WORLD ECONOMIC FORUM

COMMITTED TO IMPROVING THE STATE OF THE WORLD

Industry Agenda

Unlocking the Value of Personal Data: From Collection to Usage

Prepared in collaboration with The Boston Consulting Group

February 2013

> I'm convinced this longing to balance, this urge to climb is ingrained in all of us.

—Philippe Petit, High Wire Aerialist

Contents

Executive Summary

Our world is changing. It is complex, hyperconnected, and increasingly driven by insights derived from big data.[1] And the rate of change shows no sign of slowing. Nor does the volume of data show any sign of shrinking. But, the economic and social value of big data does not come just from its quantity. It also comes from its quality – the ways in which individual bits of data can be interconnected to reveal new insights with the potential to transform business and society. Fully tapping that potential holds much promise, and much risk. By themselves, technology and data are neutral. It is their use that can both generate great value and create significant harm, sometimes simultaneously. This requires a rethink of traditional approaches to data governance, particularly a shift from focusing away from trying to control the data itself to focusing on the uses of data. It is up to the individuals and institutions of various societies to govern and decide how to unlock the value – both economic and social – and ensure suitable protections.

As part of the multiyear initiative Rethinking Personal Data, the World Economic Forum hosted an ongoing multistakeholder dialogue on personal data throughout 2012 (See Figure 1 for more details). This dialogue invited perspectives from the US, Europe, Asia, and the Middle East and involved representatives of various social, commercial, governmental and technical sectors, who shared their views on the changes occurring within the personal data ecosystem and how these changes affect the collective ability to uphold core principles. The dialogue also addressed key regional legislative and policy approaches, particularly the proposed European Commission Data Protection Regulation[2] and the US Consumer Privacy Bill of Rights.[3]

The global dialogue centred on a set of foundational principles that are familiar across a broad range of cultures and jurisdictions.

[1] Big data is a collection of data sets so large and complex that they become difficult to process using available database management tools or traditional data-processing applications.
[2] http://ec.europa.eu/justice/data-protection/index_en.htm
[3] http://www.whitehouse.gov/sites/default/files/privacy-final.pdf

The dialogue was based primarily on three clusters building on the 1980 Organisation for Economic Co-operation and Development (OECD) Privacy Principles:[4]

- Protection and security
- Accountability
- Rights and responsibilities for using personal data

This document captures some of the key outcomes of the dialogue. It highlights areas that need to be resolved in order to achieve a sustainable balance of growth and protection in the use of personal data.

Protection and Security

Issues of protection, security and the overall stewardship of personal data remain central to the ecosystem. While the complexity of operating in a decentralized and distributed networked environment poses new challenges, ensuring data security remains crucial.

Accountability

Ensuring stakeholder accountability is a task that is increasingly challenging. Unlike the case 30 years ago, when the OECD principles were established, the questions of "Who has data about you?" and "Where is the data about you located?" are impossible to answer today. The challenge surrounding accountability focuses both on which principles to support as well as how to effectively uphold and enforce them, particularly given the lack of resolution on means of accountability. This contributes to a lack of trust throughout the ecosystem. However, technology itself has the potential to be part of the solution in ensuring accountability at scale through appropriate controls and auditing functionality. *Privacy by Design* which has been widely adopted around the world is key to ensuring privacy is proactively embedded into the technology itself.[5]

[4] http://oecdprivacy.org/
[5] http://privacybydesign.ca/

Principles can serve as a global foundation for creating an interoperable, flexible and accountable framework for coordinated multistakeholder action. Codes of conduct, technological solutions and contract law can all help translate principles into trustworthy practices that enable sustainable economic growth.

Rights and Responsibilities for Using Personal Data

Participants from the public and private sectors shared a variety of perspectives on how the rights and responsibilities for using personal data might evolve. One common concern was that policy frameworks that constrain how data can be linked, shared and used (such as collection limitations, purpose specifications, and use limitations) are increasingly less effective and anachronistic in today's hyperconnected world.

It was also pointed out that as data moves through different phases from collection, to usage and disposal, the weighting of the different principles may need to change. This approach is similar to how incremental advancements in the study of the human genome are being accomplished. Scientists explore and discover the human genome under one set of guidelines; a different set applies when those insights are put into action.

The dialogue also addressed the changing role of the individual. Three subthemes emerged:

From transparency to understanding: There is a need for new approaches that help individuals understand how and when data is being collected, how the data is being used and the implications of those actions. Simplicity, efficient design and usability must lie at the heart of the relationship between individuals and the data generated by and about them.

From passive consent to engaged individuals: Organizations need to engage and empower individuals more effectively and efficiently. Rather than merely providing a binary yes-or-no consent at the initial point of collection, individuals need new ways to exercise choice and control, especially where data uses most affect them. They need a better understanding of the overall value exchange so that they can make truly informed choices.

From black and white to shades of gray: Context matters. Given the complexity of applications, the idiosyncrasy of individual behaviours and the speed of change, there is a need for flexibility to allow different approaches to using data in different situations (See Appendix for a range of case studies of the use of personal data in different contexts).

To keep pace with the velocity of change, stakeholders need to more effectively understand the dynamics of how the personal data ecosystem operates. A better coordinated way to share learning, shorten feedback loops and improve evidence-based policy-making must be established.

Key Messages from Global Dialogue

- The world has changed, which creates new opportunities but also risks
- A new approach to personal data is needed that is flexible and adaptive to encourage innovation, but also protects the rights of individuals. Notice and consent need to be reconsidered to be equipped for this changing world.
- Key aspects of this new approach include:
  - Shifting from governing the usage of data rather than the data itself
  - Context is key in a world of increasing shades of grey. Black and white solutions won't work
  - New ways to engage the individual, help them understand and provide them the tools to make real choices based on clear value exchange
- A number of potential ways forward emerged from the dialogue:
  - The importance of establishing an updated set of principles and the means to uphold them in a hyperconnected world
  - Technology can be part of the solution – allowing permissions to flow with the data and ensuring accountability at scale
  - Need to demonstrate how a usage, contextual model can work in specific real world application

Figure 1: World Economic Forum dialogue on principles for trusted flow of personal data

San Jose
(March 2012)
- Individual rights to data
- Usability, accountability user control

London
(June 2012)
- Duties by others to ensure individual rights
- Technology enabled policy solutions

Tianjin
(September 2012)
- Reframing existing principles into new ones
- Importance of context

Brussels
(October 2012)
- Revisit openness & individual participation
- Redefine collection & use specification

Dubai
(November 2012)
- Transparency
- Empowered role for individual
- Respect for context

Davos
(January 2013)
- Shift focus from collection to use of data
- Engage individuals through real choice not binary consent
- Importance of context in which data used

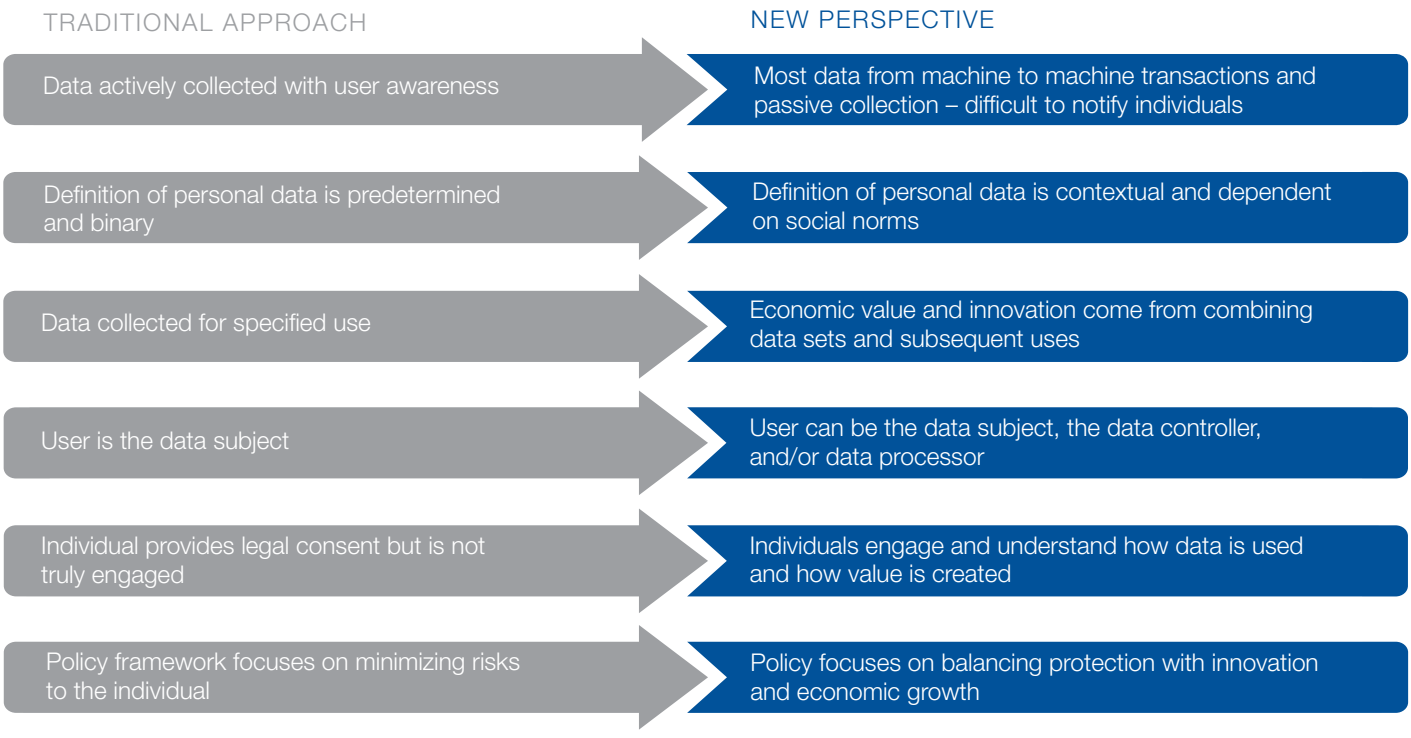
Source: World Economic Forum and The Boston Consulting Group

Chapter 1: The World Is Changing

The world is changing fast. A new computing and information-sharing architecture has emerged during the past 10 years. The policies, business models, social norms and technologies of today are simply different from what existed before. Analytics have become the new engine of economic and social value creation. The discovery and insights derived from linking previously disparate bits of data have become essential for innovation (See Figure 2 for more details).

More data is being collected, processed and transferred than ever before. Data is collected by billions of connected devices, people and sensors that record trillions of transactions and behaviours each day. The unprecedented amount of data being generated is created in multiple ways. Data is actively collected from individuals who provide it in traditional ways (by filling out forms, surveys, registrations and so on). They are also passively collected as a by-product of other activities (for example Web browsing,

Figure 2: New perspectives on the use of data



Source: World Economic Forum and The Boston Consulting Group

location information from phones and credit card purchases). The increasing use of machine-to-machine transactions, which do not involve human interaction, is generating significant amounts of data about individuals. All of this data is further analysed and commingled to create inferred data.

Data-driven opportunities are not without risk and uncertainty. The issue is how to gain new insights and make better decisions, and to do so in a manner that recognizes and protects consumers, businesses and governments against growing concerns of security, privacy and other harms.

