November 6, 2018

These comments are pursuant to Federal Register Docket Number 180821780-8780-01 on "Developing the Administration's Approach to Consumer Privacy," and are submitted by the Information Accountability Foundation ("IAF"), a non-profit corporation whose charitable purposes are research and education related to the information economy. The IAF is the preeminent global information policy think tank that successfully works with regulatory authorities, policymakers, business leaders, civil society and other key stakeholders around the world to help frame and advance data protection law and practice through accountability-based information governance. The IAF's goal, through active consultations and research, is to achieve effective information governance systems to facilitate information-driven innovation while protecting individuals' rights to privacy and control over information about them.

The comments were prepared by Martin Abrams, Peter Cullen, Stanley Crosley, Lynn Goldstein, and Barbara Lawler. For further information contact Martin Abrams at mabrams@informationaccountability.org.

The IAF congratulates the Administration for taking a leadership role in developing this request for comments ("RFC"). Privacy is no longer just about protecting people from harm. Privacy is the practice of fair processing of data that balances all the interests pertaining to people to drive balanced use of data that supports all these interests. These comments are intended to be a useful first step in defining fair processing for the next generation.

These comments begin with general comments on privacy, the current U.S. digital environment, and direction for the fourth wave of privacy legislation.[1]

General Comments

The IAF's following comments will indicate that comprehensive privacy legislation in the United States is necessary, should achieve the ability for organizations to continue to "think and learn with data," should be interoperable with forward looking privacy regimes in other regions, and should be based on clear accountability principles.

A major differentiator between the United States and other industrialized countries is the innovation and growth that has been driven by information and communications technologies fueled by data. The velocity of innovation has been accelerating at an ever-quicker pace, from relational databases, predictive sciences, online commerce, smartphones, big data, IoT to now

[1] Privacy legislation's fourth wave is a phrase the IAF has begun using, see here. Wave one was the early global privacy legislation, including the Federal Fair Credit Reporting Act. Wave two included the EU Data Protection Directive and legislation inspired by the Directive. Phase three is the EU General Data Protection Regulation. Wave four will include future law that covers the technologies associated with the fourth industrial revolution, such as artificial intelligence (AI) and autonomous vehicles.

the fourth industrial revolution that brings digital, physical and biological spheres together. If we get privacy right, our country can accelerate prosperity in many ways including better education and health outcomes. Currently our country suffers from a digital trust deficit that exacerbates our privacy challenges. The IAF lauds the Administration's endeavor to get privacy right so that people may enjoy the fruits of prosperity along with the trust and confidence that comes with effective governance.

Privacy is a complex set of rights, interests and norms. It includes our desire to have space where we are free from the gaze of others, to have control over how we are perceived, and to be treated fairly in an information economy.

Formal privacy protections in the United States date back to 1970 with the Federal Fair Credit Reporting Act (FCRA). The FCRA is particularly interesting because it balances the shared benefits of a consumer driven economy, the freedom for companies to share their impressions about individuals, and the fairness due individuals to whom the data pertains. The primacy of the First Amendment to the U.S. Constitution is reflected in the FCRA in the ability of companies to share data that is freely observed. However, that freedom is balanced by rules related to permissible purpose and the ability to dispute accuracy.

This balancing demonstrates the dynamic equilibrium that freedom of expression has had with privacy. Free expression has made it possible for Americans to be free to observe all we can see in the public commons and to be able to robustly think creatively aided by what we have seen in that public commons. This freedom to see and think and learn with data has driven American innovation for fifty years.

It is the ability to think and learn with data that differentiates our privacy regime from every other regime in the world. In the world of advanced data analytics, thinking and learning with data is where new insights are discovered which go beyond experience and intuition and come instead from correlations among data sets. Acting with data is where insights are put into effect and where individuals may be affected as these insights are employed in an individually unique manner.

The current debate is about how we preserve privacy in an environment where seclusion, a space free from observation, is increasingly limited by technology and about how we preserve the ability to think and learn with the data that is generated by the new technologies that may infringe on seclusion. While our privacy system is different, it would be desirable for our regime to be interoperable with other regimes where privacy has as much or more weight than the ability to observe, think and learn with data.

As the Administration considers its options, it must explicitly consider the nature of the observational public commons, the ability for business and researchers to think and learn with data, and a space where individuals may have control and/or seclusion.

As stated previously, the United States pioneered privacy law with the FCRA in 1970. The FCRA's purpose was to achieve the fair processing of credit data that was necessary for a dynamic economy but created the risk of negative consequences for individuals if data was inaccurate or used inappropriately. The FCRA includes the word fair. Fair processing is a concept that is captured in the other sector specific laws for which the FCRA is the precedent.

Moreover, the FCRA created the concept of permissible purpose to process data. It is established that fair was dependent on processing creating value for people in general, even if a particular individual was harmed by accurate, negative credit information. The U.S. concept of permissible purpose created the precedent for concepts in other regions for a legitimate data use and legal basis to process.

As the Administration considers how American business might have interoperable governance, it may be useful to consider the concept of permissible purpose as applied to other data sets. This application can be done without stifling creativity or prosperity.

The IAF proposed a framework for U.S. law earlier this year that includes an accountability principle of legitimate use. That principle sits squarely in the tradition of the FCRA and the Administration's consumer privacy goals. The IAF framework is attached as appendix 1.

For the past 25 years, we have seen an acceleration in the digital economy that has necessitated the Administration's objective of developing an approach to consumer privacy that reduces fragmentation nationally and increases harmonization and interoperability nationally and globally. One of the biggest changes is the composition of the data that pertains to individuals.

The RFC references provided data, the data that individuals knowingly provide when they register at a website, provide information at point of sale, or apply for a license. The RFC does not reference three other types of data pertaining to people: "observed data" collected online and from sensors, "derived data" that is the output from a simple calculation such as ratios, and "inferred data" that is the output from algorithms. A common example of inferred data is credit scores that were introduced in the late 1980s. Inferred data is sometimes referred to as predictive data, since like credit scores, it predicts the likelihood of future outcomes. It is observed and inferred data that have both driven innovation and at the same time increased consumer angst. It is this data that will be the focus of the fourth wave of privacy legislation both here in the United States and globally as well. This taxonomy of data, "Data Origins its Implications for Governance," was written at the request of the OECD and is attached as appendix 2.

The RFC lists suggested outcomes and goals for consumer privacy in the United States. The seventh outcome is accountability. Accountability is much more than an outcome; it is the foundation to enable fair processing in a data ecosystem dominated by observed, derived and inferred data. Accountability is the means for achieving the processing of data related to people that serves people as consumers, citizens, members of communities and as owners and employees of businesses.

Accountability was first articulated as a principle in the OECD Privacy Guidelines in 1980 and was revised in 2013. Accountability as we think about it today was defined by the Global Accountability Dialogue as organized by the Centre for Information Policy Leadership as the "Essential Elements of Accountability."[2] Those essential elements were captured in "Data Protection Accountability: The Essential Elements A Document for Discussion" that is attached as appendix 3 (also known as the Galway Paper). The key accountability concepts are fairly

[2] The Global Accountability Dialogue was incorporated in 2013 as the Information Accountability Foundation. Martin Abrams, IAF Executive Director, led the project.

simple: data users should be responsible in their use of data pertaining to people and should be able to demonstrate the means for responsible use to others. U.S. business played a key role in developing the essential elements, and a number of businesses have built their privacy programs based on accountability. Accountability has been updated for the new challenges. The IAF believes that Accountable Data Stewardship is required for the application of AI, connected devices, and personalized medicine. Updated Data Stewardship essential elements are contained in appendix 4.

Consistent with the IAF's work on accountability, the IAF created a framework for discussion that mirrors the Administration's project. The IAF framework, attached as appendix 1, breaks forward looking privacy protections into four principles related to individual rights and eight principles related to accountability. These principles cover the outcomes discussed in the Administration paper and break down accountability in a more granular manner. The IAF believes that legislation should not be overly prescriptive but rather create a roadmap for using data in a trusted innovative manner.

While the U.S. should have a privacy regime that serves its needs, it is also important for the U.S. to have a privacy regime that is interoperable with its trading partners. Interoperability facilitates American global companies having a single privacy approach that may be customized on the margins for each individual country. A number of U.S. trading partners have privacy regimes that restrict data transfers to countries with equivalent laws and processes. The IAF Framework accepts that cultures differ from country to country and that a country's legal system reflects its culture. Therefore, the IAF suggests a standard that is interoperable rather than a system based on equivalency.[3] For example, the concept of legitimate processing, that is part of the IAF framework, may be found in the new Brazil privacy law, the draft law in Argentina, as well as the EU General Data Protection Regulation. Adoption in the United States would create interoperability through common assessment requirements for multi-national companies that do not require some of the provisions the IAF believes restrict thinking and learning with data.

The balance of the IAF comments will respond to the specific questions asked in the RFC.

A. What are the core privacy outcomes that consumers can expect from organizations, and does the RFC have them right?

IAF's work in digital ethics suggests that data should serve people, rather than people be the passive generator of digital content that serves others. So, the key question for the Administration is whether the outcomes suggested meet that ethical test. In many cases, the proposed outcomes are more descriptive of processes than outcomes. In answering the RFC questions, the IAF will address these issues.

B. Are there other outcomes that should be included or expanded upon, are the descriptions clear, are there risks associated with the outcomes?

Transparency is listed as an outcome. Rather than label the outcome transparency, the IAF suggests the outcome is **discoverable** or **non-secret**. Transparency is a series of activities that make it possible for individuals, and also regulators, to understand what is

[3] Many laws use the term adequacy but then define the requirement to be equivalency.

being done with data and why. Furthermore, transparency precludes organizations from collecting and processing data in secret. Transparent processes act as a check on egregious processing.

Forty-eight years of enforcing privacy provisions suggest that the outcome should be **under control** rather than individual **control.** In a highly observational world with ubiquitous but necessary sensors, it is not possible for every individual to control the knowledge created around and about them. However, under the concept of accountability, individuals may, and should, expect the organizations that collect and create data that pertains to them to do so in a fashion that is subject to controls. This expectation translates into the outcome **under control. An accountable organization will have the policies, processes, and internal oversight so data use is not being used in manner prohibited by laws and norms, is consistent with the relationship between the organization and individuals, is demonstrable to others, and is secure.** This approach does not mean there are not times where individuals should have control over their data. Instead, this approach recognizes that individual control alone is just not sufficient to govern data.

Reasonable minimization means the minimal volume of data to achieve the legitimate defined end. In some cases, minimal data necessary will be high volumes; in other cases, the minimum data necessary will be low volumes. When thinking and learning, training algorithms or deriving predictive values, the volumes will be huge. When using data to make the decisions, the volumes should be reduced. A rule of thumb is that thinking and learning with data is less impactful on the individual and therefore less risky, subject to appropriate data security. Therefore, using high volumes of data for research processes, broadly defined, is "reasonable" provided there is appropriate security. Acting with data, actually making decisions, requires just the data necessary. So right sizing the data based on legitimate uses is **reasonable minimization**.

Security The IAF generally agrees with the description of security in the RFC. However, data security also includes the concept of access control. Dynamic data obscurity addresses not only when full data sets should be available to selected users but also when particular elements may be used in a protected state to achieve "thinking with data objectives." There is precedent for this approach. Credit marketing was revolutionized in the 1970's by using de-identified data from credit bureaus at third party processors. The deidentification was achieved both through technology and policy (the FCRA). This credit marketing created competition in rural communities that had not existed in the past. In many ways, the use of pseudonymous data in digital marketing has become the modern version of that 1970's innovation.

Access and Correction are processes, not an outcome. The outcomes associated with access and correction include appropriate accuracy, individual participation and control, where appropriate, and transparent processes. The Administration description limits access and correction to provided data. As the data taxonomy, referenced earlier in these comments makes clear, access should be expanded to include observed data where it is consequential to individuals. Inferred and derived data may have proprietary rights

associated with it, so access to these types of inferred data may not be reasonable. Access to inferred data is reasonable where it is highly consequential on individuals, such as credit scores.

Access and correction are not absolute rights. For example, access may not be appropriate where the data relates to other individuals. Correction may not be appropriate where it has the effect of diminishing accuracy rather than enhancing accuracy. The oldest U.S. law that requires access and correction, the FCRA, permits consumers to request corrections but allows credit bureaus to reject the suggested corrections if the data was verified as accurate.

Risk management is a process, not an outcome. **Balanced use** (or fair processing) is the outcome desired. When society discusses risk in this context, it means the risk to all participants impacted by the processing of a specific data set for a specific purpose. Those participants may be the individual, groups of individuals, society or the organization. IAF work in ethical assessments has suggested that the first step in risk management is to identify the participants and the potential consequences, both positive and negative. Risk may only be calculated by introducing both positive and negative consequences into the equation. Once risk can be gauged, then it can be determined whether balanced use has been achieved.

Accountable rather than **accountability** is the outcome. As discussed earlier, the IAF framework contains eight principles that describe how organizations may reach the state of being accountable.

C. Are the suggested high-level goals correct? Are there other goals that should be added? Are descriptors clear? Are there risks associated with the goals?

Harmonize the regulatory landscape is a sound goal. Markets span the nation and the world. Protections for individuals should not be dependent on an advanced equation of where the person lives, where he or she goes on the web, where the data user is physically or virtually located. Federal privacy legislation should preempt state law and should be interoperable with foreign privacy regimes. Achieving this goal requires an acknowledgment that privacy law must include the basics necessary to protect individuals, achieve societal goals, and facilitate organizations' learning by thinking with data.

Legal clarity while maintaining flexibility is a very reasonable objective. Bright line rules tend to drive legal clarity, but such rules almost never anticipate future opportunities to use data to produce better outcomes. The concept of accountability drives processes that balance various consequences and that openly balance those consequences in a demonstrable manner. However, legal probability and consistency are also desirable.

Comprehensive application is sensible where there is no prejudice against new technologies and where similar processing in different industries is regulated in a similar manner.

Employ a risk and outcome-based approach that leads to user centric outcomes is an appropriate goal. Achieving this goal requires a doubling down on accountability processes. Reticence risk, blocked decision making because the organization has no means to determine whether processing is in bounds or not, creates negative consequences for all stakeholders. Sound accountability processes that include ethical processing by design, assessment processes to determine if processing is legal, fair and just, and demonstrable internal oversight to ascertain whether processes are effective not only lead to legal compliance but also to a reduction in reticence risk.

Interoperability must be an overarching goal for privacy policy. Data must flow from location to location to be used effectively. The IAF could not agree more that the Administration should seek to reduce the friction with U.S. trading partners. Doing so requires an understanding of the strengths and weaknesses of other regimes. The biggest competitive advantage the U.S. has is the relative freedom that companies operating in the U.S. have to think and learn with data. Probably the biggest weakness the U.S. has is the breadth of the public commons for data associated with individuals. The IAF suggests a regime that builds on our strengths and confronts our weaknesses.

Incentivize privacy research is a sound goal, but such research should be meaningful. One U.S. trading partner has requested research grants on improving the effectiveness of consent. Increasingly, consent is not effective in business processes such as AI where there is no human involvement. The IAF believes incentives should be enhanced for research on cascading values to code in privacy by design, on means to determine that assessments are conducted in a competent and honest fashion, and on control points for internal and external oversight.

FTC enforcement seems to be a goal based on existing legal structures rather than a greenfield look at what effective enforcement might be. The overarching goal should be **effective enforcement** and the means to that goal may well be the FTC. Global privacy enforcement is increasingly linked to organizational accountability and fairness. Evaluating accountability by its very nature has some subjectivity and means first looking at whether an organization has the ability to make sound decisions about using data related to individuals and can demonstrate it does so with integrity. Organizations with sound processes may make bad decisions. Bad decisions that lead to negative consequences may require enforcement actions. However, the nature of sanctions should be impacted by whether the consequence was systematic or not. Spanish privacy law was amended in 2011 to give the Spanish agency authority to reduce fines where violations where not systematic. Fairness is not just the opposite of unfairness as described in Section 5 of the FTC Act. Fairness is a tougher standard that goes beyond the FTC's current mandate. Enforcing subjective standards such as accountability is very difficult. Any new authority to enforce accountability should be fully described to avoid prescriptive regulation that limits the objective of fair processing that protects people while enhancing productive data uses.

