WORLD ECONOMIC FORUM

COMMITTED TO IMPROVING THE STATE OF THE WORLD

Unlocking the Value of Personal Data: From Collection to Usage

Prepared in collaboration with The Boston Consulting Group

Background Paper for World Economic Forum Annual Meeting 2013, Davos, Switzerland

24 January 2013

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 nor volume of data shows no sign of slowing. But the 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 reinforces the need to focus on data usage. It is up to the individuals and institutions of various societies to govern and decide how to unlock the value and ensure suitable protections.

As part of the multiyear initiative Rethinking Personal Data, the World Economic Forum hosted an ongoing multi-stakeholder dialogue on personal data throughout 2012 (See Figure 2 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 global dialogue centered on a set of foundational principles that are familiar across a broad range of cultures and jurisdictions. The dialogue was based primarily on three clusters of the 1980 Organisation for Economic Co-operation and Development (OECD) Fair Information Practice Principles (FIPPs):[2]

- 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.

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, efficacy 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.

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.

[1] Big data is a collection of data sets so large and complex that they becomes difficult to process using available database management tools or traditional data-processing applications.
[2] http://oecdprivacy.org/

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 1 for more details).

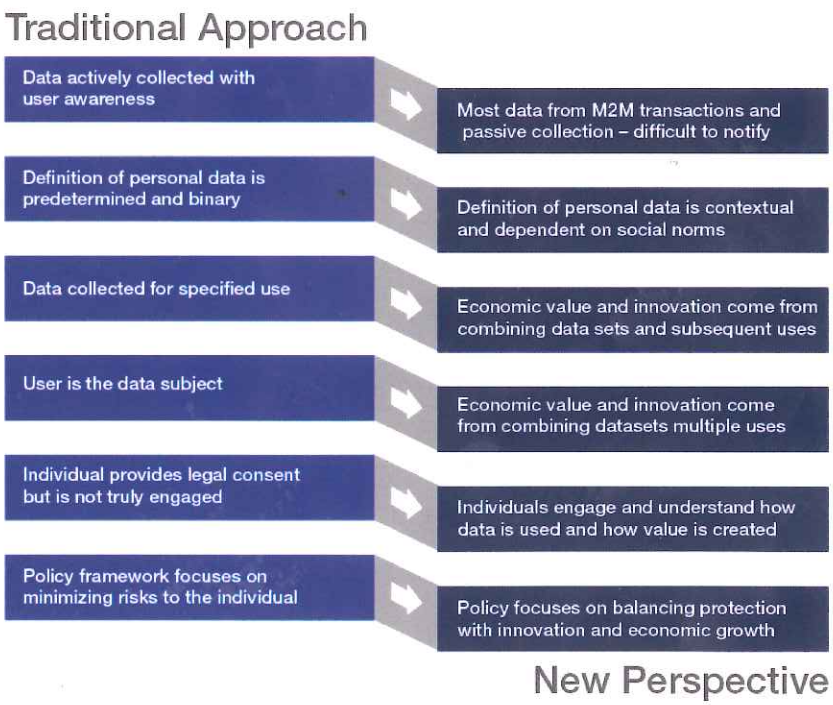
More data is being collected, processed and transferred than ever before. Data is collected by billions of connected devices, people and sensors that record trillions of transactions and behaviours each day. The unprecedented amount of data being generated is created in multiple ways. Data is actively 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, 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.

Figure 1: New Perspectives on the Use of Data



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[3] and Nike+ Fuelband[4] 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 prevent fraud and consumer identity theft. The primary purpose of collecting these data is to ensure safe and effective payment settlement. Using the data to prevent fraud creates value for all participants in the ecosystem. Scams and fraud trends are identified as they happen, not hours or days later. This results in approximately US$ 1.5 billion less in global fraud annually, and it protects consumers and merchants.

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.

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

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.[5]
- 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.

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 from thinking only about data protection to thinking also about data empowerment, and from focusing on controlling data collection to focusing on data usage, establishing appropriate permissions, controls and trustworthy data practices 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.

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

The new approach also requires a shift from focusing on protecting individuals from all possible risks to identifying plausible 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. 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.

Putting Context into Context

One of the key buzzwords in the dialogue around personal data and privacy is context. Among the phrases echoed throughout the 2012 dialogue:

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

The maturation of the discussion is requiring a re-examination and refinement of the way in which we use various common terms associated with data. 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 and valueless until it is used by someone for some purpose. 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. Research has shown that users take these elements into consideration in assessing what restrictions, consent, and notification may or may not be required.

