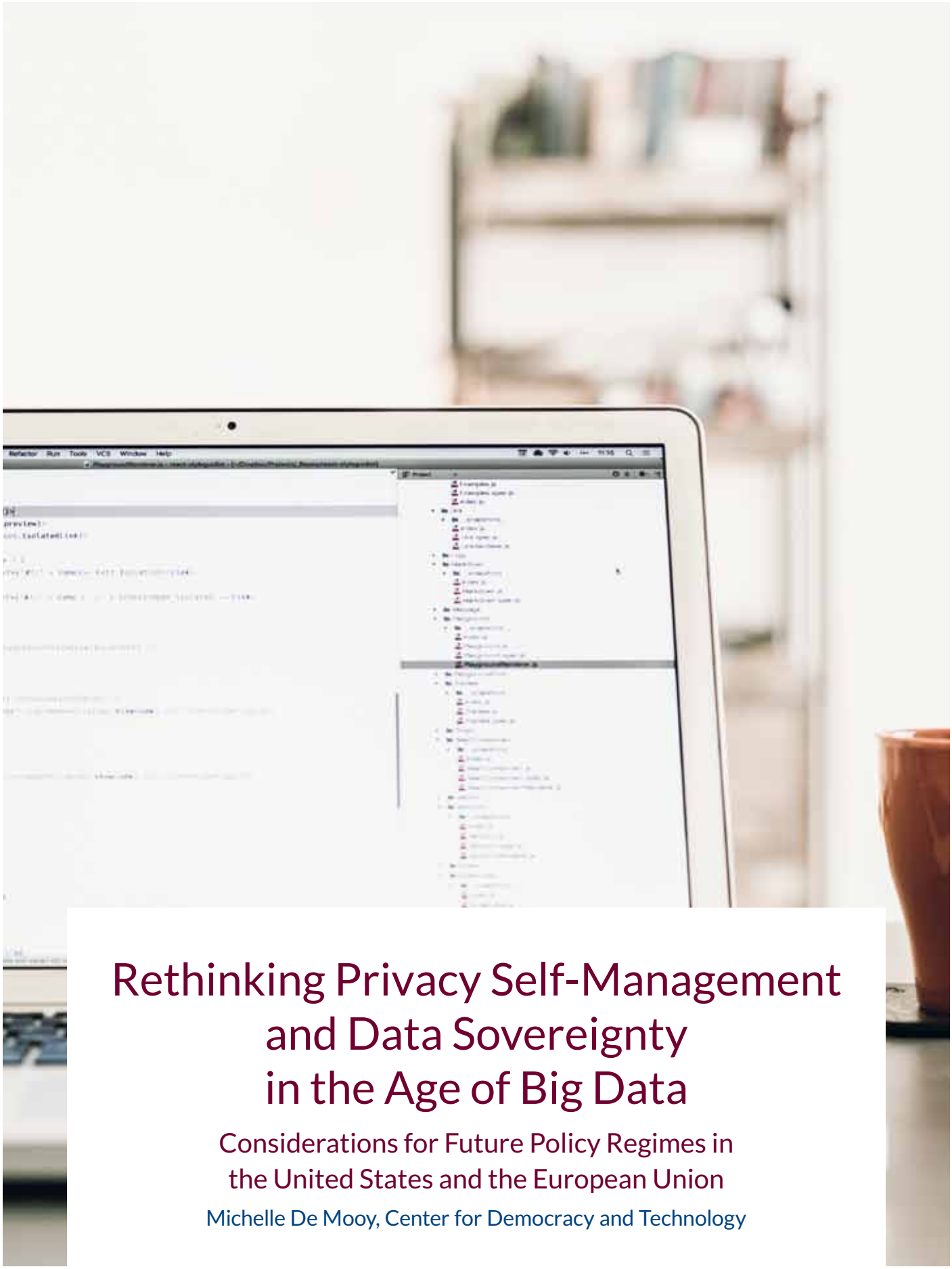


Rethinking Privacy Self-Management and Data Sovereignty in the Age of Big Data

Considerations for Future Policy Regimes in the United States and the European Union

Michelle De Mooy, Center for Democracy and Technology

BertelsmannFoundation

CENTER FOR DEMOCRACY & TECHNOLOGY

BertelsmannStiftung

Content

Preface

Every day, we leave behind a trail of vast quantities of data. Today's social media, search engines and the internet of things produce more data in only a brief period of time than were previously generated in all of human history. Experts at IBM and the University of California, Berkeley, estimate that by the year 2020, the volume of data worldwide will reach 43 zettabytes – a number with 21 zeroes – thus totaling 300 times the data that existed worldwide last year.

