour of different stakeholders. For example, Weqaya in the United Arab Emirates is a national programme aimed at improving the health of the country's 2.3 million citizens. The system universally captures data about virtually all clinical encounters, as well as behavioural data such as eating and exercise habits through at-home and mobile devices. The national system is able to drive a range of interventions and help individuals understand the actions they can take to improve their health.

Using and combining health-related data in this way offers huge potential benefits, but also raises significant questions. Should such a system be opt in or opt out, and how would the choice affect the value that can be created? Can data be used at an aggregate level for research? How do individuals react to different situations involving data capture and use?

Learning from such national-scale pilots can reveal how to strike a balance in the personal data ecosystem and integrate data into concrete policies. They can help create a more flexible and adaptable governance model than current top-down systems allow. They also provide evidence for the unintended consequences of well-meaning policy actions.

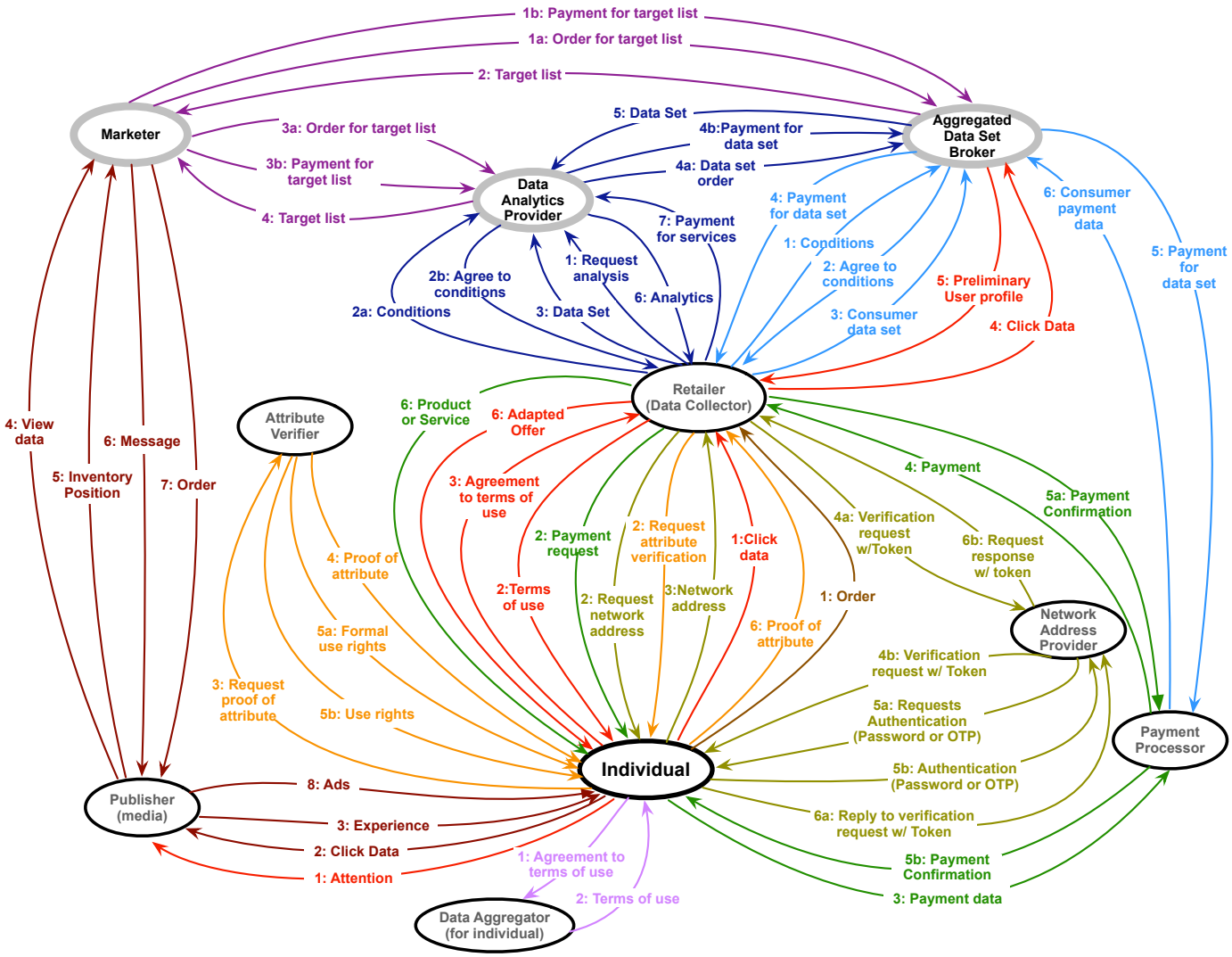
Creating mechanisms to encourage other pilots in a safe harbour learning environment would help understand the complexities of individual behaviour and therefore build a more adaptable, flexible and responsive governance model. Collectively, all the actors can share what they have learned from experiments across countries and industries as they explore innovative ways of setting and sharing rights and responsibilities for the use of personal data that unlocks value across the landscape.