Scalability is an absolute necessity. Risk is not based on the size of the organization using data but instead on the nature of the use. So, scalability means bringing the right

tools for accountability to organizations of any size that creates risk. The IAF was commissioned by the Hong Kong Privacy Commissioner for Personal Data to work with business to develop the tools for ethical assessments (attached as appendix 5: "Enhanced Data Stewardship EDIA) associated with advanced data processing such as AI. One of the most active participants was a technology incubator that works with smaller enterprises. Encouraging the development of tools for small and medium sized businesses, rather than lesser standards, is the means to achieve the scalability goal.

D. What are the appropriate next steps? Are there aspects that may be implemented or enhanced through Executive action? Should the Department convene more people? Are statutory changes needed?

The IAF believes that a comprehensive federal privacy law is needed in the United States to set the basis for the fourth industrial revolution where the digital, physical and biological spheres all come together. There is a great deal of momentum for new legislation, and it is useful to take advantage of that momentum. The IAF believes the United States should build consensus on how one achieves the full range of human interests related to an expanding information ecosystem as part of any legislative process. The IAF understands the pressure the new California Consumer Privacy Protection Act creates for national markets. There is only one opportunity to get fourth wave privacy legislation right. So, the IAF believes the Administration should continue to bring people and organizations together to explore each aspect of the desired outcomes and overarching goals. As a carpenter might say, "measure twice, or thrice, but cut once."

E. Are the definitions clear?

As mentioned above, the IAF believes some of the terms used to describe outcomes are more descriptive of processes than outcomes.

F. Are there changes or enhancements necessary for the FTC to be the primary regulator?

The information ecosystem is huge. The number of people dedicated to privacy is small in comparison to some of our major trading partners. For the FTC to be effective, the agency resources dedicated to privacy must increase. Furthermore, authority should be moved from focusing on unfair and deceptive practices to a broader responsibility to oversee organizations that are accountable for their use of data. Also, the FTC should have the means to educate the market on norms beyond enforcement actions. As stated earlier, any new powers should be carefully described to assure the appropriate balance between protection and avoiding reticence risk.

G. If the high-level goals were duplicated in other countries would it be easier for American companies to provide goods and services in those countries?

It would be easier, but it is not likely to happen. The U.S. Constitution establishes a dynamic balance between free expression and privacy. Other countries do not begin from the same starting point. However, there are commonalties between privacy regimes that could be the basis of a new regime that is interoperable with other countries. The common elements of these regimes could be attractive to other countries that are interested in thinking and learning with data.

H. Are there other ways for the U.S. to show leadership?

The United States should be very active in international fora. Key concepts related to how one might think and learn with data should be part of the agenda at APEC and at the OECD.

Thank you for the opportunity to file these comments. Please send further questions to Martin Abrams at mabrams@informationaccountability.org.

Information Accountability Comments Appendix

For Circulation August 27, 2018

Fair Processing Principles to Facilitate Privacy, Prosperity and Progress

The information ecosystem in the United States is the world's most innovative. It has not just driven economic growth, it has facilitated positive changes in all sectors. At the same time, high levels of observation along with advanced analytics have increased angst in individuals and a sense that they may be harmed by the misuse of information from them or about them. To further the discussion about a U.S. privacy regime, the Information Accountability Foundation ("IAF") puts forth these principles for a U.S. privacy framework. The framework is intended to:

- *preserve America's innovation engine,*
- *be interoperable with other new and emerging privacy regimes,*
- *protect individuals' interests in privacy, and*
- *protect all the benefits of the 21st century information age.*

While interoperable with other regimes, this framework is American in its vision and structure and is divided into two parts. The first part describes the rights necessary for individuals to function with confidence in our data driven world. The second part is focused on the obligations that organizations must honor to process and use data in a legitimate and responsible manner. While the framework outlines principles, in some cases it *includes means and outcomes to better illustrate a particular principle.*

Individual Rights

1. **Transparency** Individuals have the right to be free from secret processing of data that pertains to or will have an impact on them. Organizations should provide understandable statements about their data collection, creation, use and disclosure practices and about their policies and governance. Those statements should be directed at enforcement agencies, but they should also be publicly available. Organizations should also provide summaries and other means that make their data collection, creation, use and disclosure practices understandable to individuals.

2. **Access and Redress** As a validation there is no secret collection, creation, use or disclosure taking place and confirmation of adequate data accuracy, individuals have the right to obtain the data they provided, to understand what observational data is created

   by the organization that pertains to them, and to be told what types of data are inferred by analytical algorithms. Because intellectual property rights may prevent individuals from having the right the right to request disclosure of inferences made by the organization, and where inferences such as scores potentially have negative consequences for individuals, organizations should provide relevant explanations about their processing, appropriate opportunities for feedback, and the ability for individuals to dispute such processing.

3. **Engagement and Consent** Individuals have the right to know about data uses that are highly consequential to them, and to control those uses through an appropriate level of consent. Individuals also have the right to know that data is disclosed to third-parties beyond the context of the relationship, to request such disclosure not take place, to prohibit solicitations, and to challenge that a data use is not being undertaken in an accountable manner. Individuals have the right to object if they believe that the data about them is inaccurate or being used out of context, is not being undertaken in an accountable manner, or if they believe that uses of data are not legitimate. The right to object to processing does not pertain where data processing and use are permitted by law. Where highly consequential uses, such as health, financial standing, employment, housing and education, are governed by specific laws, those laws take priority.

4. **Beneficial Purposes** Individuals have the right to expect that organizations will process data that pertains to them in a manner that creates benefits for the individual, or if not for the individual, for a broader community of people. They also have the right to expect that data will not just serve the interests of the organization that collected the data. There may be times when objective processing does not serve the needs of each individual, but such processing does serve the broader needs of society. When this is the case, individuals may request an explanation of how processing is beneficial to the broader group. This explanation should be part of understandable summaries required under the Transparency Principle. Where there are negative consequences to individuals, individuals should expect an explanation of the results and the ability to dispute the findings, as provided in the Access and Redress Principle.

Accountable Data Stewardship

1. **Assessed and Mitigated Impacts** All collection, creating, use and disclosure of data should be compliant with all applicable laws, industry codes, and internal policies and practices, and should be subject to privacy, security and fair processing by design. Employees should receive appropriate training for their specified roles, and accountable employees should be identified to oversee privacy, security and fair processing obligations. Specifically, fair processing assessments should identify individuals and groups of individuals who are impacted, both negatively and positively, by the processing, and should guard against identifiable negative consequences. Where there

are negative consequences, organizations should mitigate those consequences to the degree possible. If unacceptable consequences still persist for some individuals or groups, the organization should document why the benefits to other individuals, groups and companies are not outweighed by the unacceptable consequences.

2. **Secure** Data should be kept secure at a level that is appropriate for the data.

3. **In Context** Data should be collected, created, used and disclosed within the context of the relationship between the individuals to whom the data pertains and the organization, based on the reasonable expectations of individuals as a group. Public safety, security and fraud prevention are considered within context.

4. **Legitimate Uses** Data should be processed only for legitimate uses that have been disclosed or are in the context of those uses, and only the data necessary for those uses should be collected, created, used or disclosed. When the data is no longer necessary for these uses, it should not be retained in an identifiable manner.

   Legitimate uses include the following:

   a. Where individuals have provided informed consent;

   b. Freely thinking and learning with data by organizations that demonstrate effective accountability, consistent with the societal objective of encouraging data driven innovation, and that honor the Onward Disclosure Responsibility Principle.

   c. Uses that create definable benefits for individuals, groups, organizations and society that are not counterbalanced by negative consequences to others, and that are based on assessments established by external criteria.

   d. Designated public purposes, including public safety and the identification and prevention of fraud, and in response to an appropriate legal request;

   e. Organizations that stand ready to demonstrate why they believe other uses not listed here that are based on assessments established by external criteria are legitimate;

   f. Where permitted by law.

5. **Accurate** Data should be accurate and appropriate for all legitimate uses, and that level of accuracy should be maintained throughout the life of the data.

6. **Onward Responsibility** Organizations that originate data should be responsible for assuring the obligations initially associated with the data are maintained when the data is disclosed to third parties. All further onward transfers should also maintain those obligations.

7. **Oversight** Organizations should monitor all uses of data to ascertain that the uses are legitimate, the data is processed fairly, the data is accurately used within the context of

the relationship with those to whom the data pertains, and processes that support individual rights and accountable data stewardship are effective and tested. The

oversight process, whether conducted by an internal body or an external agent, should be separate from and independent of those persons associated with the processing.

8. **Remediation** Organizations should stand ready to demonstrate the effectiveness of policies, practices and internal oversight to those that have external authority for oversight. Organizations should consider rectifying negative consequences where they reach a level of significant impact to individuals.

The Origins of Personal Data
And Its Implications for Governance

Martin Abrams
Executive Director and Chief Strategist
The Information Accountability Foundation
21 March 2014
Prepared for OECD Experts Roundtable
Paris, France

Executive Summary

- Legacy privacy governance regimes are based on data primarily being collected from the individual with some level of their awareness.
- Increasingly data is not collected directly from the individual but, rather, at a distance without the individual's awareness of its origination and subsequent uses.
- To understand the implication, this paper proposes a taxonomy based on the manner in which data originates. The data categories include:
    - Submitted
    - Observed
    - Derived
    - Inferred

Introduction and Purpose

Data constitutes the life blood of an information age by forming the basic building blocks of all business, government and social processes. As data growth accelerates, much of it pertains to individuals either directly or indirectly. For example, data generated by the sensors in our tires links to the vehicle which, in turn, links to the car's driver. In addition, more and more of that data is addressable by analytics processes. Those processes drive innovation and create economic and social value. They also create risks that individuals will be harmed in some tangible, inappropriate fashion, or that individual dignity will be impacted in a fashion society considers unfair. To both facilitate innovation and protect individuals, data and its uses must be governed. Governance must be effective given the true nature of data in 2014 and beyond.

The OECD documented the expansion of data and its uses in "The Evolving Privacy Landscape: 30 Years After the OEC Privacy Guidelines." The 2011 paper was published to inform the experts to make recommendations on further development of the very successful OECD Privacy Guidelines. The paper makes the case that communications and computing technologies have made more things possible, that more data flows globally, the Internet and sensors increase the amount of data, and business processes have changed to take advantage of the rapid expansion in data.

Along with the growth in data has come a fundamental change in the data itself. The computerized systems that inspired legacy privacy guidance was mostly contributed by individuals directly as those individuals participated in commerce and other facets of life. Today, more and more data originates from observations that are less obvious to the individual and are a product of processing itself. These new data will only increase as society builds out a more sensor-rich environment, and organizations make greater use of advanced analytic processes like Big Data. To get governance right, we must understand where data comes from, how it is created, and how aware and involved the individual is in its creation.

The purpose of this paper is to create a taxonomy of data based on how it first originates and tracks the policy issues that arise with new data types.[4]

Background

Collection has been the nexus for governing data since the publication of Privacy and Freedom by Alan Westin in 1967. Westin's work, along with the work of other scholars established a road map for protecting privacy when societies were in the early stages of automating information that pertains to people. The early scholarship established the contextual nature of privacy and suggested individual control the best means for governance. Early laws and guidelines put individual control in place through notifications of collection and purpose, and individual consent for the listed purposes. Further, governance guidance was designed to be supportive of the control that comes from participation in data creation. The nexus for governance would be the collection of data from the individual. The taxonomy in this paper will refer to that data type as submitted, since the individual submits the data as part of interaction with the user.

In 1967, the vast majority of the electronic data that pertained to individuals came directly from the individual's actions. The individual would apply for a loan, register a deed, open an account, apply for a license, pay a bill, or graduate from a school. All of these discrete actions would create a record that truly involved the individual. Within this setting, the actions were matched by a collection of data in which the individual participated. So, collection and origin were one in the same.

At the time, there were small observational data sets, but most were not computerized. Physicians created notes about their patients, small merchants made notes about their best customers, and early direct marketers noted similarities about their best customers. These mostly manual data sets--created without the involvement of the individual--were, for the most part, not significant enough to impact a governance model that was generally based on individual autonomy. The one exception was investigative consumer reports, in which the observations of individuals were collected as part of a report for purposes such as employment. In the United States, those reports were and still are governed by the Federal Fair Credit Reporting Act. The taxonomy will classify this category of data as observed.

---

[4] Origin is not the only lens one might use to classify data. The OECD Digital Economy Papers No. 220, "Exploring the Economics of Personal Data," contains a taxonomy of data based on the concept of data collection borrowed from the World Economic Forum[4]. The taxonomy looks at the data from a collection perspective related to a data lifecycle. The OECD paper also references Bruce Schneier's "Taxonomy of Social Networking Data" that was revised in Schneier's blog on 10 August, 2010[4]. Schneier's taxonomy does an excellent job of cataloging data from the perspective of social networking. The OECD paper also references classifications based on the nature of the relationship of the individual to the collector.

As long as there has been data that pertains to an individual, there have been others that have looked for similarities in the data. Merchants have been classifying their customers based on common attributes for as long as there has been buyers and sellers. In 19th-century North America, merchants created co-ops to share information about credit worthiness with classifications derived from shared data. The direct marketing industry began with the simple process of using transactional data to derive market segments based on look-a-likes. Furthermore, once analysts began looking for similarities, they began to conduct simple arithmetic calculations to enhance comparisons. For example, would ratios of mortgage debt to consumer debt demonstrate something interesting? The product of these simple calculations are data derived from underlying data. While the classification builds on data that comes from interactions and transactions that involve the individual, the individual is not involved in the creation of the new data. The taxonomy will classify this data as derived.

The first application of statistics against larges personal data sets was the MDS bankruptcy score in the 1980s. The MDS score made use of computerized credit reports to predict the likelihood that an individual would go bankrupt over the next five years. The MDS credit score was not just a matching of attributes of those individuals that went bankrupt but, rather, a statistically based prediction that was validated using historic data. The resulting credit score is a piece of data based on the probability of a future event taking place that is linked to an individual. While the underlying data came from interactions with the individual, the individual had no involvement in the creation of the score. The classification for this data is inferred.

Rapid Expansion of Data

The rapid increase in computing power, decrease in communications costs, and falling prices for storage all led to the expansion of data sets in the late 1980s and the 1990s. However, the most significant trigger for data expansion was the literal explosion of observational data that was sparked by the Internet in the 1990s. The Internet facilitated the collection of very granular information on how individuals behave. An observable action was no longer limited to registrations, purchases, filings but also included the micro steps that leads up to those actions. The fact that an individual paused over a pixel becomes a recordable piece of data. Much of this observational data originates in a fashion not linked to a readily identifiable individual. However, it often links to an individual in a manner that lets the non-identified individual to be characterized. So, observational data leads to the creation of both derivations such as likely responder and inferences such as 90% chance the individual is a fraudster.

The 21st century has led to sensor technologies that make granular observation possible in the physical as well as virtual world. Every major shopping mall has CCTV cameras, and images can and are transformed into data. Automobiles have sensors that read how the vehicle is operated. The combination of online and physical observation have facilitated the massive expansion of observational data. While this data begins with the actions of individuals, the individuals are not active partners in the origination itself.

Bruce McCabe published the research paper "The Future of Business Analytics" in 2007. In many ways, McCabe's paper announced the beginning of Big Data era. McCabe noted that unformatted data could now be used for analytics processes. This significantly expanded the amount of data that could be used for research, since data no longer had to be formatted in traditional fields. Diverse data sets could therefore be used to discover correlations that where less obvious in the past. Those correlations lead to predictions pertaining to individuals in almost any setting. Informatics is increasingly able to rank order individuals based on probability, which will lead to a rapid expansion of inferred data.

Taxonomy Based on Origin

In the prior section, the paper briefly described how the early work in privacy focused on the data that comes directly from the individual in a manner that involves the individual. It also discussed other forms of personal data that have a long history but only began to become impactful as technology facilitated automation. This section will begin with a table that lays out data classifications based on the manner in which the data originated.

Column 1 is the major classifications based on how the data originates.

Column 2 contains sub-classifications which help to make the analysis more granular. For example, some levels of observation are anticipated, the active sub-classification, while others are oblivious to the individual, such as the passive sub-category.

Column 3 includes examples to assist the reader in relating the categories to the data world.

Column 4 provides a simple ranking based on how aware the typical individual might be based on the distance and manner of data origination.