Big-data technologies make it possible to collect enormous quantities of data, connect diverse kinds of information and conduct rapid analyses. They enable analysts to find hidden correlations that are relevant to a social problem or business challenge – whether this is the early diagnosis of a disease or an analysis of consumer behavior and predictions of how it will change. But big data also poses a challenge to personal privacy. When data is analyzed, the findings can also affect people who have not consented to the use of their data for that specific purpose. Moreover, it is becoming increasingly common for us to be unaware of the means by which information is being collected, for example through sensors on devices connected to the internet. This is in conflict with traditional principles under which data collection is to be minimized and the use of data is to be limited to a specified purpose. Big data poses significant challenges to data-protection regimes that emphasize individual control – and may even call their very viability into question. Internet users are finding it more and more difficult to maintain control of their own data, and are feeling increasingly powerless as their data is collected and used ever more extensively.

Mindful of these trends, the Bertelsmann Stiftung commissioned the Center for Democracy and Technology to conduct this analysis. The report suggests three concepts for dealing with personal data that would meet the needs of big data while also taking into account individuals' legitimate interest in the protection of their personal data. The focus is on the ideal of data sovereignty, which can only be reached by intertwining individual, entrepreneurial and governmental responsibility.

Our thanks go to the Center for Democracy & Technology for drawing up this expert report, as well as to Prof. Dr. Iris Eisenberger and Prof. Dr. Alexander Roßnagel for their critical comments and helpful suggestions.

This analysis is the first step in our exploration of the topic of "Social Participation in an Era of Algorithms and Big Data," through which we will examine the effects of the digital environment on social participation. We look forward to your feedback and suggestions.

Eric Thode
Director
International Fora and Trends

Ralph Müller-Eiselt
Senior Expert
Taskforce Digitization

Vorwort

Tagtäglich hinterlassen wir Unmengen an Datenspuren. Durch soziale Medien, Suchmaschinen und das Internet der Dinge entstehen heute in kurzer Zeit so viele Daten wie in der gesamten Menschheitsgeschichte davor. Auf 43 Zetabyte schätzen Experten von IBM und der Universität Berkeley das weltweite Datenvolumen im Jahr 2020. Das ist eine Zahl mit 21 Nullen – und wären 300-mal mehr Daten, als im vergangenen Jahr weltweit bestanden.

Big-Data-Technologien machen es möglich, diese Datenmassen zu erfassen, verschiedenste Arten von Daten miteinander zu verknüpfen und mit hoher Geschwindigkeit auszuwerten. So lassen sich in einer Vielzahl von Einflussfaktoren verborgene Korrelationen finden, die relevant für ein gesellschaftliches Problem oder eine unternehmerische Herausforderung sind, sei es die Früherkennung von Krankheiten oder die Analyse und Prognose unseres Konsumverhaltens. Im Hinblick auf den Datenschutz wird dieses Potential gleichzeitig zur Herausforderung. Denn von den Ergebnissen solcher Auswertungen können auch Menschen betroffen sein, die nicht zugestimmt haben, dass ihre Daten zu diesem konkreten Zweck verwendet werden. Zudem werden Daten, die Grundlage für Big-Data-Berechnungen sind, zunehmend über Wege erfasst, die wir nicht bewusst wahrnehmen, zum Beispiel über Sensoren von Geräten, die mit dem Internet verbunden sind. Diese Entwicklungen hebeln bisherige Prinzipien der Zweckbindung und Datensparsamkeit aus. Big Data stellt Datenschutzkonzepte, die dem Leitmotiv individueller Kontrolle folgen, mindestens vor erhebliche Herausforderungen, wenn nicht gar grundlegend in Frage. Für die Internetnutzer wird es zunehmend schwieriger, die Kontrolle über ihre Daten zu behalten. Sie fühlen sich machtlos angesichts der umfassenden Sammlung und Verwendung ihrer Daten.