Appendix 1

Flow of Data in the Current Ecosystem around Targeted Advertising and Data Aggregation

A Value Network Map defines the roles in a business ecosystem and maps the value flows – both implicit and explicit – between them. The Personal Data Ecosystem Consortium developed the below map in meetings with industry innovators during 2011. It shows the value flows (in a simplified form) in the current ecosystem around targeted advertising and data aggregation.*

The individual is at the centre and a next ring of parties has direct engagement with the individual (publisher, attribute verifier, data collector, network address provider, payment processors) while a wider ring does not (marketers, data analytics providers, data set brokers).



Source: Created by Verna Allee of Value Network LLC, in collaboration with the Personal Data Ecosystem Consortium

The individual clicks to view content on a publisher's website. In doing so, data about the view and information about the ad inventory available in that viewing act is available to a marketer who places a message (ad) on the page, which is viewed by the individual.

Persuaded by the ad, the individual goes to a retail store advertised. In doing so, they become subject to the terms of use put forward by the site, and the data from the click is sent off to an aggregate data set broker. Using the click data (type of phone, IP address, etc.), the data broker shares with the retailer a profile of the user. The retailer then uses this information to adapt the offer it makes to the user.

The individual likes the offer and places an order for an item that requires a proof of an attribute, such as being a student in a school, or being over the age of 19. The individual must go to an attribute verifier to get proof of attribute (and they are also given explicit or implicit use rights). They then share this proof of attribute with the retailer. Alternatively, the user might choose to ask the attribute verifier to tell the retailer directly about their status, and, in a non-user centric version of this, the retailer could ask for attribute(s) from the provider without the individual's consent or awareness.

The retailer also requires a valid network end-point: a phone number or e-mail address by which to contact the individual. They request this from the individual and send a message to the address asking for confirmation.

Now the retailer is ready to charge the individual using a payment processor. Once the transaction goes through, the individual receives the item they ordered.

The data trails from the transaction continue – the retailer has a data set of transactions, which it sells to a data broker. The payment processor also has data, and it too sells it to the data broker. Note that the data broker has no direct relationship to the individual whose specific data is contained within the data set and shared with the broker.

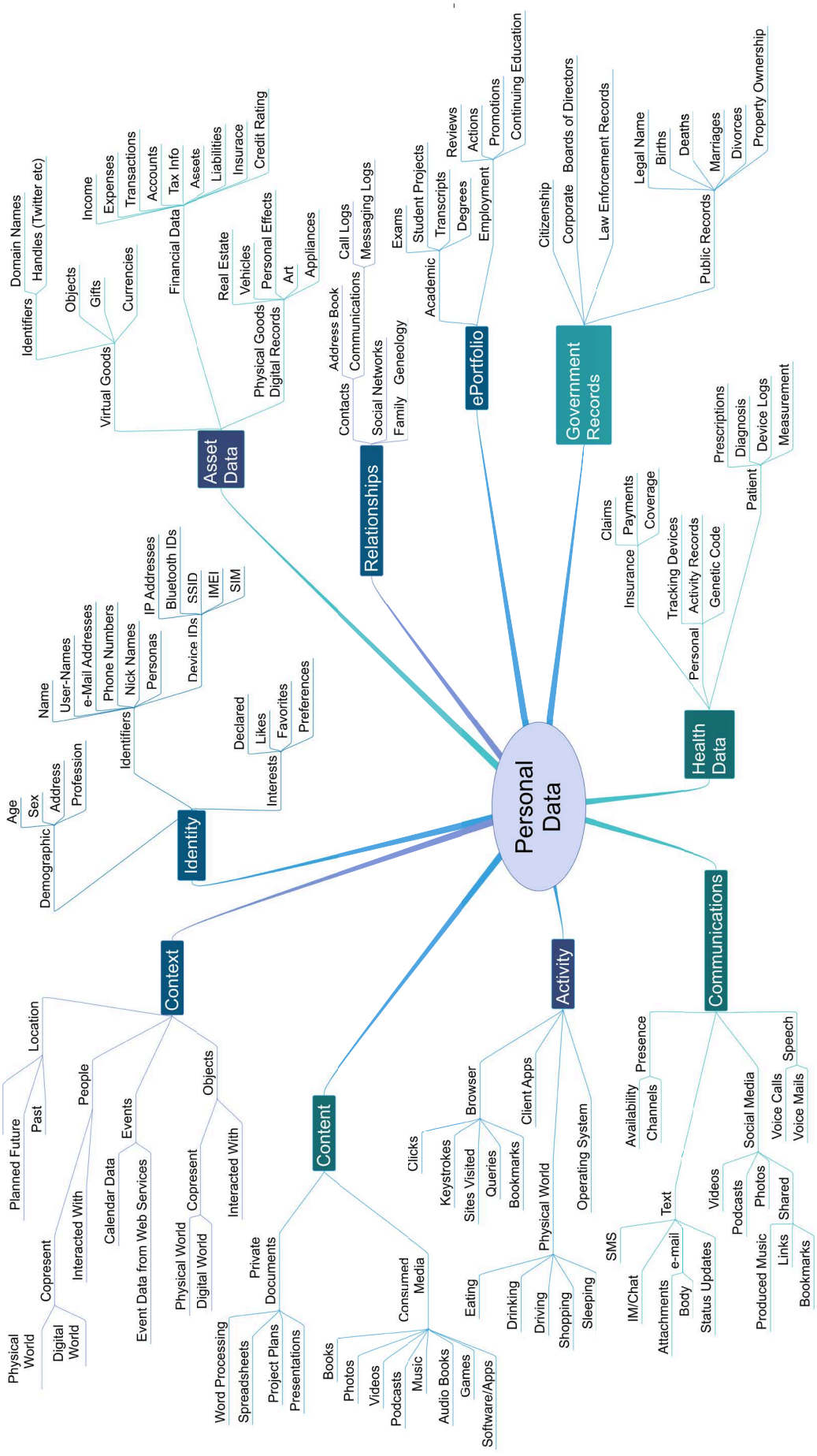
The retailer wants an analysis of their data and goes to a company that can provide that specific service. The data analytics company uses data from the data broker to help in its analysis.