Table 1: Data Categories Based on Origin

Category	Sub-Category	Example	Level of Individual Awareness

Submitted	Initiated	○ Applications
		○ Registrations
		○ Public records
		  ○ Filings
		  ○ Licenses
		○ Credit card purchases
			High

	Transactional	○ Bills paid
		○ Inquiries responded to
		○ Public records
		  ○ Health
		  ○ Schools
		  ○ Courts
		○ Surveys
			High

	Posted	○ Speeches in public settings ○ Social network postings ○ Photo services	High

		○ Video sites	
Observed	Engaged	○ Cookies on a website ○ Loyalty card ○ Enabled location sensors on personal devices	Medium
	Not Anticipated	○ Data from sensor technology on my Car ○ Time paused over a pixel on the screen of a tablet	Low
	Passive	○ Facial images from CCTV ○ Obscured web technologies ○ Wi-Fi readers in buildings that establish location	Low
Derived	Computational	○ Credit ratios ○ Average purchase per visit	Medium to Low
	Notational	○ Classification based on common attributes of buyers	Medium to Low
Inferred	Statistical	○ Credit score ○ Response score ○ Fraud scores	Low
	Advanced Analytical	○ Risk of developing a disease based multi-factor analysis ○ College success score based on multi-variable big data analysis at age 9	Low

Data Category Further Description

Submitted Data

Submitted data originates via direct actions taken by the individual in which he or she is fully aware of actions that led to the data origination.

The taxonomy breaks the category into three sub-categories, initiated, transactional, and posted.

Initiated

Initiated data is the product of individuals taking an action that begins a relationship. These actions might include applying for a loan, registering to vote, taking out a license, or registering on a website. The individual is aware of the action he or she is taking. While the individual doesn't always consider the implications, it would be intuitive to the individual that his or her actions would create data that pertains to him or her.

Transactional

Transactional data is created when an individual is involved in a transaction. Transactions may include buying a product with a credit card, paying a bill, responding to a question, or taking a test. While the individual might not be thinking about the fact

that he or she is creating a record, they understand the transaction must be recorded, records need to be updated, and histories modified. The individual is an active participant in the origin of the data.

Posted

When individuals proactively express themselves, they are aware that they are creating expression that will be seen or heard by others. In past years, the recorded data might be a newspaper or television story. The growth of social networks has actively increased the origination of data based on proactive speech. While the individual is not always aware of who might see or hear the expression, they are fully involved in its creation.

Observed Data

Observed data is simply what is observed and recorded. The emergence of the Internet as an interactive consumer medium has made it possible to observe and digitalize data in a more robust manner. On the Internet, one may observe where the individual came from, what he or she looks at, how often he or she look at it, and even the length of pauses. Facial recognition and the Internet of Things is making observation in a digital manner possible in the physical world. For the purposes of this analysis, I have three sub-categories based on the level of awareness by the individual.

Engaged

Engaged observed data includes data that originates from online cookies, loyalty cards, and other instances in which the individual is made aware of the observation at some point in time. While the individual may forget that the data is being created, there is a general awareness that it is taking place. In some cases, the individual can object to or abort the creation. For example, a person may disable the Wi-Fi on their mobile device if they don't want to be observed.

Not Anticipated

Not anticipated data creation are instances in which individuals are aware that there are sensors but have little sense that the sensors are creating data that may pertain to the individual. For example, a person may be aware that there are sensors in the tires on the car and in the oil pan in the engine, but the person might not be aware that the manner in which he or she maintains the car is a data element that might pertain to them. This sub-classification would be appropriate for many of the applications related to the Internet of Things. Typical individuals would have limited awareness of this type of data.

Passive

The last sub-category is passively created observational data. An example is CCTV in public places when combined with facial recognition. It is also applicable to any situation in which it would be very difficult for individuals to be aware that they are being observed and data pertaining to the observation is being created.

Derived

Derived data is data that is simply derived in a fairly mechanical fashion from other data and becomes a new data element related to the individual. There are two sub-categories of derived data.

Computational

Computationally derived data is the creation of new data element through an arithmetic process executed on existing numeric elements. For example, a lender might create a computational data by calculating the ratio of mortgage debt to total consumer debt, an online merchant might calculate average spend per visit, or a merchant might calculate the percentage of returned items to items bought. Each of the new computational products is a data element that might be used by an organization to better understand behavior or make decisions pertaining to the individual. The individual would not typically be aware of the creation of the new data element.

Notational

Notionally derived data are new data elements created by classifying individuals as being part of a group based on common attributes shown by members of the group. For example, a marketer might notice its customers have six common attributes and look for the same attributes in a group of potential customers.

Inferred

Inferred data is the product of a probability-based analytic process. This category name is the same as that used by the World Economic Forum. This category includes two sub-categories.

Statistical

Statistically inferred data is the product of characterization based on a statistical process. Examples include credit risk scores, most fraud scores, response scores, and profitability scores. The individual is not typically involved in the development of these scores.

Advanced Analytical

Advanced analytical data are the product of advanced analytical processes such as those found in big data. These data elements are typically the product of analysis on larger and more diverse data sets, and the elements are based on analysis that is more dependent on correlation rather causation. Early examples of such data elements are identity scores that predict the likelihood that an identity is real. While credit scores were dependent on looking at past credit failures and what correlated to and impacted those failures, identity scores were based on anomalies in the manner in which identities were structured. This required a new type of analysis that had not been possible in the past.

In the medical field, Big Data is beginning to generate insights into the likelihood of future health outcomes. The individual would not be aware of the creation of these new data that are the product of the inferences that come from analysis.

Data Begets Data

Submitted and observed data comes directly from the contributions and the observations of individuals. Derived and inferred data are the products of processing other data. However, once created, derived and inferred data then become the feed stock for future data created by ongoing processing.

If one were trying to predict the growth patterns for data, one would postulate that growth in submitted data will be fairly flat. Individuals will only apply for so many loans, register at so many websites, or pay so many bills. Growth in this category would probably be in the posted sub-category as individuals submit picture and postings.

Growth in observed data should continue to accelerate as a sensor-rich environment continues to be built out. Much of that growth will be in the unexpected and passive categories, so individual participation in its creation will be minimal.

Derived data, I believe will have a flat growth curve as business processes become more robust and analysis becomes more sophisticated. In simple terms, derived data will be replaced by inferred data.

Inferred data will accelerate as more and more organizations, both public and private, increasingly take advantage of broader data sets, more computing power, and better mathematical processes.

The bottom-line is that data begets more data. That data is increasingly created at a distance from the individual and without the individual's involvement. The data tends to be the product of more sophisticated processes, and its application has more positive implications for all parties involved. The application of the data also creates new risks that the individual is not in a position to mitigate via autonomy rights.

Key Policy Questions

In 2013, the OECD updated its privacy guidelines, first adopted in 1980. Revising the guidelines, the OECD added additional guidance on accountability. The wording of the guiding principles remained fundamentally the same as adopted in 1980 and links governance to collection. This creates challenges for applying the principles to the manner in which data originates today. This section will briefly look at each of the principles and raise possible questions for the OECD to consider.

Collection Limitation

As noted in this paper, data increasingly is created not collected. Does the OECD focus on collection make the principle less useful? If one looks beyond the principle's structure, the issues raised by the principles, lawfulness and fairness, are even more relevant in the current data rich world. The principle also acknowledges that not all data originates in a manner

where consent and knowledge are applicable. The principle also points to the need for greater individual awareness. However, the structure, focused on collection, raises questions on how those underlying issues of lawfulness and fairness might be applied to the current data classes.

Data Quality

Data quality is very relevant to the current discussion. No matter how data originates it should be appropriate for its uses. Future OECD guidance pursuant to Big Data may want to explore the governance challenges related to data quality.

Purpose Specification

Purpose specification has had two objectives over the past 34 years. The first is to provide transparency to the individual about how data will be used. The second is to provide discipline to the data user about future scope of use. With data originating at a distance and without the explicit knowledge of the individual, purpose specification is less functional as a transparency tool. The second discipline, future guidance for application seems very relevant. So a question arises on how to achieve both objectives, transparency and discipline with the guidelines if not the principles.

Use Limitation

This principle raises the same issues as the previous one. Big data processes pull data into applications to both discover trends and then build applications based on the newly identified trends. Origination of new data is sometimes the byproduct of those processes. Previous work by the Big Data Project at the Centre for Information Policy Leadership discussed governance related to discovery versus application. Future OECD governance related to privacy and Big Data may want to suggest the manner in which this principle might be applied.

Security Safeguards

Data no matter how it originates should be secure proportional to the risks associated with the data. Future OECD guidance related to security safeguards might want to reiterate the importance of security safeguards as it relates to data that originates as part of analytic processes.

Openness

Openness to the creation and use of data is increasingly important. A key question is how that might be achieved. Transparency at point of collection is relatively easy compared to transparency pertaining to data processes that are not readily apparent. The author believes additional time and resources should be dedicated to increased transparency.

Individual Participation

The author believes individual participation is also very relevant to data originating at a distance from the individual. In many ways, the issues linked to individual participation are linked to the openness principle. The question isn't whether individuals should have the

right to see data and challenge underlying data but how the mechanisms to achieve the objectives of this principle might be designed.

Accountability

Accountability is the key principle in assuring governance when data originates at a distance from the individual. The additional guidance contained in the 2013 revisions are most useful. However, there is room for even more commentary on how to be accountable. Some of the commentary has been developed by privacy enforcement agencies in Canada and more recently Hong Kong.

In summary, the nature of data more dominated by observed and inferred data challenges the concept that the nexus for governance is collection and the assumption that awareness goes naturally with collection. The OECD might want to consider additional work to tie the objectives of the privacy principles to data that originates at a distance from the individual without the individual's participation and awareness.

About the Author

Martin Abrams is Executive Director of The Information Accountability Foundation. For more than 35 years, Abrams has been an information and consumer policy innovator. His most recent work has examined big data governance and privacy compliance driven by demonstrable data stewardship.

Comments

The foundation considers data classification as a work in progress. The concepts in the paper need to be tested, and the policy implications debated. Please send your comments to mabrams@informationaccountability.org.

Data Protection Accountability: The Essential Elements
A Document for Discussion
October 2009

Prepared by the Centre for Information Policy Leadership as Secretariat to the Galway Project

Data Protection Accountability: The Essential Elements
A Document for Discussion

Preface

Martin Abrams
Executive Director
Centre for Information Policy Leadership

Innovations in technology; rapid increases in data collection, analysis and use; and the global flow and access to data have made an unprecedented array of products, resources and services available to consumers. These developments, however, in no way diminish an individual's right to the secure, protected and appropriate collection and use of their information.

The manner in which those protections are provided is often challenged by the dynamic, increasingly international environment for information. The global flow of data tests existing notions of jurisdiction and cross-border co-operation. How can companies and regulators support movement of data while providing the protections guaranteed to the individual?

Accountability, a concept first established in data protection by the Organisation for Economic Co-operation and Development ("OECD"), may provide an improved approach to transborder data governance that encourages robust data flows and provides for the protection and responsible use of information, wherever it is processed. But the practical aspects of accountability, and how it can be used to address the protection of cross-border information transfers, have not been clearly articulated.

- What will be expected of companies in an accountability system?
- How will enforcement agencies monitor and measure accountability?
- How can the protection of individuals be ensured?

The Centre for Information Policy Leadership at Hunton & Williams LLP was privileged to assemble a group of international experts from government, industry and academia to consider how an accountability-based system might be designed.[5] The experts met twice to define the essential elements of accountability, examine issues raised by the adoption of the approach and propose additional work required to facilitate establishment of accountability as a practical and credible mechanism for information governance. This report, guided by a drafting committee and reviewed by the group of experts, reflects the results of those deliberations.

[5] The group of experts is listed in the Appendix.

While this paper is focused on accountability as a mechanism for global governance of data, the issue of how accountability relates to the general oversight of privacy was raised during our discussions. It may be that accountability principles can address both international as well as domestic protection of information. Our discussion recognised that the concepts of accountability that can support an improved approach already are reflected in long-standing principles of fair information practices and are inherent in current governance in Europe, Asia and North America. Making accountability a reality requires that businesses apply those concepts so that their management of information is both safe and productive. Our talks further suggested that the growing complexity of data collection and use requires that much of the burden for protecting data must shift from the individual to the organisation.

Much of what is written about accountability in this paper can be accomplished by reinterpreting existing law. It is our hope that this paper will both chart the course forward for establishing accountability-based protection and motivate stakeholders to take the important steps to do so.

The Centre is indebted to the experts who participated in this effort for generously giving of their time and expertise, and most especially to the Office of the Data Protection Commissioner of Ireland for hosting our meetings and providing us with wise guidance. While this report reflects the results of their deliberations, the Centre alone is responsible for any errors in this paper.

Executive Summary

Accountability is a well-established principle of data protection. The principle of accountability is found in known guidance such as the OECD Guidelines[6]; in the laws of the European Union ("EU"), the EU member states, Canada and the United States; and in emerging governance such as the APEC Privacy Framework and the Spanish Data Protection Agency's Joint Proposal for an International Privacy Standard. Despite its repeated recognition as a critical component of effective data protection, how accountability is demonstrated or measured has not been clearly articulated. This paper represents the results of the Galway Project — an effort initiated in January 2009 by an international group of experts from government, industry and academia to define the essential elements of accountability and consider how an accountability approach to information privacy protection would work in practice.

Accountability does not redefine privacy, nor does it replace existing law or regulation; accountable organisations must comply with existing applicable law. But accountability shifts the focus of privacy governance to an organisation's ability to demonstrate its capacity to achieve specified privacy objectives. It involves setting privacy protection goals for companies based on criteria established in law, self-regulation and best

[6] Organisation for Economic Co-operation and Development Guidelines on the Protection of Privacy and Transborder Flows of Personal Data.

practices, and vesting the organisation with both the ability and the responsibility to determine appropriate, effective measures to reach those goals. As the complexity of data collection practices, business models, vendor relationships and technological applications in many cases outstrips the individual's ability to make decisions to control the use and sharing of information through active choice, accountability requires that organisations make responsible, disciplined decisions about data use even in the absence of traditional consent.

An accountable organisation demonstrates commitment to accountability, implements data privacy policies linked to recognised external criteria, and implements mechanisms to ensure responsible decision-making about the management and protection of data. The essential elements are:

1. **Organisation commitment to accountability and adoption of internal policies consistent with external criteria.**

2. **Mechanisms to put privacy policies into effect, including tools, training and education.**

3. **Systems for internal, ongoing oversight and assurance reviews and external verification.**

4. **Transparency and mechanisms for individual participation.**

5. **Means for remediation and external enforcement.**

While many aspects of the essential elements are already established in law, self-regulation and corporate practices, some issues remain to be resolved to encourage robust adoption of an accountability approach. Policymakers and stakeholders should address questions about how accountability would work with existing legal regimes, and whether reinterpretation or amendment of existing laws might be required to make it possible to hold organisations accountable. Third-party accountability programmes have been recognised as useful in supplementing the work of government agencies. As they may play an important part in the administration of this approach, it will be necessary to clearly describe the contours of their role and the criteria by which their credibility will be assessed. Trusted movement of data based on accountability requires that privacy enforcement agencies rely upon the oversight of enforcement bodies in jurisdictions other than their own. For the approach to work effectively, stakeholders must articulate the way in which the credibility of those programmes is established and tested. Finally, small- and medium-sized enterprises that wish to demonstrate accountability will face specific challenges that must be addressed.

While additional inquiry is needed before adoption of an accountability-based approach can be realised, its promise for international privacy protection presents an opportunity to further the long-standing goal of business, regulators and advocates — robust transfer and use of data in a fashion that is responsible and protected.

Introduction

The global flow of data drives today's information economy. Innovation, efficiency and service depend on rapid and reliable access to data, irrespective of its location. Digital technologies collect and store data in ways never before imagined, and information and telecommunications networks have evolved to provide seamless, low-cost access to data around the world.

As a result, consumers have access to an unprecedented array of personalised products and services. While previously service hours ended at 5:00 p.m., the Internet enables individuals to access customer service in the middle of the night by phoning a local number that connects them to a call centre a continent away. Today, on a single server, a company can manage its email and business records for offices located in a dozen nations; travelers can rely on their debit and credit cards wherever they go; and individuals can use the Internet to download information from around the world without ever leaving their homes.

Indeed, with the increasingly global nature of data flows and the remote storage and processing of data in the "cloud", geography and national boundaries will impose few limitations on where data can be transferred but will present more practical challenges for administering and supervising global businesses.

In this environment, individuals maintain the right to the secure and protected processing and storage of their data that does not compromise their privacy. Protection must be sufficiently flexible to allow for rapidly changing technologies, business processes and consumer demand. Regulators must be equipped to articulate clear requirements for protection, educate companies and citizens, and monitor compliance in an environment in which data processing increasingly occurs outside the practical reach of most regulators, if not their legal jurisdiction.