Diese Entwicklungen sind Anlass und Ausgangspunkt für diese Analyse, die das Center for Democracy and Technology im Auftrag der Bertelsmann Stiftung verfasst hat. Der Beitrag macht Vorschläge für drei Konzepte zum Umgang mit personenbezogenen Daten, die sowohl den Anforderungen von Big Data als auch den berechtigen Schutzinteressen des Individuums gerecht werden sollen. Im Zentrum steht dabei das Ideal der Datensouveränität, das nur durch ein Ineinandergreifen individueller, unternehmerischer und staatlicher Verantwortung erreicht werden kann.

Unser Dank gilt dem Center for Democracy and Technology für die Erstellung dieser Expertise sowie Prof. Dr. Iris Eisenberger und Prof. Dr. Alexander Roßnagel für ihre kritische Prüfung und wertvollen Anregungen.

Diese Analyse bildet den Auftakt zu einer Exploration zum Thema „Teilhabe in Zeiten von Algorithmen und Big Data“, in der wir uns näher mit den Auswirkungen von Phänomenen der digitalen Sphäre auf gesellschaftliche Teilhabe beschäftigen. Wir freuen uns über Feedback und Anregungen zu diesem Papier.

Executive Summary

This paper advances the idea that the rise of large data collection and processing, also known as big data, has challenged the validity of data-protection regimes founded on ideals of individual control. With a focus on data sovereignty, it investigates concepts able to meet the requirements of big-data technologies, while also offering guidance for future policy regimes.

We begin by looking closely at the political philosophies and legal theories grounded in the rights of individuals that have shaped data-protection frameworks in the United States, the European Union and Germany. Each of these systems approaches data protection differently, yet each is premised on the concept of an individual having some control over his or her personal information. The basis for this analysis is the American perspective. The U.S. regulates data by type and sector. The focus lies here on individual consent, which many U.S. companies apply in a "take it or leave it" approach. In Germany, the processing of personal data also needs individual consent in theory, as it interferes with the right of informational self-determination. However, individual consent is only seldom obtained in practice. Instead, numerous regulations give organizations the right to engage in personal-data processing even without the explicit agreement of the individual. Nevertheless, the principle of informational self-determination grants individuals various constitutionally protected rights - for example, the right to examine, correct or delete stored personal data - which enable them to exercise control over their data. The European Union's legal framework also bases its data-protection mechanisms on the concept of individual control, thus assigning responsibility for data management to the individual. Thus, principles such as transparency, purpose specification and data minimization have shaped existing legislation on both sides of the Atlantic.

More generally, big data has fundamentally upended the role of individuals in managing their personal information. Data-protection regimes have struggled to keep pace. In addition to looking at specific national or supranational philosophies, we explore big data's impact on traditional notions of privacy management and long-standing data-protection laws. Because public trust is a crucial component in any successful data-protection regime, we consider public opinion in the United States, the European Union and Germany on the issues of big data, individual control and privacy, highlighting commonalities that exist despite historical and cultural differences. People on both sides of the Atlantic have reported a similar sense of powerlessness with regard to the control of their personal information, though there is generally less agreement on the appropriate role of regulation and regulators in the protection of such data. Americans tend to be resigned to the commercialization of their personal information, while Europeans generally react more negatively to such uses of data.