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.

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 a combination of guiding principles, technological solutions, better evidence of what works and what does not, appropriate policy and enforceable regulation, changes in behaviour by organizations and individuals, and much more.

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. In addition, 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.

There is also an important distinction between data and information. Data is simply the raw materials. When data is extracted, combined or otherwise applied to make inferences in specific contexts, it becomes information

Principles for the Trusted Flow of Personal Data

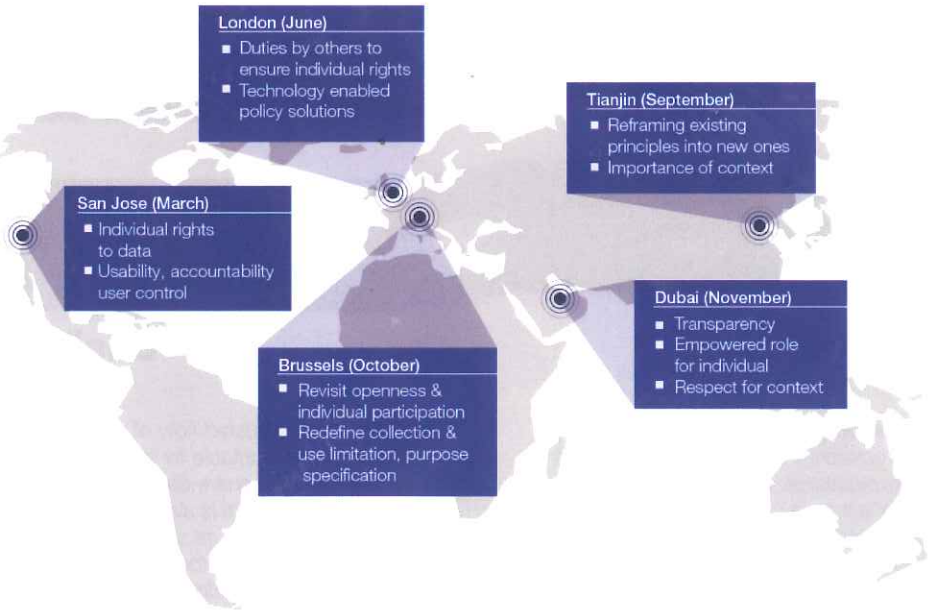
Principles have been and need to be a core part of the future governance of the personal-data ecosystem. Principles can set the foundation for trustworthy data practice and help empower users. But principles alone are not enough. Combined with technological solutions and accompanied by underlying tools 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.

Existing principles associated with the collection, handling and use of personal data (often referred to as Fair Information Practice Principles, or FIPPs) 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.

However, as discussed previously, FIPPs 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.

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 2 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.[6] 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.

Figure 2: World Economic Forum Dialogue on Principles for Trusted Flow of Personal Data



[6] 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.

This dialogue started from the broadly cited OECD FIPPS and focused on the question, "What elements of the FIPPs need to change to address current and anticipated future challenges?" There was broad consensus that change is needed to the FIPPs 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 FIPPs (See Figure 3).

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

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.

Figure 3: The World Economic Forum Dialogue Grouping of Existing Principles

Protection and Security
Security figures prominently in the original FIPPs 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 stakeholders collect, hold and use personal data. They must all take appropriate steps to secure data from accidental release, theft, unauthorized access, and misuse.

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.[7]

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

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 FIPPs 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 5)

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, efficacy and usability must lie at the heart of transparency.

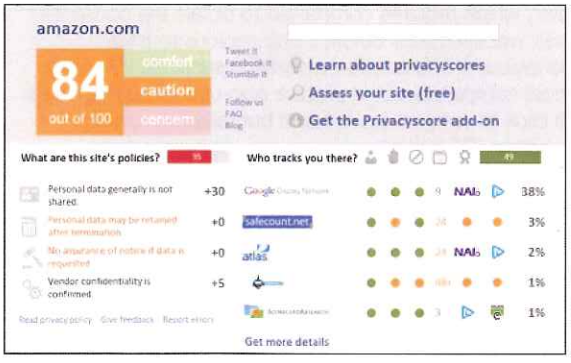
While it may be impossible to completely move away from legally derived privacy policies, there are many potential ways to help foster this shift to real understanding. There are indications that a data literacy movement is beginning to emerge in North America and Europe to help cultivate real understanding.