Currently, global data flows are governed by law and guidance, which are enacted and enforced by individual countries or through regionally adopted directives or agreed-upon principles. The EU Data Protection Directive and implementing laws of member states, for example, govern the transfer of data from the European Union. The Safeguards Rule[7] imposes legal obligations on U.S. organisations to ensure that data is properly secured, wherever it is transferred or processed. And yet global data flows often challenge the way in which we have traditionally approached information protection. Daniel Weitzner and colleagues have written that information protection policy has long relied on attempts to keep information from " 'escaping' from beyond appropriate boundaries".[8] This approach is plainly inadequate in a highly connected

[7] Under the Gramm-Leach-Bliley Act, the Safeguards Rule, enforced by the Federal Trade Commission, requires financial institutions to have a security plan to protect the confidentiality and integrity of personal consumer information.

[8] Daniel J. Weitzner, Harold Abelson, Tim Berners-Lee, Joan Feigenbaum, James Hendler and Gerald Jay Sussman, "Information Accountability," *Communications of the ACM*, June 2008, at 82.

environment in which anyone armed with a cell phone or laptop has at his or her fingertips unprecedented processing power, as well as the practical ability to collect, aggregate, transfer and use personal data around the world — and in an environment in which those capabilities are growing exponentially.

Weitzner and his colleagues lead a growing multinational call for an alternative approach to securing and governing personal data based on *accountability*. An accountability-based approach to data protection requires that organisations that collect, process or otherwise use personal data take responsibility for its protection and appropriate use beyond mere legal requirements, and are accountable for any misuse of the information that is in their care.

Adoption of an accountability-based approach to governance of privacy and information in global data flows raises significant questions for business, government and individuals.

Businesses express concerns about what might be expected of them in an accountability system, how their efforts to meet those expectations will be measured and how the rules related to accountability will be defined and enforced. Privacy enforcement agencies ask how accountability might work under local law. How do enforcement agencies measure an organisation's willingness and capacity to protect information when it is no longer in the privacy protection agency's jurisdiction? How does the agency work with and trust agencies in other jurisdictions? Consumer advocates worry that accountability will lessen the individual's ability to make his own determination about appropriate use of information pertaining to him.

The Centre for Information Policy Leadership, through a process facilitated by the Office of the Irish Data Protection Commissioner, convened experts to define the essential elements of accountability; to explore the questions raised by government, business and consumers related to adoption of an accountability approach; and to suggest additional work necessary to establish accountability as a trusted mechanism for information governance.

A small group of experts met initially in January 2009 to define the contours of the inquiry and identify existing research and legal precedents involving accountability. That meeting led to a draft paper that was presented to a larger gathering in April that included data protection experts drawn from government, industry and academia from ten countries. The April meeting identified a drafting committee that oversaw the Centre staff as they prepared this document, which was then circulated for comment among all of the participants. This paper reflects the results of that process.

Accountability in Current Guidance

Accountability as a principle of data protection is not new. It was established in 1980 in the OECD Guidelines[9] and plays an increasingly important and visible role in privacy governance. The Accountability Principle places responsibility on organisations as data controllers "for complying with measures that give effect" to all of the OECD principles.

Accountability is also fundamental to privacy protection in the European Union. While not explicitly stated in the Directive, numerous provisions require that organisations implement processes that assess how much data to collect, whether the data may be appropriate for a specified purpose and the level of protection necessary to ensure that it is secure. Accountability also has featured more prominently in data governance in Europe as binding corporate rules have served as a mechanism to ensure the trusted transfer of personal data outside the EU.

The Spanish Data Protection Agency's February 2009 Joint Proposal for an International Privacy Standard includes an accountability principle that establishes a basis for data transfers based on an organisation's demonstration that it is responsible.[10]

Accountability is also the first principle in Canada's Personal Information Protection and Electronic Documents Act ("PIPEDA"), requiring that Canadian organisations put into effect the full complement of PIPEDA principles, whether the data are processed by the organisation or outside vendors, or within or outside Canada. In doing so, the accountability principle of PIPEDA establishes in law a governance mechanism for transborder data transfers.[11]

In the United States, the Federal Trade Commission ("FTC") applies to general commerce the Safeguards Rule of the Gramm-Leach-Bliley Act ("GLBA") — an accountability-based law that places obligations on a financial services organisation to ensure personal information is secured, but that does not explicitly explain how those obligations should be met.

The Asia-Pacific Economic Cooperation ("APEC") Privacy Framework includes accountability as an explicit principle,[12] basing it on the OECD language and applying it

[9] See, Organisation for Economic Co-operation and Development, *Guidelines on the Protection of Privacy and Transborder Flows of Personal Data* (1980).

[10] "Joint Proposal for a Draft of International Standards on the Protection of Privacy with Regard to the Processing of Personal Information," version 2.3, 24 February 2009.

[11] This governance was explicitly described in a 2009 publication of the Office of the Privacy Commissioner of Canada, "Processing Personal Data Across Borders: Guidelines". In PIPEDA, accountability is an overarching principle that applies to protection and management of data, whether it is maintained and processed domestically or transferred outside Canadian borders for storage and processing.

[12] For more information about the APEC Privacy Framework and a full articulation of the principles, see http://www.apec.org_media/2004_media_releases/201104_apecminsendorseprivacyfrmwk.html

to data transfers beyond national borders. The Framework states, "A personal information controller should be accountable for complying with measures that give effect to the Principles stated above." The Framework specifically requires such accountability "when personal information is to be transferred to another person or organisation, whether domestically or internationally."

Despite the inclusion of accountability in many data protection regimes, it is often unclear how companies demonstrate accountability for purposes of cross-border data transfers, how regulators measure it or why individuals should trust it.

What is an Accountability-based Approach?

An accountability-based approach to data governance is characterised by its focus on setting privacy-protection goals for organisations based on criteria established in current public policy and on allowing organisations discretion in determining appropriate measures to reach those goals. An accountability approach enables organisations to adopt methods and practices to reach those goals in a manner that best serves their business models, technologies and the requirements of their customers.

An accountability-based approach to privacy protection offers immediate advantages to individuals, institutions and regulators alike, because it recognises and is adaptable to the rapid increases in data flows.

- It will help bridge approaches across disparate regulatory systems, by allowing countries to pursue common data protection objectives through very different — but equally reliable — means. This helps to facilitate the many benefits of allowing data to move across borders, and to assure individuals a common level of data protection — even if achieved through a variety of means — irrespective of where their information is located.

- It will also heighten the confidence of individuals that their data will be protected wherever it is located and minimise their concerns about jurisdiction or local legal protections.

- It will raise the quality of data protection, by allowing use of tools that best respond to specific risks and facilitating the rapid updating of those tools to respond quickly to new business models and emerging technologies. An accountability approach requires organisations not only to take responsibility for the data they handle but also to have the ability to demonstrate that they have the systems, policies, training and other practices in place to do so.

.

- Allowing for greater flexibility will enable organisations to more effectively conserve scarce resources allocated to privacy protection. While it is essential that an accountable organisation complies with rules, resources devoted to fulfilling requirements such as notification of data protection authorities are not available for other, often more effective, protection measures. Accountability directs scarce resources towards mechanisms that most effectively provide protection for data. Organisations will adopt the tools best suited to guarantee that protections focus on reaching substantive privacy outcomes — measurable information protection goals — and to demonstrate their ability to achieve them.

Accountability does not redefine privacy, nor does it replace existing law or regulation. Accountable organisations must comply with existing applicable law, and legal mechanisms to achieve privacy goals will continue to be the concern of both regulators and organisations. However, an accountability approach shifts the focus of privacy governance to an organisation's ability to demonstrate its capacity to achieve specified objectives.

Accountability does not replace principles of individual participation and consent that have been well established in fair information practices.[9] In many cases, consumer consent to uses of data remains essential to an organisation's decisions about data management. However, in some instances obtaining such consent may be impossible or highly impractical, and an accountability approach requires that organisations make responsible, disciplined decisions about data use even in the absence of traditional consent.

How Accountability Differs from Current Approaches

Accountability is designed to provide robust protections for data while avoiding aspects of current data protection regimes that may be of limited effect or that may burden organisations without yielding commensurate benefits. Accountability allows the organisation greater flexibility to adapt its data practices to serve emerging business models and to meet consumer demand. In exchange, it requires that the organisation commit to and demonstrate its adoption of responsible policies and its implementation of systems to ensure those policies are carried out in a fashion that protects information and the individuals to which it pertains. Accountability requires an organisation to remain accountable no matter where the information is processed. Accountability relies less on

[9] Consent is found in the OECD Guidelines principle of Use Limitation, which states: "Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with Paragraph 9 except:

a) with the consent of the data subject; or

b) by the authority of law."

The principle of individual participation is also found in the OECD Guidelines, which state:

"An individual should have the right:

a) to obtain from a data controller, or otherwise, confirmation of whether or not the data controller has data relating to him;

b) to have communicated to him, data relating to him

- within a reasonable time;
- at a charge, if any, that is not excessive;
- in a reasonable manner; and
- in a form that is readily intelligible to him;

c) to be given reasons if a request made under subparagraphs (a) and (b) is denied, and to be able to challenge such denial; and

d) to challenge data relating to him and, if the challenge is successful to have the data erased, rectified, completed or amended".

the rules that exist where the data is processed and more where the obligation is first established.[13]

Accountability relies less on specific rules but instead requires that organisations adopt policies that align with external criteria found in law — generally accepted principles or industry best practices — and foster a level of data protection commensurate with the risks to individuals raised by loss or inappropriate use of data. The accountable organisation complies with applicable law and then takes the further step to implement a programme that ensures the privacy and protection of data based on an assessment of the risks to individuals raised by its use. These risks should be assessed and measured based on guidance from regulators, advocates, individuals and other members of industry. Ultimately, regulators are responsible for ensuring that the risks to the data have been managed appropriately.

While the individual continues to play an important role in protecting his or her information, accountability shifts the primary responsibility for data protection from the individual to the organisation collecting and using data. Much of United States law, for example, is based on disclosure of the organisation's privacy policy, notification of individuals and obtaining their consent to specific uses of data. This approach is designed to enhance individual control over the manner in which data is used. Individuals are vested with responsibility for determining the manner in which their data is used and shared; organisations are obligated to provide the individual with sufficient information on which to base an informed choice.

In the U.S. the Federal Trade Commission is authorised to bring an enforcement action based on the organisation's notice when an organisation acts in an unfair or deceptive manner with respect to its privacy practices. In the absence of, and in some cases even with, an overarching privacy law, the individual is charged with policing the

[13] When, however, information security rules where data are processed are stronger than where the security obligation was incurred, they may indeed apply.

marketplace for privacy, by familiarising him- or herself with every organisation's policy and making a decision based on that information whether or not the organisation is trustworthy and using data in an appropriate manner.

Accountability does not displace the individual's ability to assert his rights, but relieves him of much of the burden of policing the marketplace for enterprises using data irresponsibly. Faced with rapid advances in data analytics and increasingly complex technologies, business models and vendor relationships, consumers find it increasingly difficult to make well-informed privacy decisions, even when they can access privacy policies. Accountability demands responsible, appropriate data use whether or not a consumer has consented to one particular use or another.

Accountability does not wait for a system failure; rather, it requires that organisations be prepared to demonstrate upon request by the proper authorities that it is securing and protecting data in accordance with the essential elements.

Enforcement of binding corporate rules ("BCRs") or the cross-border privacy rules as defined in APEC perhaps most closely approximate an accountability approach to information management and protection. BCRs, which are more fully developed, provide a legal basis for international data flows within a corporation or a group of organisations when other options are either impracticable or of limited utility. BCRs are a set of rules, backed by an implementation strategy, adopted within a company or corporate group that provides legally binding protections for data processing within the company or group. While the Directive and national laws that implement it rely on adequacy of laws and enforcement in a particular legal jurisdiction outside the EU, BCRs allow companies to write rules for data transfer that are linked to the laws where data was collected rather than look to compliance with the law of a particular geographic location where the data may be processed. Data authorities examine whether an organisation's binding rules export local European law with the data, and can determine whether its data practices and protections can be trusted to put those rules into effect — that it has in place the procedures, policies and mechanisms necessary to meet the obligations established in the BCR and to monitor and ensure compliance.[14]

Essential Elements of Accountability

An accountable organisation demonstrates commitment to accountability, implements data privacy policies linked to recognised outside criteria, and establishes performance mechanisms to ensure responsible decision-making about the management of data consistent with organisation policies. The essential elements articulate the conditions that must exist in order that an organisation establish, demonstrate and test its accountability.

[14] BCRs cover only governance of data originating in the European Union. They do not apply to data originating from other regions.

It is against these elements that an organisation's accountability is measured.

The essential elements are:

1. **Organisation commitment to accountability and adoption of internal policies consistent with external criteria.**

   An organisation must demonstrate its willingness and capacity to be both responsible and answerable for its data practices. An organisation must implement policies linked to appropriate external criteria (found in law, generally accepted principles or industry best practices) and designed to provide the individual with effective privacy protection, deploy mechanisms to act on those policies, and monitor those mechanisms. Those policies and the plans to put them into effect must be approved at the highest level of the organisation, and performance against those plans at all levels of the organisation must be visible to senior management. Commitment ensures that implementation of policies will not be subordinated to other organisation priorities. An organisational structure must demonstrate this commitment by tasking appropriate staff with implementing the policies and overseeing those activities.

   Many global organisations have established policies in accordance with accepted external criteria such as the EU Directive, OECD Guidelines or APEC Principles. These companies demonstrate high-level commitment to those policies and the internal practices that implement them by requiring their review and endorsement by members of the organisation's executive committee or board of directors.

2. **Mechanisms to put privacy policies into effect, including tools, training and education.**

   The organisation must establish performance mechanisms to implement the stated privacy policies. The mechanisms might include tools to facilitate decision making about appropriate data use and protection, training about how to use those tools, and processes to assure compliance for employees who collect, process and protect information. The tools and training must be mandatory for those key individuals involved in the collection and deployment of personal information. Accountable organisations must build privacy into all business processes that collect, use or manage personal information.

   Organisations in Europe, North America and Asia-Pacific have implemented comprehensive privacy programmes that incorporate personnel training, privacy impact assessments and oversight. In some cases, organisations have automated processes and integrated responsibility for programme obligations into all levels and across all aspects of the enterprise, while responsibility for compliance, policy development and oversight remains in the privacy office.

3. **Systems for internal ongoing oversight and assurance reviews and external verification.**

Using risk management analysis, enterprises that collect and use personal information must monitor and measure whether the policies they have adopted and implemented effectively manage, protect and secure the data. Accountable organisations establish these performance-monitoring systems based on their own business cultures. Performance systems evaluate an organisation's decisions about data across the data life cycle — from its collection, to its use for a particular application, to its transmission across borders, to its destruction when it is no longer useful — and must be subject to some form of monitoring.[15]

The organisation should establish programmes to ensure that the mechanisms are used appropriately as employees make decisions about the management of information, system security and movement of data throughout the organisation and to outside vendors and independent third parties.

The organisation should also periodically engage or be engaged by the appropriate independent entity to verify and demonstrate that it meets the requirements of accountability. Where appropriate, the organisation can enlist the services of its internal audit department to perform this function so long as the auditors report to an entity independent of the organisation being audited. Such verification could also include assessments by privacy enforcement or third-party accountability agents. The results of such assessments and any risks that might be discovered can be reported to the appropriate entity within the organisation that would take responsibility for their resolution. External verification must be both trustworthy and affordable. Privacy officers may work with their audit departments to ensure that internal audits are among the tools available to oversee the organisation's data management. Organisations may also engage firms to conduct formal external audits. Seal programmes[13] in Europe, North America and Asia-Pacific also provide external oversight by making assurance and verification reviews a requirement for participating organisations.

4. Transparency and mechanisms for individual participation.

To facilitate individual participation, the organisation's procedures must be transparent. Articulation of the organisation's information procedures and protections in a posted privacy notice remains key to individual engagement. The accountable organisation develops a strategy for prominently communicating to individuals the most important information. Successful communications provide sufficient transparency such that the individual understands an organisation's data practices as he or she requires. The accountable organisation may promote transparency through privacy notices, icons, videos and other mechanisms.