Finally, we examine possible new ways to achieve individual control in this big-data world. We investigate three complementary notions of privacy self-management that may offer a way forward in constructing modern privacy regulations, with data sovereignty playing the central role. The first concept, dealing with education and data portability, would give more responsibility to individuals, empowering as well as burdening them. However, since the empowerment of individuals alone cannot address all the challenges presented by big data, a second approach would make companies responsible for data protection in the form of voluntary industry self-regulation. This would relieve individuals of a portion of the data-management burden; however, self-regulation often fails to meet the standards of accountability and transparency fully. To account for this potential shortfall, a third concept is introduced, in which third parties would perform state-mandated impact assessments of data-management practices, advocating for users' interests and creating greater transparency. However, while these third-party assessments could help

users, there is a risk of treating users in a patronizing manner. To prevent this, users would need to engage in the education addressed in the first concept, thus enabling them to use the assessments in a self-determined manner. These collective approaches can address the challenges posed by big data. The basis for their implementation remains governmental regulation, which assigns rights to individuals, creates a dependable framework and balances power asymmetries. As regulatory systems have been stretched to their limits by the challenges of digitization, a multipronged approach of the kind advocated by this report is necessary to overcome the weaknesses inevitable in any single concept.

Zusammenfassung

Big Data stellt bisherige Datenschutzsysteme, die auf dem Ansatz der individuellen Kontrolle basieren, vor erhebliche Herausforderungen. Dieses Papier diskutiert Konzepte, die mit einem Fokus auf Datensouveränität sowohl den neuen Anforderungen als auch den Schutzinteressen des Individuums gerecht werden und zukünftige Datenregime anleiten können.

Dazu werden zunächst die bestehenden Rechtsordnungen in den USA, Europa und Deutschland betrachtet. Ausgangspunkt ist dabei die amerikanische Perspektive. Die USA regulieren Daten sowohl nach dem Typ als auch dem Sektor, in dem sie verwendet werden. Dort stehen individuelle Einwilligungen im Vordergrund, wie sie viele amerikanische Unternehmen anwenden, wenn sie ihren Kunden nur die Wahl zwischen Nutzung und damit verbundener Datenpreisgabe oder dem Verzicht auf ihre Dienste lassen. Auch in Deutschland bedarf die Verarbeitung persönlicher Daten zwar grundsätzlich der individuellen Zustimmung, weil sie einen Eingriff in das Recht auf informationelle Selbstbestimmung darstellt. In der Praxis kommen solche Einwilligungen hierzulande aber nur selten zum Einsatz. Stattdessen ist die Datenverarbeitung durch viele gesetzliche Regelungen organisiert, die jene in den meisten Fällen ohne explizite Zustimmung des Individuums erlauben. Der Staat räumt diesem aber mit der informationellen Selbstbestimmung verschiedene durch die Verfassung geschützte Rechte ein, etwa auf Auskunft, Korrektur oder Löschung, mit denen jede und jeder Kontrolle über die eigenen Daten ausüben kann. Auch die europäische Rechtsordnung gründet ihre Datenschutzkonzepte auf dem Ansatz der individuellen Kontrolle, die dem Bürger Verantwortung für die Verwaltung seiner Daten zuschreibt. Insofern prägen Prinzipien wie Transparenz, Zweckbindung und Datensparsamkeit die bisherige Rechtsprechung auf beiden Seiten des Atlantiks.

Dieser Beitrag untersucht, welchen Einfluss Big Data auf etablierte Ansätze des Datenschutzmanagements und der Datenschutzregulierung hat. Da das Vertrauen der Öffentlichkeit ein wichtiges Element erfolgreicher Datenschutzordnungen ist, berücksichtigt das Papier auch die Ansichten der Internetnutzer in den USA, Europa und Deutschland zu Big Data, individueller Kontrolle und Datenschutz. Dabei sticht trotz kultureller und historischer Unterschiede eine Gemeinsamkeit hervor: Auf beiden Seiten des Atlantiks fühlen sich die Internetnutzer ähnlich machtlos, wenn es um die Kontrolle über ihre persönlichen Daten geht. Weniger einig sind sie sich hingegen, welche Rolle Regulierung und Regulierer im Datenschutz einnehmen sollten. Amerikaner neigen bereits dazu, vor der Kommerzialisierung ihrer persönlichen Daten zu resignieren, Europäer lehnen diese Entwicklung und Nutzung ihrer Daten hingegen deutlich ab.