The forward transfer of data creates one class of uncertainties. The commercial incentive to share data with secondary and tertiary parties is strong and deeply embedded in existing Internet business models. While the transfer of data creates leverage with each additional use, it also renders the challenges of accounting for and monitoring the use of the data more complex. As more and more data is combined and commingled, the insights, discoveries, value and potential risks increase, particularly if this activity is performed by parties not directly known or necessary to the underlying transactors.

With more than 6 billion people connected to mobile devices, an increasing variety of data is also becoming capable of being linked to individual identity. Smartphones are now able to capture and track an individual's location patterns as well as help create new levels of authentication.

In addition, individuals are no longer merely the subjects of data – they are also being recognized as "producers" of data. For example, digital personal-health devices such as Fitbit[6] and Nike+ Fuelband[7] measure daily physical activities. They provide a new way of capturing a rich data set about an individual. These devices present an opportunity to combine and commingle intimate, high-resolution, activity-based health data with other data sets to provide a daily health dashboard for individuals. It helps them set wellness targets, measure progress and more effectively engage in achieving healthier lifestyles.

But such personal-health data also gives rise to new questions and challenges for individuals and institutions. For example, can these data be combined with traditional medical records for research and treatment? Is the device reliable and accurate? Can the data be authenticated and linked to only one person? Can insurance companies use the data in their coverage decisions? Such concerns are valid and need to be addressed, but preemptively fencing off certain data devices and types because of these concerns would reduce innovation, discovery, and value to individuals and businesses.

Using data for purposes in addition to those originally identified can raise privacy concerns if those uses are inconsistent with the interests of the data subject. However, as always, context matters. Restrictions on the use of data may also put the discovery of transformative innovations at risk.

For example, using a robust database of 3.2 million individuals, Kaiser Permanente addressed the biologic factors linking parental antidepressant-drug use to childhood autism spectrum disorders (ASDs). Analysis of data taken from the personal medical records of related family members from 1995 through 2002 showed that children exposed prenatally to their mother's use of antidepressants had more than twice the risk of developing ASDs. The results of the study and this rate of impact may affect the care of children and parents drawn from a total of over 4 million births per year in the US, and over 5 million births per year in EU countries together.

Another example is Visa's adaptation of its transactional data to protect consumers and merchants from fraud. The primary purpose of collecting these data is to ensure convenient, safe and reliable payment processing. Using the data to prevent fraud creates value for all participants in the payments ecosystem. Scams and fraud trends are identified as they happen, not hours or days later. This results in approximately US$ 1.5 billion global fraud identified annually.

These examples indicate that even data that is seen as particularly sensitive in some contexts can in other contexts be freed to yield important insights and value to all (See Figure 3 and Appendix for further case studies).

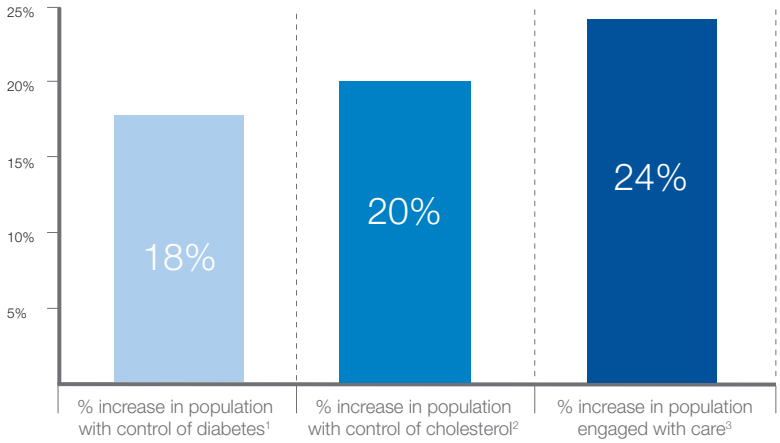
[6] www.fitbit.com
[7] http://www.nike.com/us/en_us/lp/nikeplus-fuelband

Figure 3: Use of personal data in innovative ways in the health sector can yield significant economic and social value for all stakeholders

Engaging the individual and leveraging health data to solve global health's toughest problem – tackling chronic disease

Engaging the individual and undertaking personalised individual interventions based on health data has led to significant improvements in health outcomes in Abu Dhabi

IMPROVEMENTS IN HEALTH OUTCOMES DUE TO LEVERAGING PERSONAL DATA IN TRUSTED WAY

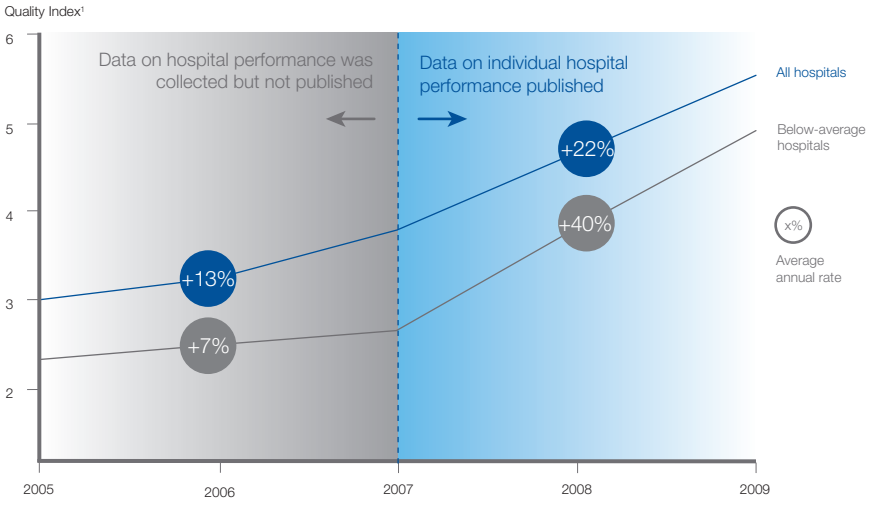


Source: Health Authority – Abu Dhabi (2010 full-year data); See Appendix for more details
1: HbA1c<7.5%; 2: LDL:HDL ratio <3.5; 3: At least one diabetes follow-up in the year (HbA1c measurement)

Systematically collecting and making publicly available health outcome data drives clinical improvement and reduces costs

Public disclosure of aggregated, anonymized patient outcome data can drive improved clinical outcomes and address healthcare costs

TRANSPARENCY OF OUTCOMES DATA IN SWEDEN LED TO IMPROVED QUALITY OF PATIENT CARE ESPECIALLY AT THE LOWEST PERFORMING HOSPITALS



Source: RIK-HIA annual reports 2005-2009; BCG analysis; See Appendix for more details
1: The quality index is based on nine process interrventions recommended by the European Cardiology Society.

Source: The Boston Consulting Group, December 2011, "Improving Health Care Value: The Case for Disease Registries", https://www.bcgperspectives.com/content/articles/health_care_payers_providers_biopharma_improving_health_care_value_disease_registries/

Chapter 2: The Need for a New Approach

Given the complexity of the personal data ecosystem, the rate of change, the potential for significant value from data and the changing role of the individual, there is a need for a flexible, adaptive and resilient approach that has at its heart the aim of enabling the global trusted flow of data.

The traditional data-protection approach, based on 1970s computing architectures in which governments and large organizations operated in discrete silos, was that the individual is involved in consenting to data use at the time of collection. The organization that collected the data then used it for a specified use, based on user consent, and then deleted the data when it was no longer needed for the specified purpose. That approach was appropriate when the data collection was often related to a specific service, a single organization or single use and when the computer data systems were not highly interconnected.

Now, however, the walls of enterprise computing have opened up along with the data flows across traditional silos.

Traditional approaches are no longer fit for the purposes for which they were designed, for several reasons:

- They fail to account for the possibility that new and beneficial uses for the data will be discovered, long after the time of collection.
- They do not account for networked data architectures that lower the cost of data collection, transfer and processing to nearly zero, and enable multiuser access to a single piece of data.
- The torrent of data being generated from and about data subjects imposes an undue cognitive burden on individual data subjects. Overwhelming them with notices is ultimately disempowering and ineffective in terms of protection – it would take the average person about 250 working hours every year, or about 30 full working days – to actually read the privacy policies of the websites they visit in a year.[8]
- In many instances (for example, while driving a car or when data is collected using many M2M methods), it is no longer practical or effective to gain the consent of individuals using traditional approaches.

[8] http://www.techdirt.com/articles/20120420/10560418585/to-read-all-privacy-policies-you-encounter-youd-need-to-take-month-off-work-each-year.shtml

Putting Context into Context

One of the key buzzwords in the dialogue around personal data and privacy is context. Some of the phrases which echoed throughout the global dialogue included:

- "Context matters"
- "Companies need to respect the context in which data was collected"
- "We need different approaches depending on the context"

But what does context mean? The formal definition is "the circumstances that form the setting for an event, statement, or idea, and in terms of which it can be fully understood and assessed." Like money left under a mattress, data is inert until it is used by someone for some purpose – creating value or potentially harm for the individual, an organization or society. The "context" is the description of the conditions of such use.

In terms of personal-data usage, context includes the type of data, the type of entity involved, the trust of the service provider, the collection method, the device context, the usage application, and the value exchange between parties.[9]

During the World Economic Forum dialogue series, this notion came up time and time again. There was widespread agreement that a more flexible approach that takes into account the data context was one of the big shifts required in adapting existing approaches.

[9] Some research has shown that users take these elements into consideration in assessing what restrictions, consent and notification may or may not be required. See for example International Institute of Communications. "Personal Data Management: The User's Perspective", http://iicom.org/resources/open-access-resources/doc_details/264-personal-data-management-the-user-s-perspective-pdf-report

But a new approach cannot be one of "anything goes". The potential for new value creation from allowing data to flow and combine with other data needs to be balanced against the potential risks and intrusions this could cause. This requires a shift in thinking from focussing on data protection to enabling data empowerment. It will require a shift from controlling data collection to focusing on data usage. Lastly permissions, controls and trustworthy data practices need to be established that enable the value-creating applications of data but prevent the intrusive and damaging ones. Data itself does not create value or cause problems; its use does.

The new approach also requires a shift from focusing on protecting individuals from all possible risks to identifying risks and facilitating responsible uses within those boundaries. In some cases, failure to use data (for example, to diagnose a medical condition) can lead to bad outcomes – not only at an individual or societal level, but also in economic terms, just as its use can create risks. It also requires acknowledging that not all data and situations are the same. As we have stated before, context matters, and one-size-fits-all approaches will not work.

This new approach also needs to carefully distinguish between using data for discovery to generate insight and the subsequent application of those insights to impact an individual. Often in the process of discovery, when combining data and looking for patterns and insights, possible applications are not always clear. Allowing data to be used for discovery more freely, but ensuring appropriate controls over the applications of that discovery to protect the individual, is one way of striking the balance between social and economic value creation and protection.

However, just as the discovery of new opportunities for growth is unknown, so are the possibilities for unleashing unintended consequences. Principled and flexible governance is required to assess the risk profile of actions taken in the use of data analytics.

Because future, yet-to-be-discovered uses of data cannot be fully anticipated, a default policy of deleting data in all contexts can be harmful. A better approach is to manage use in ways that can evolve over time, protecting both the rights and the future options of the individual, and the groups and institutions with which the individual exchanges data. Principles can provide both the foundations for such a shift and the flexibility for innovation.

But managing such a flexible, dynamic system will not be easy. It will require action by all stakeholders coming together to agree on refreshed guiding principles and ways to implement them including codes of conduct and technological solutions. There is also a need for better evidence of what works and what does not to inform the behaviour of all. And there is a need for an interoperable global policy framework that incorporates this new approach.