The marketer who is trying to find more information about who to target taps the services of the data broker and aggregator. The whole map goes full circle when this data is used to shape the ads the individual sees when he or she goes to the publisher's website.

* This builds on a submission to the FTC: http://www.ftc.gov/bcp/workshops/privacyroundtables/ personalDataEcosystem.pdf, and the graphic representation of the display and advertising industry by LUMAPartners: http://www.lumapartners.com/resource-center/lumascapes-2

Appendix 2

Personal Data Covers All Aspects of Our Lives



Source: Kaliya "Identity Woman" Hamlin and Personal Data Ecosystem Consortium derived from a list of personal data types appearing in: Davis, Marc, Ron Martinez and Chris Kalaboukis. "Rethinking Personal Information – Workshop Pre-read"" Invention Arts and World Economic Forum, June 2010.

Acknowledgements

Consistent with the World Economic Forum's mission of applying a multistakeholder approach to address issues of global impact, the creation of this report involved extensive outreach and dialogue with the private sector, academic community, governments and multilateral institutions. This dialogue included numerous interviews and interactive sessions to discuss the insights and opportunities for collaborative action.

Sincere thanks are extended to the industry experts who contributed their unique insights to this report. In particular, Scott David of K&L Gates played an invaluable role as expert and patient mentor. We would also like to thank the World Economic Forum's communities of Industry Partners, Global Agenda Council Members and academic experts.

We are also very grateful for the generous commitment and support of The Boston Consulting Group in their capacity as project adviser.

At the World Economic Forum, the energy and commitment of the report's editor and project lead, William Hoffman, is to be acknowledged as are Alan Marcus, Derek O'Halloran, Jessica Lewis and Mark Schulman. Special thanks are also extended to John Rose and Carl Kalapesi of The Boston Consulting Group whose insights were invaluable. Carl Kalapesi served as the primary author of this report and his tireless efforts are to be commended. We would also like to thank Mickey Butts for his assistance in the writing of this report.

Additional thanks are extended to:

Liz Brandt, Ctrl-shift
John Clippinger, ID^3/MIT Media Lab
Scott David, K&L Gates
Marc Davis, Microsoft Corporation
Stephen Deadman, Vodafone Group Services
Timothy Edgar, Office of the Director of National Intelligence of the United States
Jamie Ferguson, Kaiser Permanente
Michael Fertik, Reputation.com
Josh Galper, Personal
Peter Graham, Verizon
Shane Green, Personal
Kaliya Hamlin, Personal Data Ecosystem Consortium
Oliver Harrison, Government of Abu Dhabi
William Heath, Mydex
Leszek Izdebski, Cisco
Robert Kirkpatrick, UN Global Pulse
Alan Mitchell, Ctrl-shift
Mita Mitra, BT Group
Kay Nichols, FIS
Chris Osika, Cisco
Sandy Pentland, ID^3/MIT Media Lab
Robert Quinn, AT&T
Mark Read, WPP
Drummond Reed, Connect.Me
David Sable, Young & Rubicam
Alin Stanescu, Qualcomm
Don Thibeau, Open Identity Exchange
Simon Torrance, STL Partners
Owen Tripp, Reputation.com

WORLD ECONOMIC FORUM

COMMITTED TO IMPROVING THE STATE OF THE WORLD

The World Economic Forum is an independent international organization committed to improving the state of the world by engaging business, political, academic and other leaders of society to shape global, regional and industry agendas.

Incorporated as a not-for-profit foundation in 1971 and headquartered in Geneva, Switzerland, the Forum is tied to no political, partisan or national interests.

World Economic Forum
91–93 route de la Capite
CH-1223 Cologny/Geneva
Switzerland

Tel.: +41 (0) 22 869 1212
Fax: +41 (0) 22 786 2744

contact@weforum.org
www.weforum.org