Angesichts der neuen Herausforderungen durch Big Data, denen die bisherigen auf individueller Kontrolle basierenden Datenschutzregime nicht gewachsen sind, schlägt das vorliegende Papier drei Konzepte vor, die sich gegenseitig ergänzen und ineinandergreifen. Das erste gibt dem Einzelnen mehr echte Verantwortung für die Verwaltung seiner Daten, indem er nicht nur zu einem selbstbestimmten Umgang mit ihnen berechtigt, sondern durch entsprechende Bildungsangebote auch dazu befähigt wird. Die Stärkung des Individuums reicht aber nicht aus, um auf die neuen Anforderungen durch Big Data adäquat zu reagieren. Es braucht auch Ansätze freiwilliger unternehmerischer Selbstregulierung, die die Last der Verantwortung nicht allein bei den Bürgern belassen. Um diese Selbstregulierung so transparent und nachvollziehbar wie möglich zu machen, wird sie im dritten Konzept durch obligatorische Risikoabschätzungen unabhängiger Dritter ergänzt. Mit diesen kollektiven Ansätzen kann den Herausforderungen, die Big Data an den Datenschutz stellt, begegnet werden. Grundlage dafür ist und bleibt allerdings eine staatliche Regulierung, die den Einzelnen mit Rechten ausstattet, verlässliche Rahmenbedingungen schafft und Machtungleichheiten ausbalanciert. Da diese aber in Zeiten der Digitalisierung an Grenzen stößt, ist ein multiperspektivischer Ansatz, wie ihn dieser Beitrag vorschlägt, eine sinnvolle Variante, um die Stärken der einzelnen Konzepte miteinander zu verbinden.

I. Introduction

Data-protection laws have struggled to keep pace with an exciting and rapidly evolving digital environment. Regulatory regimes in the United States and Europe have taken a similar approach toward data protection since the 1970s, giving individuals the right to make decisions about how to manage their data. This concept is reflected in the term "individual control," which is defined as the extent to which a person can influence outcomes in a way that reflects their beliefs and wishes. In the context of data management, individual control refers to the ability of a person to determine "when, how and to what extent information about them is communicated to others."[1] With regard to privacy, individual control is seen as a foundational principle. As American jurist and attorney Charles Fried has said, "Privacy is not simply an absence of information about us in the minds of others; rather, it is the control we have over information about ourselves."[2] In the legal and regulatory context, the notion of individual control has often been focused on transparency (making people aware of any records kept about them) and on redress (giving them the ability to correct or amend these records). It has also been interpreted as implying a form of privacy self-management, meaning that an individual can decide whether information about her is used or shared.

In American policy, an individual-control-oriented approach made sense nearly 50 years ago when computing systems were mostly centralized and people had little need to exercise their data-control rights on an ongoing basis. Many concepts of individual control within data-protection regimes have since that time been derived from the Code of FIPs,[3] a set of data-management principles that include standards for purpose specification, use limitation and data minimization. The Code of FIPs is based on five principles: 1) the existence of personal-data record-keeping systems should not be kept a secret; 2) people must have a way to find out what information about themselves is being stored and how it is being used; 3) people must have a way to prevent information about them obtained for one purpose from being used or made available for other purposes without their consent; 4) people must have a way to correct or amend records containing personally identifiable information; and 5) all organizations creating, maintaining, using or disseminating personally identifiable data must assure that the data is reliably being used as intended, and must take precautions to prevent misuse. The FIPs also contain a collection of individual rights, such as access and consent requirements for the collection, use and disclosure of personal information, which remain de rigueur in privacy policies and practices across the globe today. These practices and rights have remained a cornerstone in the data ecosystem despite enormous changes in the use and capture of personal information through the advent of massive data generation and processing from digital sources, otherwise known as big data.

Big data has the potential to produce vast benefits for society in a variety of sectors including public health, the environment and city management. However, realizing this potential will require forward-thinking policy solutions that leave behind outdated interpretations of individual control, and instead focus on creating mechanisms that offer individuals authority, practical impact assessments and robust accountability in such a way as to build public trust and engagement. New approaches to privacy self-management, such as data sovereignty and data portability, may offer promising ways to achieve these goals. Fitting these concepts into existing data-regulation frameworks, however, will be a daunting challenge for policymakers.