The Evolution of Personal Data

The definition of personal data is evolving. Traditionally, that definition was pre-determined and governed through the use of a binary approach: In most jurisdictions, the use of personally identifiable information (PII) was subject to strict restrictions whereas the use of non-PII was often uncontrolled.

However, what is considered personal data is increasingly contextual; it changes with personal preferences, new applications, context of uses, and changes in cultural and social norms.

Traditionally, organizations have used a variety of techniques to de-identify data and create value for society while protecting an individual's privacy. Such data was not subject to the same rules as PII, as an individual could not be identified from it. But technological advances and the ability to associate data across multiple sources is shifting boundaries of what is or is not PII, including potential re-identification of previously anonymized data.

This issue is the subject of significant debate with some arguing that this means that all data is effectively personally identifiable and should be treated as such. Others urge caution, arguing that this would curtail many of the beneficial uses of anonymous data with minimal gains in privacy.

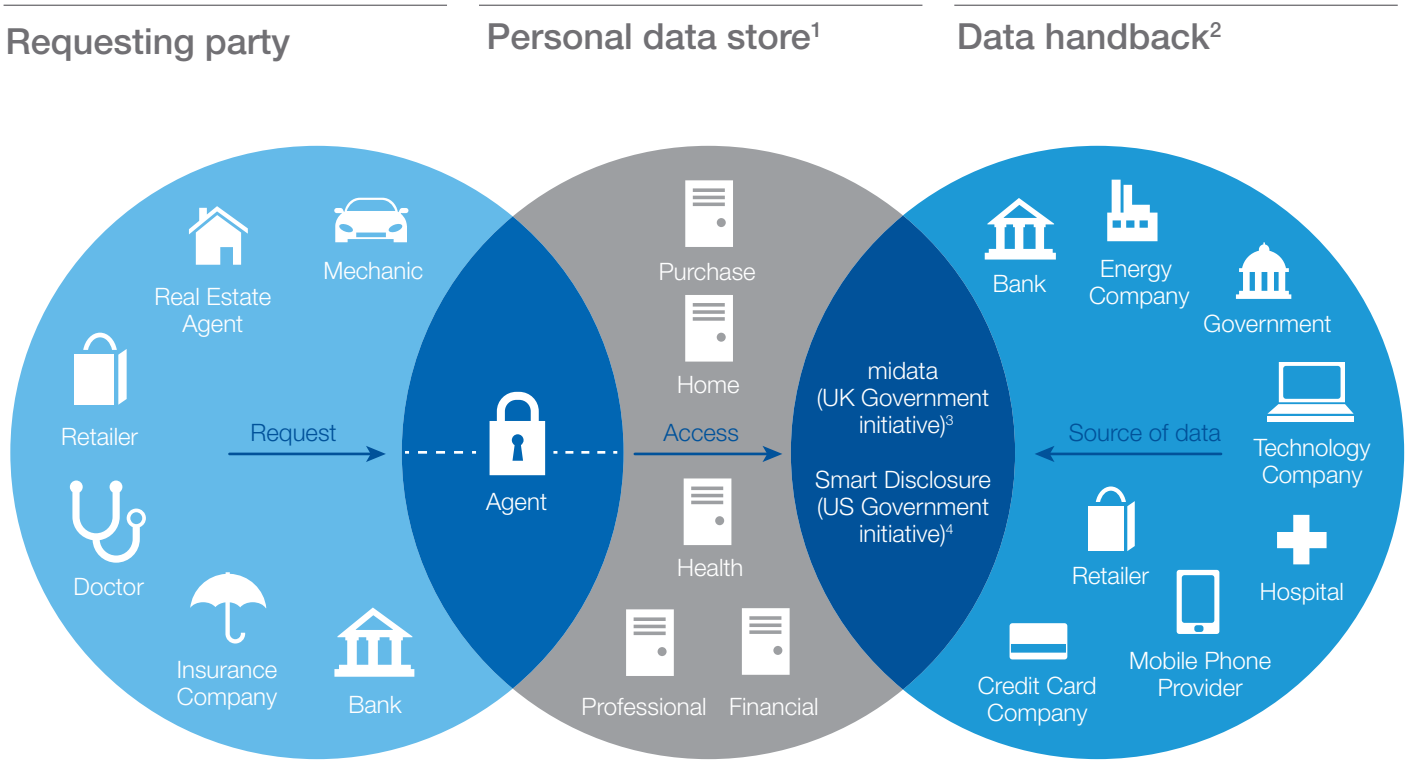
A shift in approach to thinking less about the data and more about the usage could offer a way forward. If the usage impacts an individual directly it would require different levels of governance than data which is used in an aggregated and anonymized manner.

Figure 4: Mechanisms to engage individuals and empower them are beginning to emerge

One of the missing elements of the dialogue around personal data has been how to effectively engage the individual and give them a voice and tools to express choice and control over how data about them is used.

Over the past twelve months there has been significant momentum in terms of personal data stores that provide individuals a place to store and control how a copy of their data is used and government initiatives to encourage organisations to give individuals a copy of data about them.

Market based mechanisms like this have the potential to give the individuals a real voice and say in how personal data is governed though consumer take-up remains a challenge.



Companies who want to access data about individuals can request it through data agents

Several data stores are now up and running allowing individuals to exercise control over how data about them is used

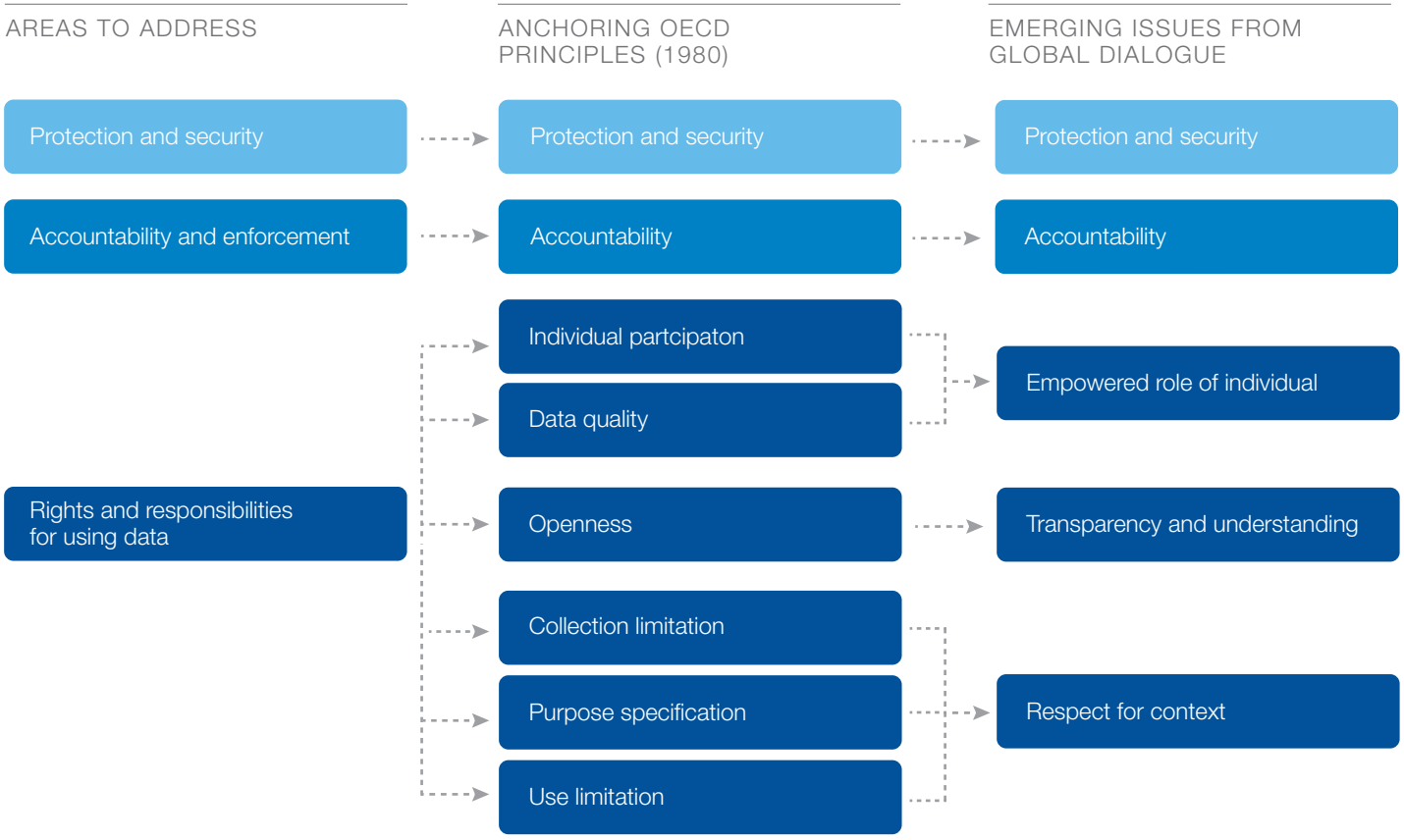
Several governments are working with the private sector to give individuals access to a copy of data about them in a usable format which can then be stored in their locker and shared with other providers

Source: World Economic Forum and The Boston Consulting Group building on original graphic by Forrester Research
1: Also known as vaults/lockers 2: Also referred to as data portability or smart disclosure 3: midata is a United Kingdom Government initiative working with the private sector which is assessing how to give people their personal data in a format that is safe to pass onto third parties, such as price comparison sites – See https://www.gov.uk/government/news/better-choices-better-deals for more information 4: Smart Disclosure is a United States Government initiative giving consumers access to the information they need to make informed decisions in usable data formats, so that innovators can create new interactive tools for consumers. See www.Consumer.Data.gov for more information

Chapter 3: Principles for the Trusted Flow of Personal Data

Principles have been and need to be a core part of the future governance of the personal-data ecosystem. Principles can set the foundation for trustworthy data practice and help empower users. But principles alone are not enough. Combined with technological solutions and accompanied by underlying tools such as codes of conduct, they can not only provide the flexibility required in a fast-moving connected world, but also enable the accountability and enforceability needed to cultivate trust. Identifying and refining the principles that reflect societal and cultural norms and ensuring ways to uphold them will enable trustworthy data practices, persuading individuals to be more willing to share data about themselves.