[15] Accountable organisations have traditionally established performance systems based on their own business culture. Successful performance systems share several characteristics:

- they are consistent with the organisation's culture and are integrated into business processes;

When appropriate, the information in the privacy notice can form the basis for the consumer's consent or choice. While the accountability approach anticipates situations in which consent and choice may not be possible, it also

- they assess risk across the entire data life cycle;
- they include training, decision tools and monitoring;
- they apply to outside vendors and other third parties to assure that the obligations that come with personal data are met no matter where data is processed;
- they allocate resources where the risk to individuals is greatest; and
- they are a function of an organisation's policies and commitment.

[13] Seal programmes are online third party accountability agents.

provides for those instances when it is feasible. In such cases it should be made available to the consumer and should form the basis for the organisation's decisions about data use.

Individuals should have the ability to see the data or types of data that the organisation collects, to stop the collection and use of that data in cases when it may be inappropriate, and to correct it when it is inaccurate. There may be some circumstances, however, in which sound public policy reasons limit that disclosure.

5. Means for remediation and external enforcement.

The organisation should establish a privacy policy that includes a means to address harm[16] to individuals caused by failure of internal policies and practices. When harm occurs due to a failure of an organisation's privacy practices or to a lapse in its compliance with its internal policies, individuals should have access to a recourse mechanism. In the first instance, the organisation should identify an individual to serve as the first point of contact for resolution of disputes and establish a process by which those complaints are reviewed and addressed.

The accountable organisation may also wish to engage the services of an outside remediation service to assist in addressing and resolving consumer complaints. Third-party agents, including seal programmes and dispute resolution services, can facilitate the consumer's interaction with the organisation and enhance its reputation for complying with its policies and meeting its obligations to individuals.

Accountability practices should be subject to the legal actions of the entity or agency with the appropriate enforcement authority. Ultimate oversight of the accountable organisation should rest with the appropriate local legal authority. The nature of that

[16] The concept of harm can include, among other things, compromise of an individual's financial or physical well-being; embarrassment; and damage to reputation. Additional work is needed to more clearly define and describe harm as it can result from violation of privacy and inappropriate use of data.

authority may vary across jurisdictions. However, it is critical that the accountable organisation recognise and respond to the legal authority exercising proper jurisdiction.

Public Policy Issues

While many aspects of the essential elements are already well established in law, selfregulation and corporate practices, consideration of several issues could usefully assist and stimulate the robust adoption of an accountability approach. These include the following:

1. **How does accountability work in currently existing legal regimes?**

   Adopting an accountability approach to global information privacy governance may require reinterpretation or amendment of existing laws to enable the use of accountability mechanisms and to make it easier and more practicable to hold organisations accountable.[17]

   It may, for example, be necessary to provide in law or regulation that organisations comply with requests to inspect or review certain privacy practices to determine whether the organisation meets the essential elements of accountability as discussed in this paper. Work may be required to provide for legal recognition of the internal rules and policies organisations adopt and the measures organisations take to be accountable.[18]

2. **What is the role of third-party accountability agents?**

   Third-party review of an organisation's practices against appropriate criteria will greatly facilitate the success of an accountability approach. Qualified, authorised accountability agents will be an important element to address resource constraints in order to make the accountability approach work in practice.

   Establishing criteria for organisations that wish to serve as accountability agents, and articulating their role and the extent of their authority, will be a

[17] In its 2008 report the Australian Law Reform Commission considered the possibility that Australian law be amended to assure an accountability approach could be used to improve governance of cross-border data transfers. A number of EU countries are exploring whether amending the law could better accommodate binding corporate rules.

[18] Such amendments are suggested in the APEC Privacy Framework, which requires that organisations comply with local data protection rules, but those amendments must enable them to write cross-border privacy rules that link to the APEC Principles to govern data transfers. Paragraph 46 of the Framework commentary encourages member economies to "endeavor to support the development and recognition or acceptance of organizations' cross-border privacy rules across the APEC region, recognizing that organizations would still be responsible for complying with the local data protection requirements, as well as with applicable laws".

key task for policymakers. It will also be necessary to determine ways to ensure that accountability agents are worthy of public trust, and to develop the criteria by which they can be judged. Such criteria would ideally be developed through a consultative process that includes businesses, government representatives, experts and advocates.

Finally, to be useful to organisations, the services of an accountability agent must be affordable from a financial and operations perspective. Accountability agents must be able to price their services in a manner that allows them to recover their cost and build working capital, but still ensure that services are affordable to the full range of organisations that wish to avail themselves of their resources. Certification processes should be meaningful and trustworthy.

They should also be designed to limit their disruption of business operations and to safeguard the confidentiality of an organisation's data assets.

3. How do regulators and accountability agents measure accountability?

An accountability approach does not rely on a breach to prompt review of an organisation's information practices and protections. Accountability agents and regulators must be empowered to review organisations' internal processes in a manner that allows them to ensure meaningful oversight. Policymakers may also wish to consider the measures to be taken by organisations to test for accountability and to be sure that it is working.

While an organisation's corporate policies must be linked to external criteria in the various countries where it does business, laws may differ from jurisdiction to jurisdiction. Accountability oversight must assess an organisation's overall privacy programme and allow for resolution of those differences in company policies in a manner that furthers the intent of a range of often conflicting laws or regulations.

Policymakers need to identify a way to measure confidence in an organisation's overall privacy accountability programme — commitment, policies and performance mechanisms — to determine whether an organisation is accountable even if its policies and practices are not a one-toone match for local law and regulation.

4. How is the credibility of enforcement bodies and third-party accountability programmes established?

Trusted movement of data based on accountability requires that privacy enforcement agencies rely upon the oversight of enforcement bodies in jurisdictions other than their own. Assessing accountability requires examining and judging an organisation's entire programme — a somewhat

subjective analysis — so that the credibility of accountability agents is critical.[19]

Third-party accountability programmes such as seal programmes may supplement the work of government agencies. The credibility of these third parties must also be established if they are to be trusted by privacy enforcement agencies and the public. Investment in robust process and experienced, thoughtful staff will be essential to their success.

Additional work should be undertaken to determine how the credibility of these organisations is tested. It will be necessary to determine ways to ensure that accountability agents are worthy of public trust, and to develop the criteria by which they can be judged. Such criteria would ideally be developed through a consultative process that includes businesses, government representatives, experts and advocates.

5. What are the special considerations that apply to small- and medium sized enterprises that wish to demonstrate accountability, and how can they be addressed?

In many cases, organisations that wish to demonstrate accountability may be small- and medium-sized enterprises, ("SMEs") for which privacy protection resources may be limited. Consideration must be given to the special needs of these organisations and the impact that fulfilling the essential element may have on these enterprises. It may be that aspects of the essential elements will need to be tailored or adapted for smaller organisations in a manner that makes them more workable but does not dilute them.

Assessment requirements provide one example. While assessments may well serve the same function for SMEs as they do for larger organisations, such assessments may pose an undue burden on smaller enterprises with scarce resources. The nature of the assessment and the parties that may carry them out may differ for such entities, depending on the nature and sensitivity of the data in question. It will be important to examine how an SME might fulfill the assessment requirement without compromising itself financially. Similar questions of scalability as they apply to these organisations will need to be considered and resolved.

Conclusion

Dramatic advances in the speed, volume and complexity of data flows across national borders challenge existing models of data protection. In the face of such complexity and

[19] Work already undertaken at the OECD may be helpful in this regard. See Organisation for Economic Cooperation and Development, *Recommendations on Cross-border Co-operation in the Enforcement of Laws Protecting Privacy* (2007).

rapid change, data protection must be robust, yet flexible. Privacy can no longer be guaranteed either through privacy notices and consent opportunities for individuals, or through direct regulatory oversight.

An accountability-based approach to data protection helps to address these concerns. It requires that organisations that collect, process or otherwise use personal information take responsibility for its protection and appropriate use beyond mere legal requirements, and that they be accountable for any misuse of the information that is in their care.

Accountability does not redefine privacy, nor does it replace existing law or regulation. While mechanisms to achieve privacy goals will remain the concern of both policymakers and organisations, an accountability approach shifts the focus of privacy governance to an organisation's ability to achieve fundamental data protection goals and to demonstrate that capability.

While there is already a greater focus on accountability in recent data protection enactments and discussion, and much can be accomplished within existing frameworks, there is also a growing awareness that organisations that use personal data need to put in place and ensure compliance with the five essential elements of accountability:

(1) Organisation commitment to accountability and adoption of internal policies consistent with external criteria;

(2) Mechanisms to put privacy policies into effect, including tools, training and education;

(3) Systems for internal, ongoing oversight and assurance reviews and external verification;

(4) Transparency and mechanisms for individual participation; and

(5) Means for remediation and external enforcement.

The path forward is clear, if at times daunting. The promise of an accountability-based approach to international privacy protection presents an opportunity to further the longstanding goal of business, regulators and advocates alike — robust transfer and use of data in a fashion that is responsible and that ensures meaningful protections for individuals. To realise this goal, policymakers and the leaders of organisations must undertake the challenging and necessary work towards greater emphasis on true accountability.

Appendix
Galway Project Participants

The following lists the participants in the Galway Project. This list indicates participation in the Galway Project deliberations only and does not imply endorsement of the contents of this document.

Joseph Alhadeff, Oracle Corporation

Rosa Barcelo, Office of the European Data Protection Supervisor

Jennifer Barrett, Acxiom Corporation

Marcus Belke, 2B Advice

Bojana Bellamy, Accenture

Daniel Burton, Salesforce.com

Emma Butler, Information Commissioner's Office, United Kingdom

Fred Cate, Indiana University, Maurer School of Law

Maureen Cooney, TRUSTe

Peter Cullen, Microsoft Corporation

Gary Davis, Office of the Data Protection Commissioner, Ireland

Elizabeth Denham, Office of the Privacy Commissioner, Canada

Michael Donohue, Organisation for Economic Co-operation and Development

Lindsey Finch, Salesforce.com

Giusella Finocchiaro, University of Bologna

Rafael Garcia Gozalo, Data Protection Agency, Spain

Connie Graham, Procter & Gamble Company

Billy Hawkes, Data Protection Commissioner, Ireland

David Hoffman, Intel Corporation

Jane Horvath, Google

Gus Hosein, Privacy International

Peter Hustinx, European Data Protection Supervisor

Takayuki Kato, Consumer Affairs Agency, Japan

Christopher Kuner, The Centre for Information Policy Leadership, Hunton & Williams LLP

Barbara Lawler, Intuit, Inc.

Artemi Rallo Lombarte, Data Protection Commissioner, Spain

Rocco Panetta, Panetta & Associates

Daniel Pradelles, Hewlett Packard Company

Florence Raynal, CNIL

Stéphanie Regnie, CNIL

Manuela Siano, Data Protection Authority, Italy

David Smith, Information Commissioner's Office, United Kingdom

Hugh Stevenson, United States Federal Trade Commission

Scott Taylor, Hewlett Packard Company

Bridget Treacy, The Centre for Information Policy Leadership, Hunton & Williams LLP

K. Krasnow Waterman, Massachusetts Institute of Technology

Armgard von Reden, IBM Corporation

Jonathan Weeks, Intel Corporation

Martin Abrams, The Centre for Information Policy Leadership, Hunton & Williams LLP

Paula J. Bruening, The Centre for Information Policy Leadership, Hunton & Williams LLP

THE CENTRE
FOR INFORMATION
POLICY LEADERSHIP
HUNTON & WILLIAMS LLP

© 2009 The Centre for Information Policy Leadership LLP. The content of this paper is strictly the view of the Centre for Information Policy Leadership and does not represent the opinion of either its individual members or Hunton & Williams LLP. The Centre does not provide legal advice. These materials have been prepared for informational purposes only and are not legal advice, nor is this information intended to create an attorney-client or similar relationship. Whether you need legal services and which lawyer you select are important decisions that should not be based solely upon these materials. Please do not send us confidential information. Visit us at www.informationpolicycentre.com.

October 8, 2018

The Essential Elements of Accountability were developed by a multi-stakeholder group that met in Dublin Ireland as the Global Accountability Dialogue. The Essential Elements provided granularity for the OECD Accountability Principle. It is the basis for the privacy accountability movement that led to new regulatory guidance and new approaches to law. The Information Accountability Foundation was the incorporation of the Global Accountability Dialogue.

The Essential Elements are still key for building an accountability-based data protection or privacy program. When organisations mature beyond core processing activities to uses beyond common understanding such as advanced analytics, the Internet of Things, artificial intelligence and advanced analytics, governance needs to move from being a data custodian to a data steward. To facilitate that change the IAF developed "Enhanced Data Stewardship Accountability Elements."

To be able to transform data into information and information into knowledge and insight and knowledge into competitive advantage, for individuals to be able to trust data processing activities that might not be within their expectations, enhanced Data Stewardship Accountability (Enhanced Accountability) is needed.

The table below compares to the 2009 Essential Elements and the 2018 Enhanced Data Stewardship Elements

Core Element	**Essential Elements of Accountability**	**Enhanced Data Stewardship Accountability Elements**
Internal Policies	Organisation commitment to accountability and adoption of internal policies consistent with external criteria.	**Organizations should define data stewardship values, and then translate into organizational policies and processes for ethical data processing.**

	An organisation must demonstrate its willingness and capacity to be both responsible and answerable for its data practices. An organisation must implement policies linked to appropriate external criteria (found in law, generally accepted principles or industry best practices) and designed to provide the individual with effective privacy protection, deploy mechanisms to act on those policies, and monitor those mechanisms. Those policies and the plans to put them into effect must be approved at the highest level of the organisation, and performance against those plans at all levels of the organisation must be visible to senior management. Commitment ensures that implementation of policies will not be subordinated to other organisation priorities. An organisational structure must demonstrate this commitment by tasking appropriate staff with implementing the policies and overseeing those activities. Many global organisations have established policies in accordance with accepted external criteria such as the EU Directive, OECD Guidelines or APEC Principles. These companies demonstrate high-level commitment to those policies and the internal practices that implement them by requiring their review and endorsement by members of the organisation's executive committee or board of directors.	As a matter of organizational commitment, organizations should define data stewardship values that are reduced to guiding principles and then are translated into organizational policies and processes for ethical data processing**.** a) These values and principles should be organizationally derived and should not be restatements of law or regulation. They may go beyond what the law requires, but at a minimum, they should be aligned, and not be inconsistent, with existing laws, regulations or formal codes of conduct.[20] Organizations should be open about their values and principles. b) Organizational policies and processes derived from these values should be anchored to clearly defined, accountable individuals within the organization and should be overseen by designated senior executives. c) The organization's data stewardship guiding[21] principles should be easily understood by all staff, and in particular by technical staff, and should be capable of being programmed into activity objectives.
Mechanisms and Assessments	Mechanisms to put privacy policies into effect, including tools, training and education.	**Organizations should use an "ethics by design" process to translate their data stewardship values into their data analytics and data use system design processes**

[20] Examples of existing professional or industry codes of conduct are those that relate to AI or ML. These Elements should work with those codes and not replace them.

[21] See IAF Blog – The Need for an Ethical Framework. http://informationaccountability.org/the-need-for-an-ethical-framework

	The organisation must establish performance mechanisms to implement the stated privacy policies. The mechanisms might include tools to facilitate decision making about appropriate data use and protection, training about how to use those tools, and processes to assure compliance for employees who collect, process and protect information. The tools and training must be mandatory for those key individuals involved in the collection and deployment of personal information. Accountable organisations must build privacy into all business processes that collect, use or manage personal information. Organisations in Europe, North America and Asia-Pacific have implemented comprehensive privacy programmes that incorporate personnel training, privacy impact assessments and oversight. In some cases, organisations have automated processes and integrated responsibility for programme obligations into all levels and across all aspects of the enterprise, while responsibility for compliance, policy development and oversight remain in the privacy office.	Organizations should use an "ethics by design" process to translate their data stewardship values into their data analytics and data use system design processes so that society, groups of individuals, or individuals themselves, and not just the organizations, gain value from the data processing activities, such as AI or ML. Advanced data processing activities, such as AI and ML, that affect individuals should have beneficial impacts accruing to individuals and communities of individuals, particularly those to whom the underlying data pertains. a) Where an analytical data driven use has potential impact at the individual level, or at a higher level, such as groups of individuals and society, the risks and benefits should be explicitly defined. The risks should be necessary and proportional to the benefits and should be mitigated to the extent possible. b) The systems, and the data that feeds those systems, should be assessed for appropriateness based on the decision the data is being used for and should be protected proportional to the risks. c) Where appropriate, organizations should follow codes of conduct that standardize processes to industry norms. d) Ethical Data Impact Assessments (EDIAs)[22] should be required when advanced data analytics may be impactful on people in a significant manner and/or when data enabled decisions are being made without the intervention of people.