Policymakers on both sides of the Atlantic have grappled with how to adapt existing data-protection regimes to the

1 Westin, Alan F. and Louis Blom-Cooper. Privacy and Freedom. London: Bodley Head, 1970: 7. ISBN 978-0370013251.

2 Fried, Charles "Privacy." Yale Law Journal 77 (3), January 1968: 475-493. doi: 10.2307/794941.

3 U.S. Department of Health, Education and Welfare, Secretary's Advisory Committee on Automated Personal Data Systems, Records, computers, and the Rights of Citizens viii (1973).

complexities of big data, mitigating the harms that can arise from big-data analytical procedures such as predictive modeling while also providing individuals with more say in how their data is used. Many data-protection laws emphasize traditional mechanisms of individual control such as transparency, despite the fact that big data's opacity and speed of processing have limited an individual's practical ability to evaluate the consequences of data-sharing choices or indeed to provide meaningful consent to such sharing.

The public has expressed confusion over about how personal information is used, as well as a sense of powerlessness regarding the ability to control such information in the big-data environment.[4] Commercial and noncommercial entities that rely on big data have suggested that members of the public are aware of and accept the fact that free online services are "paid for" through the provision of personal information, but this argument evades the fact that it has become almost impossible for individuals to evaluate such trade-offs,[5] much less implicitly agree to them. Surveys consistently find that individuals feel resigned to the ubiquitous collection and use of their personal information.[6] Concerns about privacy and security have eroded the public's trust in data systems. This gradual loss undermines the potential of big data, and hampers the ability of the commercial and noncommercial entities that rely on it to innovate.

Against this background, this paper will also scrutinize how data regimes have incorporated concepts of individual control, how big data has impacted such concepts and how public policy must evolve to address these impacts. We will examine how regulation in the United States, the European Union and Germany incorporate concepts of individual control, considering the role of influential bodies such as the United States Federal Trade Commission and the Article 29 Working Party.[7] The paper also reviews key legislation and legislative instruments, such as the European Union Data Protection Directive (DPD) and the Data Protection Regulation (DPR), which have shaped data-protection regimes.

Finally, we will contemplate how data-usage and data-governance models, including data sovereignty and data portability, might evolve beyond a reliance on individual control by employing collective efforts to assess data processing and uses, and by placing an emphasis on understanding the impact of these activities on individuals and society. The paper will provide an analysis of three potential data-usage frameworks – education and data portability, industry self-regulation, and legally mandated impact assessments – that address the challenges big data poses to effective individual control. Finally, we will conclude with a discussion of the strengths and weaknesses of each framework.

4 Madden, Mary. Public Perceptions of Privacy and Security in the Post-Snowden Era. Pew Research Center, Nov. 12, 2014. www.pewinternet.org/2014/11/12/public-privacy-perceptions/

5 Acquisti, Alessandro et al. "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study." Carnegie Mellon University.

6 Pew Research Center. "Americans' Attitudes About Privacy, Security and Surveillance." May 2015. www.pewinternet.org/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/

7 The Article 29 Working Party is an advisory body set up under the 1995 Data Protection Directive.

II. A history of individual control in the United States, the European Union and Germany

Data-protection laws in the United States and Europe rely heavily on the FIPs, the collection of data-management principles that have guided international law, policy and standards since the 1970s. Concerns about the rising use of automated computer systems in that era prompted the creation of the FIPs. After the inventions of the internet and the personal computer in the late 1970s, regulators in the United States and the European Union were forced to consider how to protect individual privacy in a vast global communications network, and how to contend with the rise of a robust market for personal information. In 1980, the Council of Europe adopted a Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data. This convention stated that "it is desirable to extend the safeguards for everyone's rights and fundamental freedoms, and in particular the right to the respect for privacy, taking account of the increasing flow across frontiers of personal information undergoing automatic processing," a sentiment that continues to resonate today. Since then, despite massive changes with regard to the power, speed and size of such automated systems, the FIPs have persisted over time as the gold standard for cross-border data protection.