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,[8] 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,[9] 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 4 for an example).

Figure 4: Privacyscore Helps Individuals Understand How Different Websites Use Personal Data



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.[10]

[8] http://security.intuit.com/privacy/data-stewardship.html
[9] http://privacyscore.com/
[10] 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.

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, 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. Consider that, if the collector deletes data after the first use, the original value of such data as information will have a leverage of 1x. In the alternative scenario, the data, if not deleted, would have the potential to be further leveraged by either the original collector or others. One further use of data generates 2x value, two further uses generate 3x value, and so on. Each such use is independent and can generate independent value. Consider further that each such use of data as "information" generates further data for use by yet other third parties, thus yielding a potentially exponential rate of increase in data and information production and value. In this way, a single item of data can generate significant value, as long as it is not deleted prior to such leverage being realized.

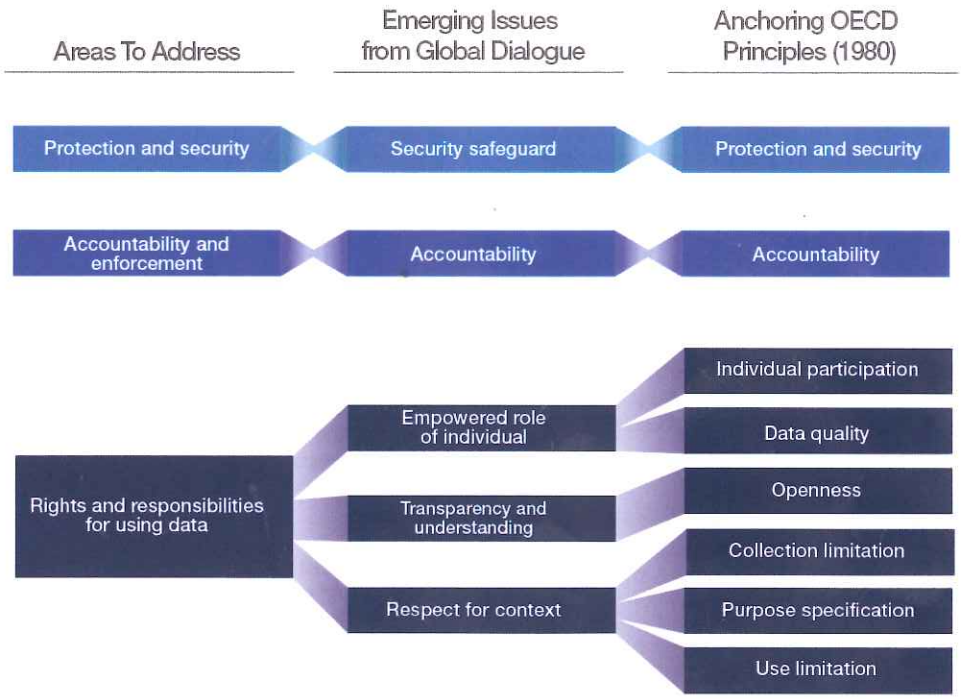
It is also important to distinguish between data being used to generate insights and discover new patterns and how they are applied to the individual. One way of dealing with this ambiguity is to more clearly identify specific risks and intrusions of concern. Once these are known, actions to manage and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

Figure 5: Areas to Focus on to Achieve Trusted Flow of Data Emerging from Dialogue Series



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.

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.

Continued Areas of Focus for the Global Dialogue

- How can the collective agreement on shared principles be best advanced?
- What constitutes "use" of personal data and what are the implications of different uses?
- How are the various uses of data most appropriately governed?
- What actions should be permitted and what should be constrained?
- By what standards and processes should these determinations be made?

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.

Acknowledgments

The World Economic Forum would like to acknowledge the support of all those who contributed to this initiative in 2012. The global dialogue series included sessions in San Jose, USA; London, UK; Tianjin, People's Republic of China; Brussels, Belgium; and Dubai, United Arab Emirates. Special thanks are extended to all those who supported these events with their insights and collaborative spirit.

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

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.

Contact

For more information, contact:

William Hoffman, Associate Director, at william.hoffman@weforum.org

Carl Kalapesi, Project Manager, at carl.kalapesi@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