Figure 5: Areas to focus on to achieve trusted flow of data emerging from dialogue series



Source: World Economic Forum and The Boston Consulting Group

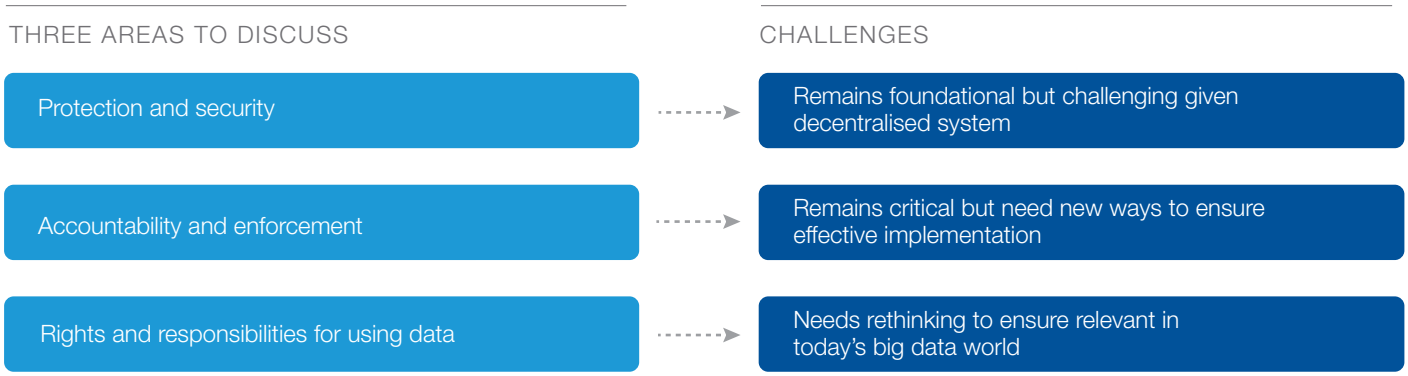
Existing principles associated with the collection, handling and use of personal data have formed the basis of most privacy and data-protection legislation around the world. A version of these principles was agreed to internationally in 1980 in the form of the OECD Privacy Principles.[10]

However, as discussed previously, principles need to be periodically revisited and updated to reflect current practices and to address changed circumstances in technology and society. The world has changed dramatically in the last five years, let alone the three decades since the OECD principles were agreed upon. It is therefore important to reconsider how these principles can be upheld and updated in a way that is appropriate for a hyperconnected world.

This dialogue started from the broadly cited OECD principles and focused on the question, "What elements of these principles need to change to address current and anticipated future challenges?" There was broad consensus that change is needed to these principles to ensure they are relevant for this changing world.

The World Economic Forum's dialogue clustered existing OECD principles into three broad categories. This initial clustering exercise enabled insight into a current view of the overall purpose of the individual principles, which served to inform how a given principle might appropriately be updated, while still maintaining maximum "backward compatibility" with the original aims of the principles (See Figure 5).

Figure 6: The World Economic Forum dialogue grouping of existing principles



Source: World Economic Forum and The Boston Consulting Group

To support this process, the World Economic Forum held a global, multistakeholder dialogue on personal data throughout 2012 in the US, Europe, Asia and the Middle East (See Figure 1 for more details). This dialogue has involved extensive participation from the private and public sectors involving more than 40 companies from IT, telecommunications, health, financial services, logistics, aviation and professional services as well as policy-makers, advocacy groups, and others from the US, the EU and beyond. Results from primary research by leading academics were also incorporated into this series of discussions.[11] The unique perspective of representatives from international organizations such as the World Bank and the United Nations added additional perspective on the challenges they face and the increasing need for trusted information flows.

The OECD principles were very carefully thought out. This fact is reflected in the observation that a number of the OECD principles remain relevant today. The issue is that they need updating in terms of the way in which they are applied and upheld in today's hyperconnected world

Protection and Security

Security figures prominently in the original principles and continues to be foundational. However, approaches to security need to reflect today's decentralized world. Securing personal data is increasingly difficult in a distributed network system with multiple parties involved in storage and management – no one party can do it alone. Dependent on the behaviour of others, all stake-holders collect, hold and use personal data. They must all take appropriate steps to secure data from accidental release, theft, unauthorized access, and misuse.

[10] In some jurisdictions these principles are known as Fair Information Practice Principles, or FIPPs.

[11] See for example International Institute of Communications. "Personal Data Management: The User's Perspective", http://iicom.org/resources/open-access-resources/doc_details/264-personal-data-management-the-user-s-perspective-pdf-report ; Rubinstein, Ira. "Big Data: The End of Privacy or a New Beginning?" October 2012. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2157659 ; Tene, Omer and Jules Polonetsky. "Big Data for All: Privacy and User Control in the Age of Analytics". Northwestern Journal of Technology and Intellectual Property, Forthcoming. http://ssrn.com/abstract=2149364.

Accountability

Accountability remains critical, but we need new ways to ensure effective implementation in a hyperconnected world. For the trusted flow of data, all stakeholders should be accountable for how they collect, store, secure, use and share data. But accountability alone is not sufficient. There is also a need for effective enforcement to ensure systemic trust. Yet creating accountability and enforcement in a rapidly changing, hyperconnected world is increasingly difficult given the external pressure for increased flexibility in design of rules. There are numerous ongoing efforts focused on how to build such accountability working with data protection regulators and companies, including exploring co-regulatory approaches as a way to develop a more flexible, contextually relevant, and efficient approach.[12]

Rights and Responsibilities for Using Personal Data

However, other principles, particularly those that establish rights and responsibility for using data, need significant rethinking to reflect the changes in the world. These changes include the increasing recognition of the role of individuals as both producers and consumers of data, the number of new beneficial uses of data discovered long after the point of collection, and the sheer volume of data being created. Other emerging concepts that were not anticipated at the time of the original drafting of the OECD principles include the recognition that all data is "dual use" (it can be used for good or bad purposes), and the understanding that there is a direct correlation between the value of data and the potential intrusiveness of its use.

In particular, reliance on mechanisms of "notice and consent" to ensure individual participation are seen as increasingly anachronistic. The current manifestation of the principles through notice and consent as a binary, one-time only involvement of the individual at the point of data collection was identified in the dialogue as an area ripe for reconsideration to better empower individuals, build trust in the system, and encourage the reliable, predictable and more valuable flow of data into and within the system.

Other areas identified as candidates for reconsideration include requirements to specify, in detail, the purpose of usage at the time of collection and to restrict future uses to that purpose. In the past, this was a viable solution when collected data was much more isolated and was not subject to the correlations that can reveal valuable new information. Given that much of the innovation and therefore economic and social value come from subsequent uses of data, there is work to be done in balancing the rights of individuals yet recognizing that notions of the "single use" of data are increasingly difficult to embrace.

In addition to identifying areas where existing principles need to be refined, the dialogue pinpointed three key areas in establishing the rights and responsibilities for using personal data that could form the basis for the evolution of existing principles. (See Figure 6)

From transparency to understanding: New ways to inform individuals and help them understand how data about them is being collected and used are needed. This does not mean that individuals have to understand every detail of every data flow, but they do need to have a broad understanding and a greater sense of control of what is happening to data about them to ensure trust.

The current approach to providing transparency through lengthy and complex legalistic privacy policies overwhelms individuals rather than informs them. The challenges are compounded as more data is being collected by more and more devices, many of which are not within the direct control of the subject of the data. Simplicity, efficient design and usability must lie at the heart of transparency. As one participant in Davos noted, effective design should be applied to engaging the individual not just to make the website look better.

While it may be impossible to completely move away from legally derived privacy policies, there are many potential ways to help foster this shift to real understanding. There are indications that a data literacy movement is beginning to emerge in North America and Europe to help cultivate real understanding (See Figure 7).

Some companies are aiming to develop simple language explanations of their approach to data use so that the individual can more quickly understand the main elements of how data is being used without having to wade through the legal privacy policy. Intuit,[13] for example, has established "Data Stewardship Principles" that are at the heart of how the company deals with personal data. This sets out in clear simple language what Intuit stands for, what it will do, and what it will not do. For example, the principles make clear that Intuit will not sell, publish or share data that identifies any person. But the company will use data to help customers improve their financial lives and to operate its business. And it will give customers a choice about how Intuit uses data that identify them.

A challenge for companies in preparing these "simplified" approaches is whether the new approaches will be sufficiently detailed to pass muster from a "full and conspicuous disclosure" perspective. That problem can be mitigated through the adoption by companies of more standardized language for part or all of their policies. This would allow all parties (businesses and consumers) to enjoy the benefits of more familiar and predictable systems, including the legal and rules portion of networked systems.

Recognizing the benefits of standard legal language in helping to normalize the user experience and in reducing risk for both users and businesses in existing markets, others are aiming to standardize and score protection approaches by different companies. Privacyscore,[14] for example, analyses the privacy policies of companies along four clear criteria and gives each website a colour-coded rating and score. In this way, Privacyscore is able to help translate legalese into a clear and easy-to-understand guide (See Figure 7).

Mozilla has proposed a symbols-based approach to presentation of legal terms that features a number of icons that signal, for example, how long data is retained, whether data is used by third parties, if and how data is shared with advertisers, and whether law enforcement can access the data (see Figure 7).[15]

[12] For example the CIPL project on Accountability, - http://www.informationpolicycentre.com/accountability-based_privacy_governance.

[13] http://security.intuit.com/privacy/data-stewardship.html
[14] http://privacyscore.com/
[15] https://wiki.mozilla.org/Privacy_Icons

Although transparency is not a new principle, it warrants revisiting in light of the complexity inherent in a hyperconnected world. A refined principle that focuses less on whether information is disclosed and more on truly seeking to help individuals understand how data about them is being collected and used is foundational to the accomplishment of other fair principles of usage. If the individual does not understand a system, he or she cannot effectively engage with it. When it comes to transparency, less can sometimes be more.

From passive consent to engaged individuals: Organizations (the operations of which depend on a relationship of positive engagement with their customers) need to understand and accommodate the changing role of the individual by engaging with and empowering them. The individual was historically seen as a "data subject" – a passive consumer of products and services who was tracked for customer relationship management purposes only and needed to be notified about how data about them is being used and consent to that use. However, increasingly individuals are being understood to act as both producers and consumers of data. The current model of notice and consent at the point of collection has not led to a level of engagement by individuals in terms of how data about them is used; nor is it necessarily commensurate with the value that the assets provide.

Given the sheer volume of data and the various ways that data is collected and used today, it is, as a practical matter, physically impossible for an individual to consent to all the different data uses. Rather than relying on yes-or-no consent at the point of collection, individuals need new ways to exercise more effective choice and control when data is being used in a way that impacts them. As part of this, organizations must be clear to individuals about the value exchange that is taking place for data, in terms of monetary and other benefits, so that those individuals can make truly informed choices between different options based on what they consider fair.

Consider how BT implemented the recent update to the EU e-privacy directive, often referred to as the "cookie law", which requires companies to obtain the consent of their website users before using cookies to track behaviour online and to personalize services. Whereas most companies put in place a pop-up box asking users to click to consent (a standard but relatively opaque process), BT implemented an easy-to-understand practice for visitors to its website. A simple pop-up screen allows users to discern the strictly necessary cookies required for the site to operate properly (from which customers do not have the right to opt out) and the functional and targeting cookies that enable potentially "intrusive" social sharing and behavioural tracking, but that also enable the best experience for site users. The company clearly explained what customers get for the information they give, helping individuals to make an informed and engaged choice (See Figure 7).

Technology and new approaches can clearly help. Organizations need to build simple-to-use tools that encourage individuals to become engaged in setting the policy governing use of data and to be able to change those settings over time without being overwhelmed. Usability and simplicity are key to effectively engaging the individual and enabling users to see and understand equitable benefits, keeping in mind that the benefits may sometimes be shared between the organization and the individual and even with society in general. In addition, organizations can make better use of metadata and leverage existing contract law to create simpler and more engaging ways to empower the individual.

Finally, enabling and encouraging forms of "peer support" becomes increasingly possible in social-network settings, as evidenced by the multiple rating sites, blogs, FAQs and so on that are increasingly available to enable consumer choice. In this latter case, businesses that encourage the formation of community around their customers can leverage those relationships to help solve business problems. As is often the case in networked information systems, the source of the problem can also be the source of the solution.

Potentially, markets can encourage a "race to the top" in which user control and understanding of how data is used and leveraged become competitive differentiators. Various trust marks and independent scoring systems will help stimulate this kind of response.