[22] See here for A Model EDIA.

		1. An EDIA is a process that looks at the full range of benefits, risks, rights, obligations and interests of all individuals, groups of individuals, society and other data stakeholders such as regulators. 2. An EDIA is a means of determining whether an instance of processing is in accordance with the data stewardship values and guiding principles established by the organization. Processing includes all steps necessary to achieve an outcome, from the collection of data through the implementation of data driven outcomes. 3. Organizations should have EDIAs that achieve an "ethics by design" process that is integrated into systems development. e) All staff involved in data impacting processing should receive training so that they may competently participate in an "ethics by design" process.
<u>Oversight</u>	Systems for internal ongoing oversight and assurance reviews and external verification. Using risk management analysis, enterprises that collect and use personal information must monitor and measure whether the policies they have adopted and implemented effectively manage, protect and secure the data. Accountable organisations establish these performance-monitoring systems based on their own business cultures. Performance	**There should be an internal review process that assesses whether EDIAs have been conducted with integrity and competency** There should be an internal review process that assesses whether EDIAs have been conducted with integrity and competency, if the issues raised as part of

	systems evaluate an organisation's decisions about data across the data life cycle — from its collection, to its use for a particular application, to its transmission across borders, to its destruction when it is no longer useful — and must be subject to some form of monitoring. The organisation should establish programmes to ensure that the mechanisms are used appropriately as employees make decisions about the management of information, system security and movement of data throughout the organisation and to the outside vendors and independent third parties. The organisation should also periodically engage or be engaged by the appropriate independent entity to verify and demonstrate that it meets the requirements of accountability. Where appropriate, the organisation can enlist the services of its internal audit department to perform this function so long as the auditors report to an entity independent of the organisation being audited. Such verification could also include assessments by privacy enforcement or third-party accountability agents. The results of such assessments and any risks that might be discovered can be reported to the appropriate entity within the organisation that would take responsibility for their resolution. External verification must be both trustworthy and affordable. Privacy officers may work with their audit departments to ensure that internal audits are among the tools available to	the EDIA have been resolved, and if the advanced data processing activities are conducted as planned.[23] a) Where data processes begin with analytic insights, those insights should be tested for accuracy, predictability, and consistency with organizational values. b) Intensive data impacting systems should be reviewed so that outcomes are as intended with the objectives of the activity, risks are mitigated as planned, harms are reduced, and unintended consequences are understood. c) Where internal reviewers need external expertise, that expertise should be sought. d) The review of the EDIA process is separate and independent from the EDIA, ethical data impact assessment, process.

[23] See here for A Model Oversight Assessment.

		oversee the organisation's data management. Organisations may also engage firms to conduct formal external audits. Seal programmes13 in Europe, North America and Asia-Pacific also provide external oversight by making assurance and verification reviews a requirement for participating organisations.	

Individual Participation and Engagement	Transparency and mechanisms for individual participation. To facilitate individual participation, the organisation's procedures must be transparent. Articulation of the organisation's information procedures and protections in a posted privacy notice remains key to individual engagement. The accountable organisation develops a strategy for prominently communicating to individuals the most important information. Successful communications provide sufficient transparency such that the individual understands an organisation's data practices as he or she requires. The accountable organisation may promote transparency through privacy notices, icons, videos and other mechanisms. When appropriate, the information in the privacy notice can form the basis for the consumer's consent or choice. While the accountability approach anticipates situations in which consent and choice may not be possible, it also provides for those instances when it is feasible. In such cases it should be made available to the consumer and should form the basis for the organisation's decisions about data use. Individuals should have the ability to see the data or types of data that the organisation collects, to stop the collection and use of that data in cases when it may be inappropriate, and to correct it when it is inaccurate. There may be some circumstances, however, in which sound public policy reasons limit that disclosure.	**Processes should be transparent and where possible should enhance societal, groups of individual or individual interests**. D**ata stewardship values should be communicated widely**. Processes should be transparent and where possible should enhance societal, groups of individual or individual interests. The data stewardship values that govern the advanced data processing activities, such as AI or ML systems developed, and that underpin decisions, should be communicated widely. Furthermore, all societal and individual concerns should be addressed and documented as part of the EDIA process. *a)* Organizations should be able to explain how data is used, how the benefits and risks to society, groups of individuals, or individuals themselves are associated with the processing, and how society, groups of individuals and individuals themselves may participate and object where appropriate and permitted. *b)* Individual accountability systems that provide appropriate opportunities for feedback, relevant explanations, and appeal options for impacted individuals should be designed and be effective, and effectiveness should be tested. *c)* Organizations should be open about how analytical data use and advanced data processing activities, such as AI or ML systems, have been developed. Individual and societal concerns should be part of the data system evaluation lifecycle.

Enforcement and Remediation	<u>Means for remediation and external enforcement.</u> The organisation should establish a privacy policy that includes a means to address harm14 to individuals caused by failure of internal policies and practices. When harm occurs due to a failure of an organisation's privacy practices or to a lapse in its compliance with its internal policies, individuals should have access to a recourse mechanism. In the first instance, the organisation should identify an individual to serve as the first point of contact for resolution of disputes and establish a process by which those complaints are reviewed and addressed. The accountable organisation may also wish to engage the services of an outside remediation service to assist in addressing and resolving consumer complaints. Third-party agents, including seal programmes and dispute resolution services, can facilitate the consumer's interaction with the organisation and enhance its reputation for complying with its policies and meeting its obligations to individuals. Accountability practices should be subject to the legal actions of the entity or agency with the appropriate enforcement authority. Ultimate oversight of the accountable organisation should rest with the appropriate local legal authority. The nature of that authority may vary across jurisdictions. However, it is critical that the accountable organisation recognise and respond to the legal authority exercising proper jurisdiction.	**Organizations should stand ready to demonstrate the soundness of internal processes** Organizations should stand ready to demonstrate the soundness of internal processes to the regulatory agencies that have authority over advanced data processing activities, such as AI or ML processes, as well as certifying bodies to which they are subject, when data processing is or may be impactful on people in a significant manner. *a.* Organizations should be open about core values in regulator facing disclosures. *b.* Organizations should stand ready to demonstrate the soundness of the policies and processes they use and how data and data use systems are consistent with their data stewardship values and guiding principles. Depending on how data is used and what type of data is used, soundness of internal processes may be demonstrated by privacy impact assessments (PIAs), data protection impact assessments (DPIAs) or EDIAs.

Data Stewardship Accountability, Data Impact Assessments and Oversight Models

Detailed Support for an Ethical Accountability Framework

I. Enhanced Data Stewardship

The proliferation of information and communication technologies (ICT), like the Internet of Things, big data analytics, and artificial intelligence (AI), in recent years has brought significant changes to the scale and way personal data is collected, processed and used. ICT is bound to drive economic growth in the data economy of the 21st century and to bring tremendous benefits to both organizations and society by improving, for example, communications, resource allocation, productivity, and customer/client satisfaction. Data, in particular personal data, is the key element that fuels this growth engine. In addition, data-intensive activities that can involve advanced technologies are increasingly accessing and using nonpersonal data and yet can still have an impact on individuals. This use of ICT poses challenges to privacy-protection laws that rely heavily on the notions of "collection, transparency, notice and consent" and that focus just on personal data to protect the individual's right to personal data privacy.

Against this backdrop, the question arises: What would an accountable, trustworthy data-processing model look like in which data intensive activities and technologies that may have an impact on individuals are conducted in a fair and ethical manner? For example, uses of data by an organization where the use does not easily enable meaningful consent, uses that may not be within the individual's expectation, uses that cannot be explained effectively through transparency alone can raise issues about trustworthiness of advanced data-processing activities. How does the individual trust that the organization is not using the data in a way that adversely impacts his or her rights or interests yet may also provide substantial benefits?

In order to encourage innovation in various global regions, digital information strategies are being adopted that recognize that the Internet and digital technologies are transforming the world, that the needs of business, government, and the general public impact the competitiveness of their country's economy, and that the protection of personal data and fair data processing are needed for the development of Internet-based economies.[24] If individuals do not trust how organizations are using their data, and how organizations are transforming data into information and information into knowledge, and the law is not keeping up with the technology, organizations need guidance on how to act ethically and apply equitable principles particularly in advanced data-processing activities, such as AI and machine learning (ML), and the application of knowledge to enable data-driven innovation to reach its full potential.[25] Acting ethically means organizations need to understand and evaluate advanced data-processing activities and their positive and negative impacts on all parties. This approach means organizations will need to be effective data stewards not just data custodians. Data stewards consider the interests of all

[24] . E.g. Hong Kong Government's ICT Strategy & Initiatives, Hong Kong Digital 21 Strategy, March 2018. https://www.gov.hk/en/residents/communication/government/governmentpolicy.htm , and EU Digital Single Market Strategy, https://ec.europa.eu/commission/priorities/digital-single-market_en .

[25] . PDPO s 8(1)(c) charges the Privacy Commissioner with promoting awareness and understanding of, and compliance with, the provisions of the PDPO, particularly the Data Protection Principles. PDPO Data Protection Principles can be construed widely to include some principles of equity at law such as "mutual

parties and use data in ways that create maximum benefits for all while minimizing risks to individuals and other parties. They ask whether the outcomes of their advanced data processing activities are legal, fair and just[3]. Legal, fair and just is a proxy for ethical and associated and describable values. In order to determine whether advanced dataprocessing activities, such as AI and ML, that may impact people in a significant manner and/or that directly impact people, are ethical or fair, organizations should define values that are condensed to core or guiding principles and then are translated into organizational policies and processes including Ethical Data Impact Assessments (EDIAs) and appropriate independent oversight. Ultimately, data stewardship is predominantly driven by organizational policies, culture, and conduct and not technological controls.

What does an appropriate trustworthy accountability framework look like for an ethical data steward?

Ethical Data Stewardship accountability is at the foundation layer.

---

interests" between parties. W v Registrar of Marriages [2010] HKEC 1518 at 1218 ("The absence of any relevant definition in the Ordinance itself or elsewhere would also support the view that the relevant provisions should be construed in the light of moral, ethical and societal values as they are now rather than as they were at the date of first enactment and that Parliament intended some judicial license."); Consultation Document , 1.06 (The review of the PDPO was guided by (amongst other guiding principles) the principle that ". . . the rights of individuals to privacy . . . must be balanced against other rights, as well as certain public and social interests and with reference to the particular circumstances in which they arise" and "the need to balance the interests of different sectors/stakeholders. For instance, a suitable balance is needed between safeguarding personal data privacy and facilitating continued development of information and communications technology.")

[3].IAF, "Artificial Intelligence, Ethics and Enhanced Data Stewardship", September 20, 2017, 5-7. http://informationaccountability.org/wpcontent/uploads/Artificial-Intelligence-Ethics-and-Enhanced-Data-Stewardship.pdf .

II. Enhanced Data Stewardship Accountability Elements

In 2009, the accountability principle in the OECD Privacy Principles formed the basis for the Essential Elements of Accountability (Essential Elements).[26] In 2010, the EU Article 29 Data Protection Working Party issued opinion 3/2010 on the principle of accountability.[27] The Office of the Privacy Commissioner of Canada and provincial commissioners in Alberta and British Colombia adopted accountability guidance in 2012.[28] Hong Kong issued accountability guidance in 2014 and updated it in 2018,[29] and Colombia issued accountability guidance in 2015.[30] Now, accountability is the foundation of the General Data Protection Regulation (GDPR).[31] The guidance and the adoption of the GDPR has elevated accountability from check-box compliance to a risk-based approach but has not kept up with the advanced data-processing activities, such as AI and ML, that may impact people in a significant manner. In order to be able to transform data into information and information into knowledge and insight and knowledge into competitive advantage, in order for individuals to be able to trust data processing activities that might not be within their expectations, enhanced data stewardship accountability elements are needed.[10]

Working with approximately 20 Hong Kong businesses, the Enhanced Data Stewardship Accountability Elements for Data Processing Activities, such as AI and ML, that Directly Impacts People (Enhanced Elements) were drafted. The Enhanced Elements (see Appendix 1 for the complete text) call for organizations to:

1. Define data stewardship values that are reduced to guiding principles and then translated into organizational policies and processes for ethical data processing.
2. Use an "ethics by design" process to translate their data stewardship values into their data analytics and data use design processes so that society, groups of individuals, or individuals themselves, and not just the organization, gain value from the data processing activities, such as AI and ML, and require Ethical Data Impact Assessments (EDIAs) when advanced data analytics may be impactful on people in a significant manner and/or when data enabled decisions are being made without the intervention of people.

---

[26] . Essential Elements. http://www.informationaccountability.org

[27] · Article 29 Data Protection Working Party, Opinion 3/2010 on the principle of accountability, WP 173, 13 July 2010.

[28] . The Office of the Privacy Commissioner of Canada (OPC) and the Offices of the Information and Privacy Commissioners (OIPCs) of Alberta and British Columbia, "Getting Accountability Right with a Privacy Management Program," April 17, 2012. https://www.priv.gc.ca/media/2102/gl_acc_201204_e.pdf .

[29] . Hong Kong Privacy Management Programme guidance was issued in 2014 and reissued in 2018. https://www.pcpd.org.hk/pmp/files/pmp_guide2018.pdf .

[30] . Columbia Superintendence of Industry and Commerce, "Guidelines for the Implementation of the Accountability Principle," May 2015. https://iapp.org/media/pdf/resource_center/Colombian_Accountability_Guidelines.pdf .

[31] . General Data Protection Regulation 2016/679. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN .

3. Use an internal review process that assesses whether EDIAs have been conducted with integrity and competency, if the issues raised as part of the EDIA have been resolved and if the data processing activities are conducted as planned.
4. Be transparent about processes and where possible enhance societal, groups of individual or individual interests; communicate the data stewardship values that govern the data processing activities, such as AI or ML systems developed, and that underpin decisions widely; address and document all societal and individual concerns as part of the EDIA process and design individual accountability systems that provide appropriate opportunities for feedback, relevant explanations and appeal options for impacted individuals.
5. Stand ready to demonstrate the soundness of internal processes to the regulatory agencies that have authority over data processing activities, including AI or ML processes, as well as certifying bodies to which they are subject, when data processing is or may be impactful on people in a significant manner.

GDPR Article 5(2).

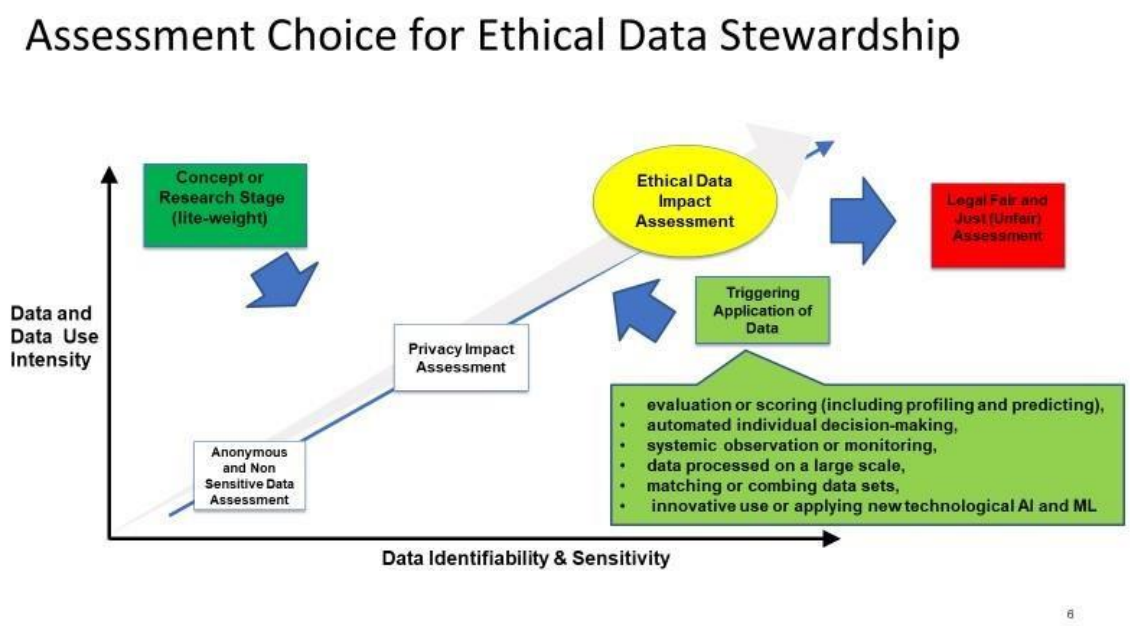
[10]. Stephen Wong, "Protecting Consumers & Competition – International Emerging Technologies," 66th ABA Section of Antitrust Law Spring Meeting, April 11, 2018, 20 ("[A]ccountability represents a perfect balance between seemingly irreconcilable interests of personal data protection and innovative use of data in data-driven economies. It helps data protection regulators realise abstract privacy principles and allows businesses to make innovative uses of data so long as they use data responsibly, minimize risks and prevent harms to data subjects.")