Given the complexity of choices, there is also potential for the development of "agency type" services to be offered to help individuals. In such a scenario, parties would assist others (often for a commission or other fee) in a variety of complex settings. Financial advisers, real estate agents, bankers, insurance brokers and other similar "agency" roles are familiar examples of situations when one party exercises choice and control for another party via intermediary arrangements. Just as individuals have banks and financial advisers to leverage their financial assets and take care of their interests for them, the same type of "on behalf of" services are already starting to be offered with respect to data.

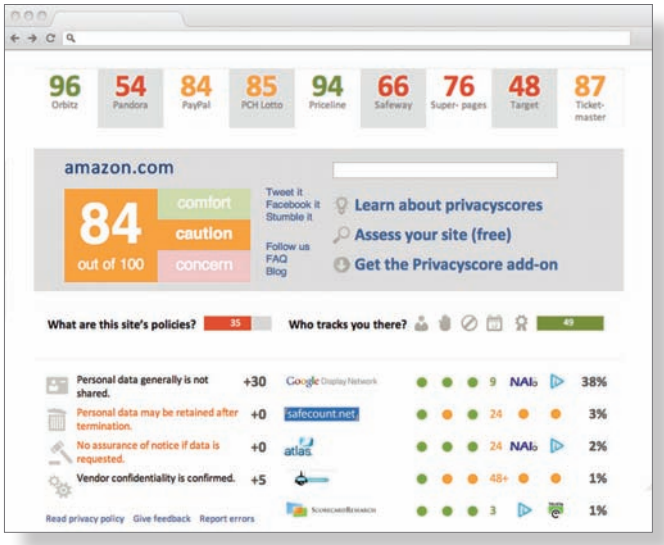
From black and white to shades of gray: Given the complexity and speed of change, flexibility is needed to enable the simultaneous deployment of different but complementary approaches depending on the context in which data is used. For example, an appropriate data-usage practice for treating an individual as a patient in a medical emergency situation may not be appropriate for that individual in financial services settings or for targeted advertising.

The challenges of contextual complexity have at least two implications. First, there is a clear need to avoid a one-size-fits-all approach to issues including consent, notice, what is and is not personal data, and more, given that the context of the data use is crucial to determining what is and isn't appropriate. For example, permissions to use data within a company to fulfil a customer order will differ from those associated with using this data for a completely unrelated purpose that may have been created through subsequent analysis. This is not to say that all potential uses of data need to be mapped out in detail for every data collection, but clearly there is work to be done in improving the information flow between data subjects and data collectors so that individuals can form reasonable expectations, and have those expectations met.

Second, the importance of context strengthens the need to shift the focus of engagement for the individual from the point of collection to the point of usage. In the past, when data was not networked and was used only once, it was possible to declare specifically why a particular set of data was being collected. There may still be situations where data is appropriately collected for only a single purpose and a single use, but in the era of big data,

Figure 7: Emergence of different tools that help individuals understand how personal data is used

Privacyscore helps individuals understand how different websites use data

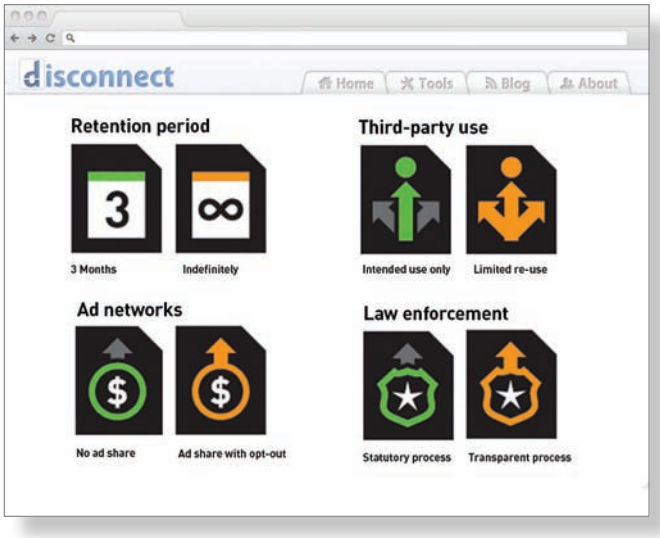


Source: www.privacyscore.com

WolframAlpha helps individuals visualize vast quantities of social data about them

Source: http://www.wolframalpha.com/facebook/

Mozilla has proposed icons which would create a common set of symbols for how websites use data



Source: https://wiki.mozilla.org/Privacy_Icons

BT offers users a sliding scale of choices with a clear explanation of the value exchange taking place

Source: http://www.bt.com/; https://github.com/BTplc/Cookies

single-use collections reflect a decreasing percentage of overall data collection. This point is fundamental to the design of future data systems that can harness the value of multiple instances of data leverage while still protecting stakeholder rights.

The advent of systems to enable the replication of this type of leverage on a large scale would have potentially dramatic economic benefits. If data is deleted after its first use then the potential future economic and social value that could be created from subsequent uses is lost. Yet, this potential leverage is not without potential future risks of misuse of the data.

There are legitimate reasons why individuals and organizations may want to delete data. Retention of data involves both costs and risks including of it being breached or misused. The current approach to data does not effectively ensure the security of data, does not ensure accountable trustworthy data practices by data handlers nor does it give individuals an effective understanding and control for how data can and cannot be used. Under such circumstances, it is only natural that many people see data deletion as the only tool an individual has to combat misuse. Without clearly established trustworthy data practices, many organizations too are increasingly seeing that data is a potential liability as well as an asset and are choosing to delete data to prevent the downside risks.

It is also important to distinguish between data being used to generate insights and discover new patterns and how they are applied to the individual. One way of dealing with this ambiguity is to more clearly identify specific risks and intrusions of concern. Once these are known, actions to manage and prevent them can be addressed.

The approach of "reasonableness" has also been advanced as one possible way to help manage the risk of harms. Establishing easily understood and reasonable expectations could help increase trust and reliability.

Although a contextual approach is more flexible and able to strike the balance between using data to create value and protecting the individual, it is difficult to implement. The challenge is defining data permissions and allowable use contexts.

To address this challenge, it is critical to be able to answer a number of questions. What is the provenance of the data? What are the associated permissions for accessing and using it? And what are the allowable circumstances of use?

Global Momentum to Establish New Norms for Personal Data

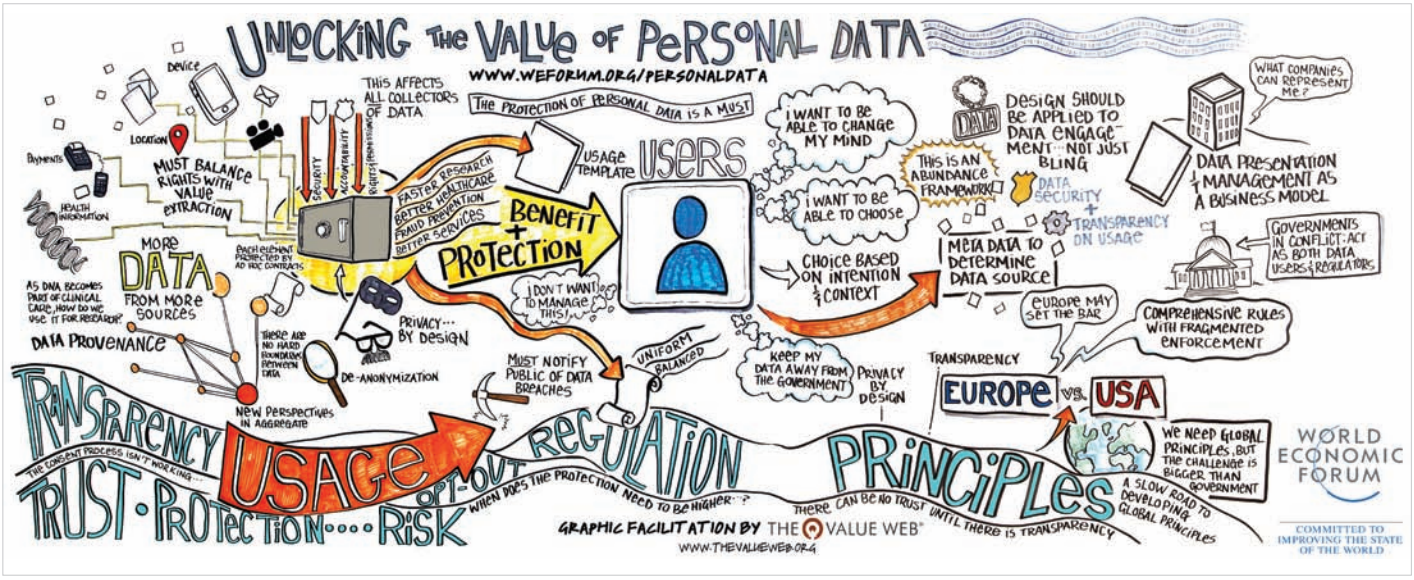
In addition to the World Economic Forum's efforts to convene a multistakeholder dialogue, various other groups are exhibiting increasing momentum to establish new and evolving norms to guide how personal data can be used to create value.

For example, the OECD and its member governments have been discussing how to refresh the OECD principles for a hyperconnected world. Other groups such as the Centre for Information Policy Leadership (CIPL) have been focusing on accountability, one of the key aspects of the principles. In addition, different sector groupings and regional authorities have been considering how these principles apply to their particular applications. The GSMA has developed principles for mobile privacy, and the Digital Advertising Alliance has developed principles for the use of data in online behavioural advertising. In addition, the Asia-Pacific Economic Cooperation (APEC) forum is establishing a cross-border privacy rules system to harmonize approaches throughout the region.

Privacy by Design was adopted, in October 2010, as the global privacy standard in a resolution by the International Data Protection and Privacy Commissioners in Jerusalem at their annual conference. It has since been incorporated in various regulations around the world and is a real and recent evolution to privacy principles on an international level.

The proposed European Commission Data Protection Regulation currently under discussion by the European Council and Parliament is the most comprehensive attempt to establish new norms for the flow of personal data. While differing views were expressed throughout the dialogue on these proposals and the underlying principles, it was clear that these rules when agreed will have a significant impact on the global governance of personal data.