The Enhanced Elements of Data Stewardship are the foundation of trustworthy data intensive activities. They support model Data-Stewardship Values[32] and a Model Ethical Data-Impact Assessment that are further enabled by a Process-Oversight Model.

III. The Model EDIA

A triage process determines the type of assessment necessary for advanced data processing activities.

[32] . See IAF Research Report for an example of Ethical Values.



If data processing is very similar to processing that has been done in the past, no additional assessment may be necessary provided that the appropriate assessment has been conducted already. If the processing is less complex a more simplified Privacy Impact Assessment (PIA) may be more appropriate. At the concept or research stage of a data processing activity a light-weight version of a PIA might be appropriate to identify issues early in the development life-cycle. As data uses get more complex and/or are less obvious to the parties, a more rigorous PIA is likely required. Where the uses are most complex, under either a third-party or an in-house solution, an assessment that weighs the risks and benefits may be required. It is in these latter situations where an EDIA may be more appropriate in addition to a PIA (if the EDIA does not include all the elements of a PIA).

An EDIA is a process that looks at the full range of rights and interests of all parties in a data processing activity to achieve an outcome when advanced data analytics may impact people in a significant manner and/or when data enabled decisions are being made without the intervention of people. An EDIA assists an organization in looking at the rights and interests impacted by the data collection, use and disclosure in data-driven activities. In order to determine whether an EDIA may be necessary, the organization should consider, before the activity begins and when there are any changes that affect the scope of the activity, whether the data processing activity involves **advanced analytics such as: evaluation or scoring (including profiling and predicting), automated individual decision-making, systemic observation or monitoring, data processed on a large scale, matching or combing data sets, innovative use or applying new technological or organizational solutions (such as AI and ML).** If the data processing activity may have an impact on an individual or on a group of individuals that may not be anticipated or easily known, then an EDIA should be considered either at the concept stage or at

the service/product/analytical development stage or at both stages. If the data processing activity does not require an EDIA, then only a PIA may need to be completed.[33]

The Model EDIA consists of four sections:

I. **Purpose of the activity**
II. **Data – a full understanding of the data, data use and parties involved**
III. **Impact to parties and in particular individuals**
IV. **Decision – whether an appropriate balance of benefits and mitigated risks supports the data processing activity**.

The very nature of an ethical and values-based assessment requires a careful consideration of the data activity benefits as well as the risks to individuals and society, considering the interests of all the parties who may be part of the activity. While open, structured questions can help, a way to organize the ultimate decision as to whether to proceed can be evaluated by using a well-established risk modeling process where the outcome of the analysis (significance, likelihood and effectiveness of controls) is depicted in a "net benefit/risk heat map". This quantitative portion uses a standardized risk assessment process often found in many organizations Enterprise Risk Management (ERM) programs.

Successful implementation of an EDIA assumes and depends on the full implementation of the Enhanced Elements and, in particular, on highly qualified and competent, accountable roles and responsibilities with appropriate separation of duties. For example, EDIAs could be conducted by the privacy group. The structure of the overall Model EDIA and the questions in each section are illustrative, and the Model EDIA should be adapted as appropriate for each organization and/or industry as well as the different data-processing contexts. In particular, in the section that determines and describes how the data-processing could potentially impact the rights and freedoms of individuals, the impact should be assessed against a context-based set of issues. The Omidyar Network and Institute for the Future have established a comprehensive Ethical Framework for Tech–Techonomy organized around "risk zones"[34] that could be used as a reference guide.

The EDIA is broader in scope than the typical PIA; however, the EDIA could be used in conjunction with the PIA. For example, all data are considered in an EDIA and not just personal data. However, to the extent the EDIA can be used to consider and appropriately mitigate the impact of a personal data practice, the EDIA process may supplement (or be woven into) the organization's PIA process. In this regard, the EDIA process may enhance an organization's

[33] . A PIA Template example can be found at the CNIL. See https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-2-en-templates.pdf. The PCPD's PIA information leaflet, https://www.pcpd.org.hk//english/resources_centre/publications/files/InfoLeaflet_PIA_ENG_web.pdf, also contains information on how to conduct a PIA. The Office of the Privacy Commissioner for Personal Data, Hong Kong, conducted a Privacy Compliance Assessment Report on the Smart Identity Card System (SMARTICS), https://www.immd.gov.hk/pdf/PCAReport.pdf

[34] . https://ethicalos.org/

privacy management program and compliance with its legal obligations under various regulatory frameworks.

An EDIA does not replace a PIA; it is designed to be used in conjunction with PIAs; it is not a complete PIA. Organizations may incorporate the EDIA in whole or in part into their own unique processes and programs so as to supplement or evolve with their PIA processes.

As a Model EDIA, other relevant authorities and/or regulatory bodies may provide input into its content and format. The goal of the EDIA is to encourage ICT innovation and competition by demonstrating that an organization has considered the interests of all parties before deciding to pursue an advanced data-processing activity.

Model Ethical Data Impact Assessment

EDIA Question
Section 1: Purpose of the Activity
A. Business objective and purpose of the data activity
1. What is the business need/goal/objective for this data activity? *If the purpose of the activity is to solve a question/problem, what particular question/problem is the activity trying to solve? Does the activity fit within a larger theme of work that is currently being contemplated or undertaken?*
2. Is this activity an expansion of a previous activity? If yes, determine whether a previous assessment has been done. If a previous assessment has been done, what has changed in this data activity and why (refer to previous assessment)? *Does the activity fit within a larger theme of work that is currently being contemplated or undertaken?*
B. Accountability for the data activity
1. Who has ultimate decision-making authority for the data activity? *Who else needs to be involved in making the decision regarding the activity?*
2. Who is accountable for the various phases of the data activity? *Who are the leaders that are responsible for the activity?*

C. Legal and Other obligations regarding data collection, analysis and use(s)
1. What laws apply to the collection, analysis and use(s) of data?
2. Does the data activity comply with all organizational policies and self-regulatory commitments?
3. Are there other legal, cross-border, policy, contractual, industry or other obligations linked to the collection, analysis and use(s) of data?
4. How will all these obligations be managed and satisfied? *Have appropriate governance and accountability measures and processes been implemented?*
Section 2: Data- A Full Understanding of the Data, Data Use, and Parties Involved
A. The nature of the data
1. What specific types of data will be collected, tracked, transferred, used, stored or processed?
2. Is the data identifiable to a person? *Determine whether there has been data linking of an identifiable individual's data or the data is reasonably linkable to an individual.*
3. Is the data anonymous? *Determine how and what the anonymizing process was? Is it sufficient? Has the data been aggregated such that it is no longer identifiable personal data? Is reidentification possible? What policy, processes and/or technical measures have been used to minimize the reidentification of the data to an individual*
4. Are there data elements that are the product of a probability-based process, such as a score?
5. Is the data or anticipated use of the data sensitive?

Sensitive categories of data and/or use include Information associated with personal data that is used to decide or discriminate based on race, ethnic origin, religion or philosophical belief, sexual orientation, physical or mental health, information or data that could be used to facilitate identity theft, information associated with personal data that is used to permit access to an individual's account, precise location and/or there is a reasonable expectation the use of the data would be embarrassing to the individual whose data it is.

Would any of the data use be considered sensitive to the individual?

B. The sources of the data to be used in the activity

1. What are all the sources and governance of the data, internal and external?

Determine how the data was originated from each source and whether each source is a legitimate entity? How reliable is the source for the data activity? Is the source data permissible for the purposes of the activity? Who has custody or control over the source data and what are the governance arrangements?

2. Determine if the data is provided by the individual (originated in direct action taken by the individual) and whether:
    - The data is initiated (the product of individuals taking an action that begins a relationship)
    - The data is transactional (created when the individual is involved in a transaction)
    - The data is posted (created when individuals proactively express themselves)

3. Determine if the data is observed (created as the result of individuals being observed and recorded), whether:

    - The data is engaged (instances in which individuals are aware of observation at some point in time)
    - The data is not anticipated (instances in which individuals are aware there are sensors but have little awareness that sensors are creating data pertaining to the individuals)
    - The data is passive (instances in which it is very difficult for the individuals to be aware they are being observed and data pertaining to observation of them is being created)

4. Determine if the data is derived (created in a mechanical fashion from other data and becomes a new data element related to the individual), whether:
    - The data is computational (creation of a new data element through an arithmetic process executed on existing numeric elements)
    - The data is notational (creation of a new data element by classifying individuals as being part of a group based on common attributes shown by members of the group)

5. Determine if the data is inferred (product of a probability-based analytic process), whether:
    - The data is statistical (the product of characterization based on a statistical process)
    - The data is advanced analytical (the product of an advanced analytical process)

C. The accuracy of the data

1. Is the data accurate enough for the purpose of the activity?

Determine what steps are being taken to determine the accuracy of source data and if the source data will be accurate enough over time? Has consolidation/transformation impacted the data in such a way the accuracy is affected? Are there concerns about the quality of the final data set relative to the purpose of the activity?

2. What preprocessing will be done on the data before the analysis and will this affect the accuracy and appropriateness for the data activity?

Determine what work will be done to put the source data used in the analysis in a consistent format? How will the data sources be consolidated for analysis? Will errors and redundancy in the data to be used in the analysis be identified and dealt with during preprocessing? If yes, describe how these errors and redundancies will be identified and addressed.

3. Will preprocessing be done with data that is linkable to an individual? Describe how the preprocessing will be done and if there is any impact?

Determine if there are any sensitivity issues or unique data protection issues with respect to the preparation of the data used in the analysis? What security is appropriate for preprocessing of the data? Will the preparation steps be accurate enough over time?

D. The governance of the data

1. Outside of individuals, who are all of the possible stakeholders and parties involved or related to the data activity? What are their interests and potential concerns?

Stakeholders are very broad and apply to any party impacted by the data. A stakeholder for a framework could be a regulator or advocacy organization. Stakeholders for data and data uses include data partners. However, stakeholders can also include those interested in the success of a data use.

2. If the data has been collected by, shared with and/or received from others, do those parties have authority to share?

Determine whether the authority of those parties can be relied upon to protect impacted parties.

3. Are there restrictions on data that would affect the use of the data?

Section 3: Impact to Parties and in Particular to Individuals

A. Identify all the impacted parties and the impacts on those parties

1. During the activity, how will data be used and are there identifiable expectations of individuals, groups of individuals, and society for each use of the data?

For example, could there be an impact (real or perceived) to social or reputation status?

2. Could the data be used in a way that may result in a group of individuals being treated differently from other groups of individuals?

Determine what the goal of the difference in treatment is.

3. What are the benefits to the individual or groups of individuals?

Determine and describe what the positive impacts on the parties are that are expected to come from the data activity. Consider factors such as:
more objective or safer interactions, better product selection and utilization, better access to new products and services, significant discounts, improved service or ease of use, more convenience or improved health and well-being. Improved financial condition, lower cost alternatives or increased availability.

4. What are the benefits to society?

Determine and describe what the benefits are that could be realized by someone beyond the immediate individual whose data is being processed. The processing of data will be more legitimate if the community or society can benefit from the usefulness of the data. Consider factors such as: *better/lower cost health care, greater access to health services, or better health outcomes or an improved ability to track and assess health outcomes; more accurate sensors or devices to detect or diagnose health conditions or to improve general wellness; improved education; environmental enhancements such as water conservation, energy cost reduction; infrastructure enhancements; economic improvement; more accessible/usable technology; increased job opportunities; protection of reasonable expectation of privacy, including anonymity; protection of freedom of religion, thought and speech or protection of prohibition against discrimination.*	
5. How significant is the benefit? (1-Low; 3-Medium;5-High)	
6. Are the benefits likely to occur? How likely? (1-Low; 3-Medium;5High)	

7. What are the benefits to the organization? *Consider factors such as increased revenue; lower costs; improved profitability; greater market share; enhanced employee satisfaction; engagement and productivity; enhanced customer relationships; enhanced or maintenance of brand or reputation; assurance of compliance; fraud prevention; enhanced or maintenance of cyber or physical security; new or improved products or services or customer service; improved manner of marketing; improved ability to assess customer preferences; improvements*	

to innovation or enabling greater, faster, more efficient innovation; improved research processes; improved ability to conduct research and find or enroll study subjects or improved efficiency with studies; and innovative ways to conduct research. **NOTE:** *The benefits to the organization are not factored in the numerical assessment of significance and likelihood.*	

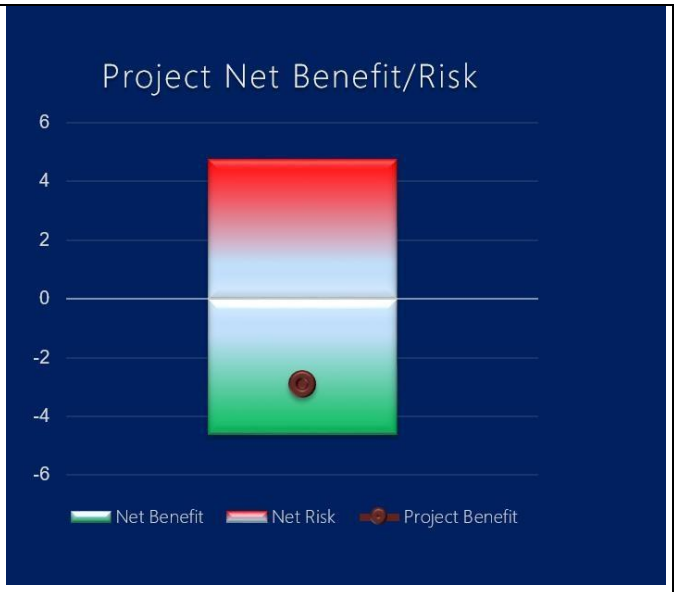
8. Considering all the factors relating to the data, the likely data use, the associated data activity, the identifiability and sensitivity of the data and the data activity objective, what are the risks to the individual, groups of individuals, and society? Determine and describe how the data processing could potentially impact the rights and freedoms to individuals. *Consider the risks or increase in risks to the individual whose data is being used and those risks that occur because of the processing being considered. Areas to consider include: physical harm; financial harm; reduced health and well-being or reduced ability to move freely in society; damage to reputation or embarrassment; shock or surprise at the processing activity or the results of the processing; inappropriate discrimination, such as where the discrimination is based on a legally protected class such as race, age, religion or politics; the possibility of inappropriate access to or misuse of data by the company, including sensitive or special categories of data and directly identifiable data; manipulation of needs or desires/wants of the individual (i.e. creation of a need where one previously did not exist); a negative impact of data that are the product of a probability-based process, such as a score; data subjects who may be in a more vulnerable position than the organisation processing the data, such as children or elderly or less-educated or impoverished individuals; larger volume processing (versus a small-scale pilot).*	

9. Is it foreseeable that the potential data analytical insights or the data activity might seem surprising, inappropriate or discriminatory or might be considered offensive causing distress or humiliation? *Would individuals be surprised by the data activity about them? Would the data activity about individuals align with the choices they have provided and the choices they have made? Determine whether there are other sensitivity issues with the potential insights and what aspect of collection/processing/analysis or use of potential insights might be considered unfair to the individual or society.*	.
10. Is the accuracy and/or quality of the data appropriate for the data activity? *Determine the impact of inaccurate data.*	
11. How significant is the risk? (1 – Low; 3 – Medium; 5 – High)	 Description For Significance or Impact / Impact Score The risk or circumstance is **Highly Impactful** — High Impact 5 The risk or circumstance is **Moderately High Impact** — Moderately High Impact 4 The risk or circumstance is **Moderately Impactful** — Moderate Impact 3 The risk or circumstance is **Moderately Low Impact** — Moderately low Impact 2 The risk or circumstance is **Minimally Impactful** — Low Impact 1
12. What factors about the activity have the highest impact on the likelihood any of these risks could be realized?	
13. How likely is the risk to be realized? (1 – Low; 3 – Medium; 5 – High)	 Description For Likelihood / Probability / Likelihood Score The risk or circumstance is **relatively certain to occur** — 90-100% — Expected 5 The risk or circumstance is **highly likely to occur** — 65-90% — Highly Likely 4 The risk or circumstance is **likely to occur** — 35-65% — Likely 3 The risk or circumstance is **possible but not likely** — 5-35% — Not Likely 2 The risk or circumstance is **only remotely probable** — <5% — Slight 1 Microsoft Confidential
14. Are there technical and procedural safeguards (mitigating controls) that could be implemented to prevent and mitigate risks should they occur (e.g. encryption and delinking of data or increased transparency)?	