Panel of executives addresses participants in Davos workshop

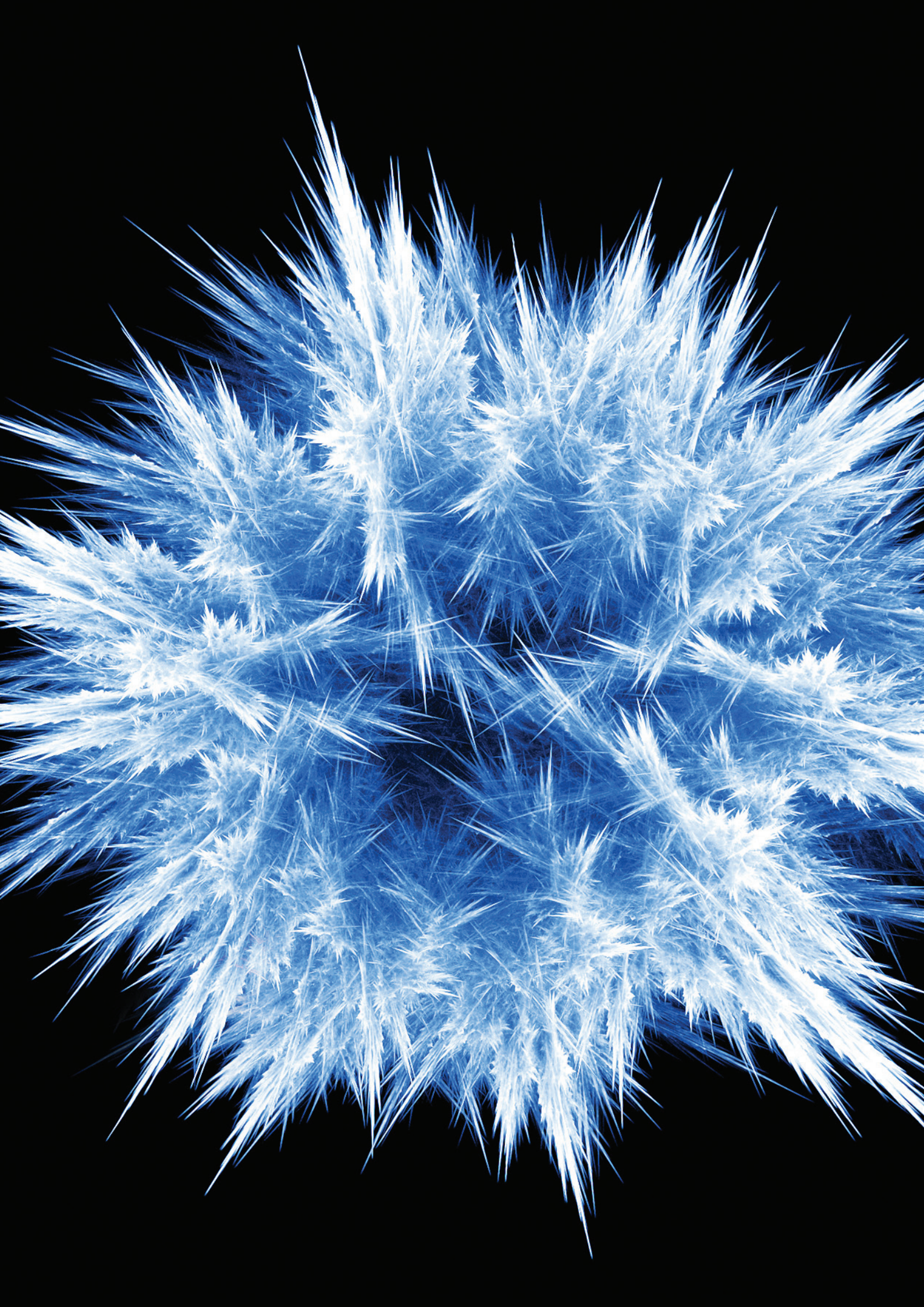


Graphic summary of Davos workshop on personal data

Ray Baxter, Kaiser Permanente (left); Ellen Richey, Visa Inc.; Craig Mundie, Microsoft Corporation

Lynn St Amour, Internet Society, leads a breakout discussion in Tianjin, China

Hamadoun Touré, International Telecommunication Union (left); Viviane Reding, European Commission



Chapter 4: Principles into Practice

Principles by themselves are not enough. To translate principles into practice, a number of steps must be taken.

It is important to build a better evidence base that informs all stakeholders about how the managed use of data can create socioeconomic value, and to better understand user attitudes and behaviours regarding the use of data in different contexts. The achievement of sustainable economic growth requires clear insights on both fronts. This evidence will help make the case for the value of a trusted flow of personal data. In the absence of such evidence, public debate will continue to be dominated by speculation, uninformed fear, uncertainty and doubt.

An evidence base can help facilitate an informed dialogue among stakeholders in the personal data ecosystem with the aim of developing an appropriate policy framework. However, given the pace of change in technologies, society and institutional structures, this evidence will need to be updated constantly, rather than just periodically. An ongoing feedback loop indicating what is working and not working is needed to guide better decision-making and actions by all stakeholders. One of the most broadly shared data system needs across cultures, jurisdictions and contexts is a transparent, simple, responsive and empowering rule-making processes and system operations.

Also needed is an agreed upon set of rights and duties based on principles for trusted flow of data. With a shift in the focus to frameworks focused on the use of data, additional work is needed to design, define and come to agreement on the specifics of "duties of care" associated with data actions such as collection, use, processing, transfer and the like. These can form the basis of industry codes of conduct.

Technology can play a role in the crafting of solutions that will help to enable and facilitate policy alternatives. For example, the metadata-based infrastructure (see sidebar), in which descriptions of actual usage practices are captured, could support increased transparency, predictability and trust and could help establish a strong foundation for trustworthy data practices.

Technology Can Support and Uphold Policy Aims

One potentially promising way to use technology to help address policy goals is to use data system functionality to provide information about the data itself. The generation of so-called metadata is an example of this approach. Metadata is the term used for "data about data". The generation of metadata can enable the system to answer such questions about collection history, uses, and more. Such a system would be structured so that each bit of data actually carries (or is virtually linked to) information about its provenance, permissions, and so on. This approach enables real-time, periodic verification of usage consistent with established restrictions and the maintenance of contextual integrity in the use of the data as they flows through the value chain. However, governance of metadata needs to be carefully managed as its use can also create potential risks and misuse.

The "law" can have effects on data systems both through public law (such as legislation and regulation) and private law (such as through contracts and self-regulatory organization structures). Both types of laws must be flexible.

Government legislation and regulation have a crucial role to play in establishing trusted flow of data, but given the speed of change and complexity, it can never be relied upon to cover everything. As noted above, there is a strong role for co-regulation, including enforceable industry codes of conduct.

Organizations of all types will be well served if they do not continue to try to "go it alone", but instead move to agree on, adopt and ensure compliance with uniform, consensus-based trustworthy data practices. In particular, institutions will benefit if they rethink how they engage with individuals so that those individuals both trust how data about them is being used and have a real stake in those uses.

It is clear that there is a role for simplification of complex systems to engender trust and adoption. Individuals should be provided with access to simple tools that enable them to either understand or set the policy to be applied to the use of data, and be able to change that selection over time. It should be possible for them to delegate the detailed specification of their policy choices to third parties, perhaps through agency arrangements with organizations that can further their values and norms.

But care must be taken. The growth of data collection, transfer and processing is proceeding at exponential rates worldwide, and data uses are also growing at a nonlinear pace. However, people's attention span and cognitive capacity are not keeping pace. Organizations need to avoid overwhelming individuals with information and choice in the name of engagement. The amount of information and choice must also be driven by context.

Making these changes will not be easy, but nor is it optional. Change is never neutral and can often create "winners" and "losers". Such an approach will require changes by all stakeholders to their traditional approaches and a willingness to work together to unlock the value of personal data and to balance growth with protection.

Key Areas for Further Work

- Establishing an updated set of shared principles, and the means to uphold them in a hyperconnected world
- Need to demonstrate how a usage-based contextual model can work in specific real world applications. This will require further thinking on how to establish which uses should be permitted and which should be constrained especially given new uses of data are being discovered so quickly
- Developing usage-based codes of conduct that establish clear norms for trustworthy data practices
- Establishing how technology can be part of the solution – allowing permissions to flow with the data and ensuring accountability at scale
- The space to test and learn to work out what works and what does not given the speed of change and complexity in the personal data ecosystem
- Examine how the most current evolution of privacy principles, *Privacy by Design*, has already assisted in finding privacy protective features for the personal data ecosystem

Appendix

Relevant Use Cases

As discussed in the report, personal data can create significant economic and social value for individuals, companies, government and society at large. However such uses can also raise significant risks of harm to individuals. They also raise challenges that need to be overcome to allow this value to be created.

By considering the issues discussed in this paper in the context of a range of case studies, it is possible to start to see the difficulties of one-size fits all approaches and the importance of a contextual, usage approach to personal data. All of the cases aim to create a trusted flow of data to strike a balance between growth and protection.

The following set of case studies have been produced by members of the World Economic Forum's Working Group on personal data. They cover a range of sectors including health, financial services, telecoms, marketing, automotive, global development, government and more.

Managing chronic conditions through better access to and sharing of data – collaborative cardiac care service

What is it?

- Physicians, nurses and pharmacists collaborate with coronary artery disease (CAD) patients to coordinate care
- Lifestyle modification, medication management, patient education, lab results monitoring and adverse events are coordinated across diverse virtual teams
- Patients have an 88% reduced risk of dying compared to those not in the programme
- Patients at a cholesterol goal went from 26% to 73%
- Cholesterol screening went from 55% to 97%

How was personal data used to create value?

- Immediate access to reliable, evidence-based information at all points of care
- This enables each care team member to support a given patient's care plan, encourage treatment adherence and allow disparate care teams – from primary care to pharmacy to rehabilitation centres – to coordinate care, regardless of setting
- Integrated uses of data aggregated from different entities, caregivers and patient touch points were all used together in coordinated workflows

Scale of the impact

- Clear Results – better survival rates and reduced need for emergency interventions: The results were impressive
- Nationwide, research indicates that fewer than 20% of CAD patients are expected to survive 10 years after their first heart attack. The coordinated, evidence-based care, enabled by Kaiser Permanente HealthConnect (an electronic care registry), and cross-functional coordination increased that survival rate dramatically

Constraints/issues this raises

- Privacy and security: Individual consent and authorization for the use of confidential medical records data for purposes other than treatment, and the security of those data systems
- Ability to accurately, reliably and consistently match individuals to their data from different sources over time

Kaiser Permanente
Source: www.kaiserpermanente.org

Leveraging data to solve global health's toughest problem – tackling chronic disease

What is it?

- Public health initiative launched 2008 aimed at addressing Abu Dhabi's high chronic disease burden
- Comprises tailored whole-population screening then targeted interventions in well-stratified groups delivered at scale, enabled by eHealth technology
- Personalized individual interventions plus group and population interventions driven from hard data (attributable individual-level screening)
- Consent at time of screening and data guardianship model
- In absence of Data Protection Law, voluntary adoption of international standards (Health Authority is Data Guardian)

How was personal data used to create value?

- Weqaya data used to determine impact of different interventions, generating a unique, scalable health feedback loop to identify the most effective interventions
- Weqaya data (with the consent of screened individuals) can also be used for a range of secondary uses, for example, academic research, pharmaceutical R&D and the refinement of risk prediction methods
- Data enables proactive management of health risk
- Opens up game-changing model for outsourcing contingent liability of worsening health to Disease Management entities, enabling governments/payers to directly "Pay for Health"

Scale of the impact

- 200,000 adults screened through the programme at least every three years (more frequently if higher risk)
- ~70% of this population has increased cardiovascular disease risk
- 1/3 with diabetes, 1/2 with high blood pressure and 2/3 with high cholesterol were unaware before screening
- Programme participants show better control of diabetes and dyslipidaemia
- Since Weqaya compliance with care standards far higher for diabetes than for other chronic disease – interest in expanding scope to cover other chronic diseases

Constraints/ issues this raises

- Determining which data may be shared with whom
- Need to strengthen mechanisms and rules for:
  - Data protection, including rights and responsibilities for
  - data use (primary and secondary)
  - Accountability and enforcement
  - Gain share from data use
- Building a cadre of qualified data counsellors to explain complex risk-benefit
- Exploration of dynamic consent beyond original use case
- Opportunities to serve other health markets that have expressed an interest in the model

Weqaya, Abu Dhabi Health Authority
Source: "eHealth as foundation for tackling NCDs" (2012); "Measuring health data management in Abu Dhabi",' http://www.weqaya.ae/en/index.php; www.shafafiya.org; http://www.haad.ae/haad/tabid/819/Default.aspx

Use of mobile phone data to map the spread of malaria in Kenya

What is it?

- Scientists from Harvard School of Public Health conducted a study in Kenya to measure how human travel affects malaria infections
- Scientists showed that the spread of malaria is driven primarily by movement of infected individuals rather than movement of infected mosquitoes
- Mobile phone data used to identify the location of "parasite Hotspots": locations that are highly trafficked by infected Humans, driving disease spread across the nation.

How was personal data used to create value?

- Scientists collected data from almost 15 million mobile phones over the course of one year between June 2008 and June 2009
- Researchers mapped calls and text messages made in 692 settlements to 11.912 cell phone towers to determine duration of trips and population movement patterns
- Individuals were assigned to their primary settlement and data was anonymized
- The data was compared with the information about malaria infection spread provided by the Kenya Medical Research Institute and the Malaria Atlas Project

Scale of the impact

- The study found areas, where targeted malaria intervention programme would be most effective
- Lake Victoria was found to be the area with the highest endemic rate and a major source of further transition of infections to the regions
- Mobile phone data also allowed to show other areas expected to receive most of the infections and demonstrated that the travel patterns were very stable throughout the year
- According to the WHO, malaria death rates went down by 25% since 2000 due to targeted prevention measures

Constraints/issues this raises

- The data used for the study is retrospective, which can be used to identify existing patterns
- The approach cannot be used to map a similar dependency in case of a disease outbreak
- The limit of the approach of using mobile phones to map infection spread is that it only accounts for people who own a phone (around 76% of population)
- The density of cell phone towers may not always allow for further localization of population movement

Harvard School of Public Health

Source: "Quantifying the Impact of Human Mobility on Malaria", Science, 338, 267 (2012).

Systematically collecting and using health outcome data drives clinical improvement and reduces costs

What is it?

- Healthcare systems across developed markets are largely inadequate to meet future demand
- Growth of healthcare cost has outpaced GDP growth and current cost containment efforts have been ineffective
- Estimates of excess or unnecessary costs in the US health systems are as high as US$ 750 billion
- Outcome registries systematically capture data on outcomes for all patients with a given medical condition in a given country, region or clinic
- Value-based healthcare efforts have shown that transparent outcomes data can drive clinical improvement, to improve and save lives and save costs

How was personal data used to create value?

- Transparent, high-quality outcome registries create a platform for clinical improvement
- Transparent outcomes enable comparisons of performance, creating incentives to change and driving faster improvement of low performers
- Active, structured and clinician-led work with outcomes data can drive clinical improvement and increase overall value in the system
- Outcomes registries enable identification, dissemination and adoption of best practice
- Outcomes data analysis allows for improvement more effectively than isolated process metrics

Scale of the impact

- Public disclosure of aggregated, anonymized patient outcomes data at hospital level in Sweden led to a sharp improvement in quality of care
- Below average hospitals improved their quality index by 40% per year as opposed to 7% per year before the disclosure of outcomes data
- If the lower performing half of US hospitals, treating Medicare patients had been able to achieve the median mortality rate for just one medical condition (heart attacks); this would have saved 6,300 quality-adjusted statistical life years – an economic value of US $1.1 billion
- Similar estimates on the cost side suggest that the economic impact of reducing variation in health outcomes in the US health system could save around US$ 200 billion per year

Constraints/issues this raises

- Registries which systematically capture data are more effective in driving improvement than isolated process metrics, but require significantly large datasets – which raises questions about the rights of individuals to opt in or out of such data collection
- Access to outcomes data is a prerequisite to understand performance at all levels – teams of doctors/departments, hospitals, regions/countries, but requires strong top-down mandate or consensus to making data transparent at some levels
- Making the data public has a big incentive effect in motivating and implementing change, but can raise concerns over privacy of data even though it is aggregated and anonymised

The Boston Consulting Group

Source: The Boston Consulting Group, December 2011, "Improving Health Care Value: The Case for Disease Registries", https://www.bcgperspectives.com/content/articles/health_care_payers_providers_biopharma_improving_health_care_value_disease_registries/; The Boston Consulting Group, October 2012, Health Reform Should Focus on Outcomes, Not Costs https://www.bcgperspectives.com/content/articles/health_care_payors_providors_health_reform_should_focus_on_outcomes/

Using transaction data to protect consumers and merchants from fraud

What is it?

- Payment card authorization and transaction information can be used to create patterns of card use, such as purchase size, frequency and type of transaction.
- Services like Advanced Authorization from Visa can evaluate worldwide authorization data and alerts payment card issuers to potential fraudulent purchases in real time – both at checkout and at the ATM
- Payment card issuers have the ability to immediately notify the consumer of fraudulent or suspicious account activity, thereby blocking future transactions and minimizing potential losses
- Detects domestic and international fraud schemes that range from single incidents to large scale assaults

How is card data used to create value?

- Key to detecting fraud is the ability to identify patterns of card use behaviour based on past usage. For example, if a card is generally used for small, everyday purchases and a large authorization is requested for jewellery and electronics, the risk score for potential fraud is higher than usual
- A centralized network is able to instantly recall and analyse millions of pieces of information in its memory; Visa is able to identify emerging fraud trends as they happen, not hours or days later
- Issuers may decline the purchase authorization, ask to speak to the cardholder, send a text to the cardholder asking for confirmation, or monitor the account for similar out-of-pattern purchases

Scale of the impact

- An analysis of past global transactions suggests the Advanced Authorization programme could help identify US $1.5 billion in fraud around the world.
- Thousands of issuers globally utilize risk scores at time of purchase to detect fraudulent activity

Constraints/issues this raises

- Minimizing false declines, whereby the actual consumer is attempting a purchase but is declined by the issuer, is an ongoing process that must balance approvals, fraud losses and customer satisfaction
- A key constraint to the effectiveness of the risk detection is having real-time access to the cardholder's purchase activity in order to determine normal account usage patterns
- Different jurisdictions have database laws affecting financial services transactions, such as federal and state credit reporting acts

Visa Inc.
Source: www.visa.com

Exploring the opportunities and risk of using personal data in a real world context through living labs

What is it?

- Mobile Territorial Lab (MTL) is an experimental "living lab" for understanding opportunities, risks and balance between protection and exploitation of personal data
- MTL aims at creating an open infrastructure and a real community to perform experiments to understand people approaches, attitudes and feelings to user-centric Personal Data Store (PDS) paradigm
- MTL is developed in cooperation with FBK Trento, the Human Dynamics group at MIT Media Lab, the Institute ID3 and Telefonica I+D

How was personal data used to create value?

- MTL is exploring how people can take advantage of their personal data according to their wills and needs
- People can exploit data collected in their PDS by means of personal applications for life monitoring, behaviour awareness and social behaviour comparison
- People can freely decide to contribute with their personal data to research analyses and city monitoring
- MTL is investigating innovative data marketplace models enabling people in "selling " their personal data

Scale of the impact

- Outcomes of the lab will impact the definition of services on both personal data protection and its usage. These insights will be provided to Telecom Italia's residential customers
- MTL will provide feedbacks on: the experience of the involved people in controlling the protection and exploitation of their personal data through a PDS service, and on the proposed business models

Constraints/issues this raises

- Hosting PDS in the Cloud: the regulation concerning PDS service providers need to be further explored in order to understand the repartition of responsibility on data treatment between the provider and the service user, and possible constraints on the kinds of personal data to be treated by PDS services
- Privacy: issues such as the "treatment" of personal data concerning multiple persons (e.g. a phone call record or a payment receipt) must be further investigated

Telecom Italia
Source: http://www.mobileterritoriallab.eu

Using personal data vaults (lockers) can save the world 10 billion hours and improve the delivery of public and private sector services

What is it?

- Personal data vaults (also called data lockers or stores) are secure, private clouds that individuals can access via Web and mobile apps
- Benefits: they empower individuals with their data, allowing them to aggregate, store, find, securely share and get value from data about them and their lives
- One example of data reuse: Personal.com's Data Vault and Fill It app allows individuals to automatically save data, including passwords, when completing forms or registering for sites and fill out new forms in seconds

How was personal data used to create value?

- Automated form filling could:
  - Make "form filling" a thing of the past for individuals and businesses, allowing far more efficient online, mobile and in-person interactions
  - Unlock new permission-based, data-driven services from both public and private organizations
  - Address a leading security vulnerability in the online world - the use of repetitive and/or simple passwords

Scale of the impact

- Automated form filling could:
  - Save over 10 billion hours annually
  - Drive major new economic benefits for both individuals and businesses and improve the delivery of government services
  - Potentially save hundreds of millions (if not billions) of dollars in reduced security breaches
  - Personal is working with the World Economic Forum Global Agenda Council and founding partner companies to make this a reality

Constraints/issues this raises

- Most online forms have been created without such a capability in mind, making implementation and awareness harder than necessary
- Simple changes to online and mobile forms – including converting print and PDF forms to HTML – would allow them to be mapped by personal data vaults
- Data vault companies must be designed with privacy and security built in[1]: user-driven; strict and transparent permissioning to people and companies; and strong security/ encryption safeguards

Personal
Source: www.personal.com
1 This is consistent with the 7 Foundational Principles of *Privacy by Design* (See http://privacybydesign.ca/ for more details)

Using personal data to create a Professional Reputation Score helps individuals manage their professional reputations and career goals

What is it?

- Professional Reputation Scoring (created by Reputation.com) helps individuals instantly assess their reputation against others like them – and offers useful insights on how they can achieve their career goals.
- The data provides insight to users (and, upon user-approved request, to third parties) on how they compare to their cohorts (education, employment, etc.) and what they can do to improve or change trajectories.
- Reputation.com has three issued patents (and roughly 35 more in the pipeline) directly related to this technology

How was personal data used to create value?

- Personal data can be used to gain unexpected insights that guide, strengthen and effectively accelerate career and life decisions
- Personal professional data, publicly available and collected via Reputation.com's existing customer database, create a foundation for a patented scoring technology
- Professional Reputation Score is based on a variety of factors – education level, school, industry, company, trajectories, income levels, peer networks, etc.
- Individuals can evaluate where they currently stand professionally, identify steps they may take to improve their future, assess their current and prospective employers, industries and companies, and use this intelligence to take action for the future

Scale of the impact

- The product/feature is now available to millions of Reputation.com customers, users, and partner customers and users on an opt-in basis
- A reputation score makes it possible for people to see how to achieve their goals: universities successful people in their field attended, what programmes are the best, different positions to a career height, etc.
- Third party (opt-in, data vault-style) applications include: giving offers of employment, products, services, etc., to individuals with specified cohort analysis, improving Web search results based on the profiles attaching to various searchers, etc.

Constraints/issues this raises

- Scoring is only as powerful as the data foundation it rests upon so the information must constantly be updated to ensure the scoring remains meaningful
- This data is widely and publicly available, but aggregated so that individuals are not able to be discerned
- All third-party sharing is done on pre-approved, user-centric basis. Users can opt-in to a specific data vault application, retaining control over when , how and to whom the data is shared

Reputation.com
Source: www.reputation.com

Greater price transparency in the retail automotive market place has led to time savings, cost savings and efficiencies for consumers

What is it?