A mitigating control Is a type of control used to discover and prevent mistakes that may lead to uncorrected and/or unrecorded misstatements that would generally be related to control deficiencies. A mitigating control may help to remedy any elevated risk identified in the analyses above. Determine what risks can be mitigated and how these risks can be mitigated	
15. In the case of analytical driven models, insights or algorithmic decision-making, what is the useful life of each insight for each user? (Periodic recalibration of the insight might be necessary.) Are there appropriate testing and review mechanisms in place? How has the risk of bias in the data activity been addressed? *Determine how long the potential insight might endure and determine whether potential insights could become less useful or valuable over time. Are potential insights progressive and sustainable (repeatable over time) and for how long are potential insights sustainable? Application of potential insights could impact behavior in a manner that could reduce predictive value of insights over time.*	
16. Is there a less data-intensive way to achieve the goals of the data activity (including potential insights)? *Determine whether the minimum possible amount of data has been used in the data activity or to obtain potential insights.*	
17. Have all the stakeholder concerns identified in the Governance of Data section been appropriately addressed?	
18. If data is to be shared with any identified stakeholder have appropriate mechanisms to ensure	

adherence to data obligations been put in place?	

A PIA should be done even in the case of an EDIA, and core third- party sharing controls should be evaluated for effectiveness.	
19. Does the data activity include mechanisms that explain how data is used, how benefits and risks to individuals are associated with the processing, and how individuals may participate and object where appropriate? *Determine what the transparency and individual accountability mechanisms are and whether they are appropriate for the data activity use.*	
20. How effective are these controls and safeguards in reducing risk (1 – Low; 3 – Medium; 5 – High)	

The OUTOME of the assessment of benefits, risks and controls reflected in a Residual BENEFT/RISK HEAT MAP [35]	
Section 4 – Decision: Whether an Appropriate Balance of Benefits and Mitigated Risks Supports the Data-Processing Activity.	
A. Outcome	
1. Are there any other factors that should be considered? Determine whether the interests, expectations and rights of individuals have been effectively addressed and what additional contextual based individual participation and choice factors should be considered. *Consider if the risks are necessary and proportional to the benefits? Have the risks have been mitigated to the extent possible? Are the mitigated risks sufficiently balanced by the benefits?*	
2. Does the purpose of the activity fit within the values of the organization?	
3. Does the purpose of the activity fit within the values of society?	

[35] . Net or Residual Benefit/Risk model is for illustration purposes. Individual organizations can develop and modify consistent with their own Enterprise Risk Management system; the illustrative model consists of a numerical assessment of benefits (Significance and Likelihood) – Risks (Significance and Likelihood) = Inherent Risk – Effectiveness of controls = Residual Risk.

4. After considering all the above factors, is the activity a "go," "no go," or should some aspect of the activity be recalibrated to reduce the residual risk?	
B. Approvals	
1. Have all the individuals described in I.B.1. through I.B.2. above been involved in the decision?	

IV. The Process Oversight Model

Assessments conducted solely by the parts of a business implementing intensive data activities may raise issues of trustworthiness. Where the oversight of the assessment and accountability process is done by the organization itself (versus the accountability or regulatory agency), then the oversight should be conducted pursuant to a common framework.[36] Until such an approach is established, the Process Oversight Model looks at how an organization has translated organizational ethical values into principles and policies and into an "ethics by design" program. It considers how well established internal review processes, such as EDIAs and effective individual accountability systems, have been implemented. It presumes the oversight process is independent from the assessment process. It could be a function performed by, for example, an internal audit group. It may be likened to an assessment of "controls and controls effectiveness" by the internal audit group.

The internal audit group usually is established by the Audit Committee of the Board of Directors or the highest level of the governing body. The Chief Audit Executive reports functionally to the Board, and the internal audit function is independent and objective. The scope of internal audit's responsibilities encompasses, but is not limited to, the examination and evaluation of the adequacy and effectiveness of the organization's governance, risk management, and internal controls.[37] The Process Oversight Model can be thought of as analogous to a set of control definitions against which the capability and effectiveness of the organization's assessment process is tested. A set of control parameters across functional assessment domains is established and then, through a set of audits, the effectiveness of the relative controls is tested. While this oversight could be performed by internal audit, it could also be accomplished by way of an assessment or test conducted by an external resource (e.g. a consulting firm). This sort of audit and

[36] . See IAF, Report for Comprehensive Assessment Oversight Dialog: Canadian Ethical Data Review Boards Project, March 31, 2018, 18-24 [IAF Oversight Report]. http://informationaccountability.org/wp-content/uploads/Report-for-the-Comprehensive-Assessment-Oversight-Dialog-Canadian-Ethical-DataReview-Boards-Project.pdf

[37] . "Model Internal Audit Activity Charter," The Institute of Internal Auditors (rev. 05/2013) https://global.theiia.org/standards-guidance/public%20documents/modelcharter.pdf .

testing work is similar to work already performed by these external firms in other domain areas.

The Oversight Model consists of questions in seven sections**:**

I. **Accountability for the oversight process**
II. **Translation of organization values into principles and policies**
III. **Translation of organizational values into an "ethics by design" program**
IV. **Utilization of the EDIA**
V. **Internal review process**
VI. **Individual accountability system**
VII. **Transparency of process.**

The questions in each section of the Process Oversight Model are illustrative, and the Process Oversight Model should be adapted as appropriate for each organization to oversee the trustworthiness of its assessment process.[38] The Process Oversight Mode is designed to address the ethics part of data stewardship and assumes other internal oversight processes exist to address core elements of privacy programs.

Evidence of oversight is important. Oversight provides rigor to the assessment process and demonstrates that oversight of the EDIA process has occurred. Whether this oversight occurs internally, for example by the audit group, or externally, for example by a consulting firm, it is necessary that documentation exists that demonstrates how the oversight was conducted and that, in fact, it was conducted. The oversight process should measure whether the EDIA process is being conducted with honesty and recognizes the full range of interests of all parties in order to demonstrate that the interests of the organization were not placed in front of the interests of other parties.[18] The organization should stand ready to demonstrate its assessment governance process and individual assessments to regulators with appropriate authority.[39] The Process Oversight Model provides guidance regarding how such oversight should be conducted and documentation that the oversight actually occurred.

The questions in each section of the Process Oversight Model are illustrative and should be adapted as appropriate for each organization to oversee the trustworthiness of its assessment process[40]. The Process Oversight Mode is designed to address the ethics part of data stewardship and assumes other internal oversight processes exist to address core elements of privacy programs. As process oversight models evolve, there may be input and guidance from other relevant authorities and/or regulatory bodies. Such input and guidance will increase the trustworthiness of the EDIA process.

[38] . An assessment of the process is designed to be different than a secondary assessment of a specific data-intensive activity. [18]. IAF Oversight Report p. 21
[39] . Id. pp. 23-24.
[40] . An assessment of the process is designed to be different than a secondary assessment of a specific data intensive activity.

Process Oversight Model

Oversight Question
I. Accountability for Oversight Process
1. Are accountability and responsibility for achieving outcomes established through clearly defined roles throughout the organization? *Are the accountable and responsible roles carried out by competent and capable individuals? Is there a clear separation of duties between data activity roles?*
II. Translation of Organizational Values into Principles and Policies
1. Are shared organizational values described and/or articulated and have they been integrated into the organization? *Have the values have been condensed to core and guiding principles that are understood by technical staff? Have they been fully translated into organizational policies and processes? Have they been programmed into data and activity objectives?*
2. Have the articulated values been aligned to the varied geographic-values across the organization's reach and footprint?

Could design choices become international standards or norms?
III. Translation of Organizational Values into an "Ethics by Design" Program
1. Does the organization have an "ethics by design process" that is part of its products/service development process? *Determine whether Core or Guiding Principles are understood by staff (in particular by technical staff) and have been programmed into activity objectives and the full product/service development lifecycle.*
2. Does the product/service development process ascertain whether there is benefit to individuals and society in addition to the organization?
IV. Use of an EDIA

1. Does the organization use an EDIA to achieve a principlesbased outcome of data? Is the assessment process effective?

Does the organization assess all risks and benefits to an individual, group of individuals, and society? Are the risks effectively mitigated? Does the EDIA process effectively evaluate that data use avoids actions that seem inappropriate or discriminatory, might be seen as generating unequal treatment, might be considered offensive or causing distress or humiliation?

2. Does the EDIA process effectively assess the complexity and potential impact of the data and data use?

Does the EDIA process consider all the factors relating to the data, the likely data use, the associated data activity, the identifiability and sensitivity of the data, as well as the potential impact of the data activity?

3. Does the EDIA process effectively evaluate if the purpose of the activity fits within the values of the organization and society?

4. Is there an effective triage process to determine what type of assessment is appropriate? Is this process effectively employed?

A triage process determines the level of review of the process necessary. Where data processing is very similar to processing that has been done in the past and therefore it was concluded no assessment was necessary, only a quick review may be required to confirm those understandings. Where data uses are more complex and/or less obvious to the parties and more rigorous assessments were conducted, a more rigorous review should be required. Where the uses are most complex, an EDIAs that effectively weighs the risks and benefits should be used.

V. The Internal Review Process

1. What kind of periodic assurance reviews will occur over time?

Do the periodic reviews appropriately consider the data and datause objectives? Is the periodic review process established at appropriate timeframes?

2. For intensive data impacting systems, does the review assess that outcomes are as intended with the objectives of the activity and impacts are mitigated as planned, harms are reduced, and unintended consequences are understood?

Determine whether this analysis includes the likelihood of benefits being achieved and risks effectively mitigated. Does the post review include an assessment of if the anticipated outcomes were achieved?

3. Have analytic models and insights been tested for their accuracy and predictability?

Is there an ongoing systematic process to ascertain whether analytic models are tested for their consistency with organizational values and principles? Are data-intensive technologies subject to appropriate human direction and control?

4. Does the review process include a risk review by senior accountable leadership? Are higher risk activities approved by senior-accountable leadership?

Is there a formalized, risk-ranked review process where higher impacting data activities are reviewed? Where internal reviewers need external expertise? Is this expertise sought?

5. Does the assessment and review process ascertain whether all parties' concerns are assessed and appropriately addressed as part of the data-system lifecycle?

6. Have systems themselves, and the data that feed those systems, been assessed and protected proportionate to the risks?

7. Does the review evaluate whether only the minimum data that is needed is used?

VI. Individual Accountability System

1. Are there effective systems that provide appropriate opportunities for feedback, relevant explanations, and appeal options for the individuals impacted?

Will individuals have some ability to engage in how their data is used? How will individual situations be remediated, if necessary?

VII. The Transparency of the Process

1. Does the organization have mechanisms that explain how data is used, how benefits and risks to individuals are

associated with the processing, and how individuals may participate and object where appropriate?

Is the use of the data transparent and effectively made available for all data activities?

2. Is the organization ready to demonstrate the soundness of the processes they use so that data and data-use systems are consistent with established values and principles?

Can the organization demonstrate its data stewardship accountability processes?

Appendix 1

Enhanced Data Stewardship Accountability Elements for Advanced Data Processing Activities, such as Artificial Intelligence (AI) and Machine Learning (ML), that Directly Impacts People

1. As a matter of organizational commitment, organizations should define data-stewardship values that are condensed to guiding principles and then are translated into organizational policies and processes for ethical data processing.
    a. These values and principles should be organizationally derived and should not be restatements of law or regulation. They may go beyond what the law requires, but at a minimum, they should be aligned, and not be inconsistent, with existing laws, regulations, or formal codes of conduct.[41] Organizations should be open about their values and principles.
    b. Organizational policies and processes derived from these values should be anchored to clearly defined, accountable individuals within the organization and should be overseen by designated senior executives.
    c. The organization's data stewardship guiding[42] principles should be easily understood by all staff, and in particular by technical staff, and should be capable of being programmed into activity objectives.
2. Organizations should use an "ethics by design" process to translate their data-stewardship values into their data-analytics and data-use system design processes so that society, groups of individuals, or individuals themselves, and not just the organizations, gain value from the data processing activities, such as AI or ML.
    a. Advanced data-processing activities, such as AI and ML, that affect individuals should have beneficial impacts accruing to individuals and communities of individuals, particularly those to whom the underlying data pertains.
    b. Where an analytical data driven use has potential impact at the individual level, or at a higher level, such as groups of individuals and society, the risks and benefits should be explicitly defined. The risks should be necessary and proportional to the benefits and should be mitigated to the extent possible.
    c. The systems, and the data that feeds those systems, should be assessed for appropriateness based on the decision the data is being used for and should be protected proportional to the risks.
    d. Where appropriate, organizations should follow codes of conduct that standardize processes to industry norms.
    e. Ethical Data Impact Assessments (EDIAs)[43] should be required when advanced-data analytics may impact people in a significant manner and/or when data-enabled decisions are being made without the intervention of people.
        i. An EDIA is a process that looks at the full range of benefits, risks, rights, obligations, and interests of all individuals, groups of individuals, society and other data stakeholders, such as regulators.

---

[41] . Examples of existing professional or industry codes of conduct are those that relate to AI or ML. These Elements should work with those codes and not replace them.

[42] . See IAF Blog: The Need for an Ethical Framework. http://informationaccountability.org/the-need-for-an-ethical-framework/

[43] . See here for A Model EDIA.

ii. An EDIA is a means of determining whether an instance of processing is in accordance with the data stewardship values and guiding principles established by the organization. Processing includes all steps necessary to achieve an outcome, from the collection of data through the implementation of data-driven outcomes.
iii. Organizations should have EDIAs that achieve an "ethics by design" process that is integrated into systems development.

f. All staff involved in data impacting processing should receive training so that they may competently participate in an "ethics by design" process.

3. There should be an internal review process that assesses whether EDIAs have been conducted with integrity and competency, if the issues raised as part of the EDIA have been resolved, and if the advanced data processing activities are conducted as planned.[44]
    a. Where data processes begin with analytic insights, those insights should be tested for accuracy, predictability, and consistency with organizational values.
    b. Intensive data impacting systems should be reviewed so that outcomes are as intended with the objectives of the activity, risks are mitigated as planned, harms are reduced, and unintended consequences are understood.
    c. Where internal reviewers need external expertise, that expertise should be sought.
    d. The review of the EDIA process is separate and independent from the EDIA process.
4. Processes should be transparent and, when possible, should enhance societal, groups of individual or individual interests. The data-stewardship values that govern the advanced data-processing activities, such as AI or ML systems developed, and that underpin decisions, should be communicated widely. Furthermore, all societal and individual concerns should be addressed and documented as part of the EDIA process.
    a. Organizations should be able to explain how data is used, how the use may benefit and potentially pose risks to society, groups of individuals, or individuals themselves are associated with the processing, and how society, groups of individuals and individuals themselves may participate and object.
    b. Individual accountability systems that provide appropriate opportunities for feedback, relevant explanations, and appeal options for impacted individuals should be designed and be effective, and effectiveness should be tested.
    c. Organizations should be open about how analytical data use and advanced data processing activities, such as AI or ML systems, have been developed. Individual and societal concerns should be part of the data system evaluation lifecycle.
5. Organizations should stand ready to demonstrate the soundness of internal processes to the regulatory agencies that have authority over advanced data-processing activities, such as AI or ML processes, as well as certifying bodies to which they are subject, when data processing is or may impact people in a significant manner.
    a. Organizations should be open about core values in regulator-facing disclosures.
    b. Organizations should stand ready to demonstrate the soundness of the policies and processes they use and how data and data-use systems are consistent with their data stewardship values and guiding principles. Depending on how data is used and what type of data is used, soundness of internal processes may be demonstrated by privacy-impact assessments (PIAs), data protection impact assessments (DPIAs) or EDIAs.

[44] . See here for A Model Oversight Assessment.