- Transparency around price transactions data produces value for consumers
- TrueCar creates value for the consumer by analysing data and then visually presenting it in a format that enables consumers to understand what they should be paying for a new car
- Focus on delivering 3 Ps of transparency:
  - Product: Availability and product differentiation
  - Pricing: Full disclosure of historical and current pricing to demonstrate fair pricing in the auto marketplace
  - Provider: Dealership location, reviews, differentiating features

How was personal data used to create value?

- Traditionally, pricing information from car sales has been difficult to obtain. This leads to wide variability in sales prices for the same car
- TrueCar collects and analyses individual transaction data to provide an accurate reflection of local vehicle-specific prices The presentation of this data shows car shoppers what others have paid for the same car to help them understand a fair price, based on current market conditions
- Uninformed shoppers can pay as much as 20% or more for the same exact vehicle within the same local area – online fixed pricing services prevent this disparity
- This data also helps dealers understand how to price vehicles in order to sell cars more efficiently and profitably

Scale of the impact

- Over the past 10 years, gross margins on new cars have dropped by 25%, saving US consumers over US$ 5 billion per year.
- Over the past 3 years, price ranges (low to high for the same vehicle) have narrowed by 35%. Of that range of pricing, the most disadvantaged buyers are typically the ones who pay the most[1]. Price narrowing is especially beneficial to these disadvantaged (lowest income) buyers
- Data transparency has reduced structural inefficiencies leading to:
  - Time Savings: equivalent of 60.4 lives saved per year[1] (reduction in time from negotiation[1] + contact/drive to dealers[2])
  - Cost Savings: US$ 1 billion/year (with no loss in dealer margin; from time savings with negotiation[1] + contact/drive to dealers[2])
  - Cost Efficiency Redistribution: US$ 150 billion in costs since 1999

Constraints/issues this raises

- Personal and big data are different: The responsibility when using and collecting data is contextual. Circumstances matter. How the data is used is important. Responsibilities lie with data handlers to ensure consistent and helpful applications of use (and not abuse)
- Our process:
  - We collect data from multiple sources (Personal)
  - Scrub data of personal identifiers (Personal -> Anonymous)
  - Analyse and aggregate to generate useful insights to assist consumers in their buying process (Big Data)

TrueCar

Source: http://www.truecar.com/

1 TrueCar Inc. 2012 Dealer Distance Study

2 J.D. Power and Associates, 2010 US Sales Satisfaction Index (SSI) Study

Addressing 21st century development challenges using the power of big data and real-time analytics

What is it?

- A UN initiative exploring how digital data and real-time analytics technologies can help policy-makers understand human well-being and emerging vulnerabilities in real-time
- R&D: Projects to discover new proxy indicators and analytical technologies
- “Big Data for Development” Partnerships: Forging partnerships around data, technology and analytical expertise
- Pulse Lab Network: Working with governments to pilot the approaches at country level and drive broad adoption of useful innovations, methodologies and technology tools

How was personal data used to create value?

- Since its inception in 2009, Global Pulse has been exploring utility of new digital data sources to support development goals (For example, analysis of online search data, blogs and social media chatter can help understand opinions and perceptions about issues such as unemployment, education, health, migration)
- Analysis of anonymized mobile phone data can help understand socioeconomic well-being of a community, or population movement patterns in the aftermath of a disaster or disease-outbreak

Scale of the impact

- The development of a new set of tools and techniques that allow decision-makers to harness big data to understand changes in human well-being in real time will contribute to:
- Enhanced early warning: Detection of anomalies, trends and events allows earlier response to emerging crises
- Real-time awareness: A more up-to-date picture of what a population needs and wants can lead to better, more effective programme planning and implementation
- Real-time feedback: Understanding sooner where needs are changing or are not being met will allow for rapid, adaptive course correction in development programmes

Constraints/issues this raises

- Data access: Ad hoc precedents of “Data Philanthropy”, or sharing data through non-disclosure agreements for research purposes, have been explored but mechanisms for regularly and safely sharing private sector data at an aggregated and anonymized level still are yet to be developed
- Data privacy as a human rights issue: generally accepted guidelines/parameters on ethical use of personal/digital data, and protection of privacy are yet to be developed
- Mainstreaming/operationalizing: Expensive computing power and infrastructure required to digest and process real-time big data on an ongoing basis create high barrier to entry

UN Global Pulse

Source: www.unglobalpulse.org

Using personal data to reconnect refugees with their families can significantly amplify the number of refugees helped

What is it?

- Refugees United has developed a Web platform and mobile tools that drastically streamline the refugee family tracing process for both NGOs and individuals
- Prior to Refugees United, refugee family tracing was carried out via pen and paper and without the use of mobile platforms connecting refugee organizations and refugees across conflicts and borders

How was personal data used to create value?

- Personal data is used to connect family members
- Information on village, tribe, clan, sub-clan, places family was last seen and names, ages and family-known traits are used to create a discoverable profile, keeping data access and collection in difficult settings in mind
- Refugees can register on the Refugees United search site using information such as nicknames, scars, former locations and the like that is recognizable only to family and close friends

Scale of the impact

- Refugees United is currently helping 183,000 refugees in their search for missing loved ones
- Prior to the formation of our platform, a typical NGO could assist app. 750 refugees with their tracing needs per year, a number we have many-fold amplified with numerous organizations
- The project covers the majority of East African refugee camps and urban refugee dwellings, in addition to our work in Egypt

Constraints/issues this raises

- Constraints include displaced population that may not completely understand the concept of identity, privacy, technology and the security issues revealing this may present
- With an at times strong need for anonymity among displaced communities, personal data is a challenge to utilize safely and effectively within refugee family tracing

Refugees United
Source: http://info.refunite.org/; http://www.fmreview.org/technology/kiama-et-al.html

Expanding government services and engaging citizens through open data approaches in emerging economies

What is it?

- Moldova became among the first 16 countries to launch an open data initiative in April 2011. The initiative was part of the government's efforts to modernize public sector, stimulate economic growth and improve citizens engagement
- It was supported by the World Bank's Government e-Transformation project (US$ 23 million loan)
- The government opened up public expenditures information and income declarations of public servants along with the data from various ministries
- Open data initiative was strengthened by the launch of a variety of new digital services for citizens and businesses

How was personal data used to create value?

- As of February 2013, 600 datasets were released and geo-portal was launched
- 43 agencies and ministries opened datasets on education, economics, finance, healthcare and agriculture
- The data is intended to hold government accountable in front of citizen and make its work more transparent
- Citizens are encouraged to create apps using open data
- Apps are uploaded and distributed through open data portal
- The government is now looking to use citizens' data to perform identity management, authentication, transaction authorization tasks through mobile and electronic platforms

Scale of the impact

- New services were introduced through servicii.gov.md
- A government cloud is being set up to process and store data and deliver virtual services
- Agencies are transforming the way they operate to accommodate open data standards
- Citizens are showing growing interest in government work
- Between 2010 and 2012, Moldova moved up 11 positions in the UN e-Government readiness index (from 80 to 69) and 19 positions in the World Economic Forum's networked readiness index (from 97 to 78)

Constraints/issues this raises

- Personal data privacy concerns are comparatively low among citizens
- Government agencies are very sensitive about releasing their data and often unwilling to cooperate due to security concerns, closed culture, concerns about political and economic consequences of data release
- Standards of data collection and management remain poor
- Data does not exist in digital formats
- Agencies do not have budget for open data work
- The cost of maintaining and developing open data initiatives is underestimated

Moldova Open Data Initiative
Source: http://data.gov.md ; The World Bank; The Journey of Moldova Open Government and Open Data publication (World Bank and e-Government Center of Moldova); /; www.geoportal.md

Strengthening citizen engagement in the Democratic Republic of Congo (DRC) using mobile technologies

What is it?

- Much of the DRC's infrastructure is still weak and most of the country has poor access to electricity and utilizes portable generators and charging stations
- Mobile phones are helping increase citizen participation and positively transforming the relationship between citizens and their government
- The program uses mobile phones to help citizens stay informed and engaged

How was personal data used to create value?

- Using mobile technology, citizens can:
  - Now vote on the priorities that are most pressing for their communities
  - Receive announcements on voting results, making the process more transparent and inclusive
  - Request feedback from citizens about the projects as the projects are implemented to improve accountability.

Scale of the impact

- Through mobile phones, citizens in the DRC are changing the way they engage with their governments, with their communities, and with one another
- Communities involved in this programme have seen an increase in the transfer of funds from the provincial to the local level
- Preliminary results of an external evaluation suggest a reduction in tax evasion as citizens are now more willing to pay taxes as they link government spending to improvement in the delivery of services

Constraints/issues this raises

- As the country rebuilds itself, citizens need a voice and a larger role in helping the government provide bridges, roads, electricity, and water in the places with most need as efficiently as possible
- The usage of mobile phones has been estimated to reach 47% in 2013. Also, 55% of the country's population resides in areas currently covered by mobile networks, including most rural areas from the Eastern province of South Kivu

World Bank Institute
Source: http://wbi.worldbank.org/wbi

Acknowledgments

The World Economic Forum would like to acknowledge the support of all those who contributed to this initiative in 2012-13. The global dialogue series included sessions in San Jose, USA; London, UK; Tianjin, People's Republic of China; Brussels, Belgium; Dubai, United Arab Emirates and Davos, Switzerland.

The project engaged a multistakeholder community across government, the private sector, civil society, academia and others throughout the dialogue series and through its Working Group.

Special thanks are extended to all those who supported these events and joined regular Working Group calls with their insights and collaborative spirit including: Alcatel-Lucent, Allianz SE, AT&T, BT Group Plc, Centre for Information Policy Leadership, Cisco, Ctrl-Shift, European Commission, FIS, Future of Privacy Forum, GSM Association, Health Authority of Abu Dhabi, HP, ID3 - MIT, Intel, Internet Society, Kaiser Permanente, Microsoft Corporation, MIT, Mydex CIC, Personal, Personal Data Ecosystem Consortium, Qualcomm, Refugees United, Reputation.com, Respect Network, STL Partners, SWIFT, Telecom Italia, TrueCar, UN Global Pulse, University of Washington, US Department of Treasury, US National Institute of Standards and Technology, US Office of the Director of National Intelligence, Verizon, VimpelCom, Visa Inc., WeTheData, World Bank and WPP.

Thanks also to all the organizations who provided case studies for this report.

Editorial input for this report was provided by the Steering Board members and their respective teams. Members of the Steering Board are as follows:

Robert Quinn, Senior Vice-President, Federal Regulatory and Chief Privacy Officer, AT&T

George Halvorson, Chairman and Chief Executive Officer, Kaiser Permanente

Craig Mundie, Chief Research and Strategy Officer, Microsoft Corporation

Augie K. Fabela II, Chairman and Co-founder, VimpelCom

Ellen Richey, Chief Enterprise Risk Officer, Visa Inc.

At the World Economic Forum, William Hoffman leads the Rethinking Personal Data initiative. The Boston Consulting Group (BCG) served as the project adviser in 2012 under the leadership of John Rose, David Dean, and Carl Kalapesi. Carl Kalapesi was seconded to the World Economic Forum and was the primary author of this report. Thanks also to Global Agenda Council on Data-Driven Development, especially Scott David, University of Washington, for his advice and guidance. Thanks to lowercase for their graphic design and layout and Mark Schulman for his editing support.

For more information, contact:
personaldata@weforum.org

Visit http://www.weforum.org/personaldata

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