Specifically, both the United States and the European Union enacted laws that provide a set of rights enabling individuals to make decisions about how to manage their data. These rights were derived from the FIPs, which center on regulation focused on individuals, and consist primarily of the rights to notice, access and consent regarding the collection, use and disclosure of personal information. Individual control is a concept deeply rooted in social theories of self-determination and autonomy, which in turn are cornerstones of most privacy laws in the United States and the European Union. Article 8 of the European Convention on Human Rights, which addresses the "right to respect for private and family life," for example, established the principle of individual autonomy with regard to privacy. In the United States, the Third, Fourth and Fifth Amendments of the Constitution protect the pursuit of liberty through an array of privacy rights in the home and in communications. U.S. law considers privacy to exist in a physical or personal place, such as in the home or on the job; European privacy law often includes collective rights, and is thus informed by notions of privacy that conform with social values.

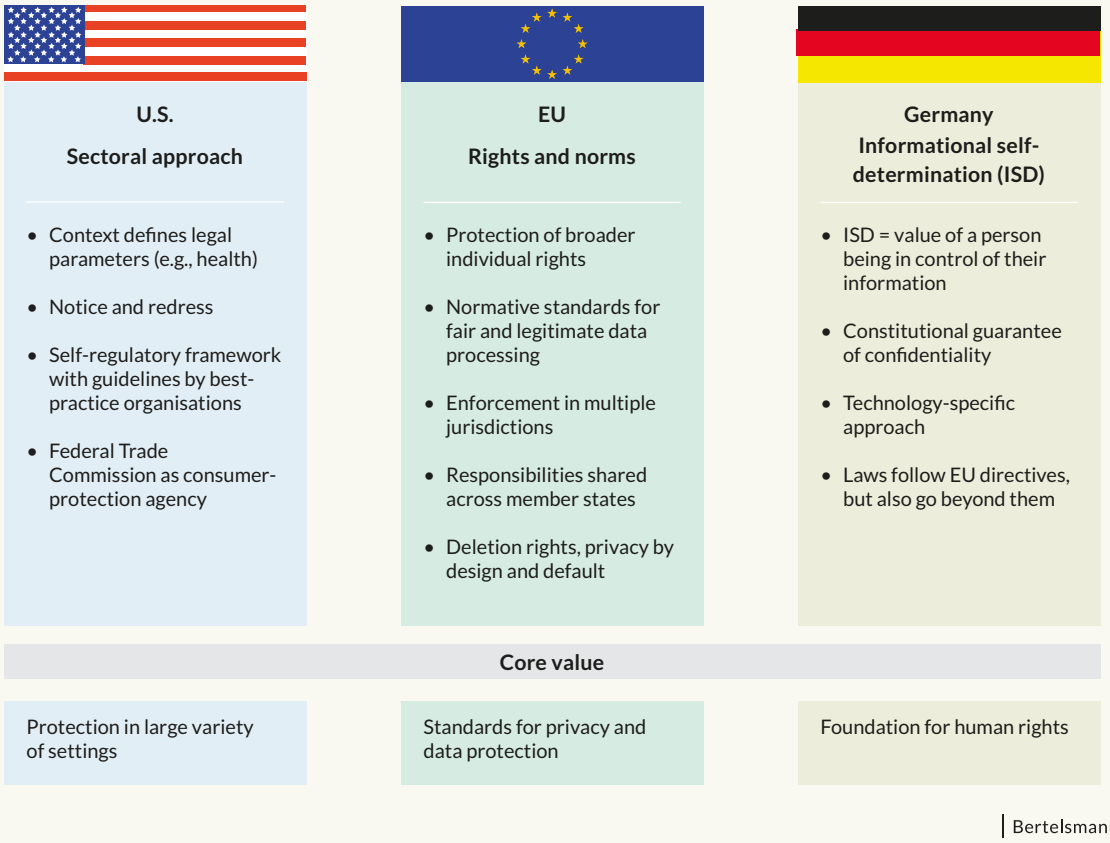
Though U.S. and EU legal foundations are similarly anchored in individual control, the interpretation and application of these concepts in law have been very different (Figure 1).

1. United States: A sectoral approach

American privacy laws are sector-specific, meaning they use the context of how and where the data is moving to define relevant legal parameters. Laws such as the Health Insurance Portability and Accountability Act (HIPAA) of 1996 and the Gramm-Leach-Bliley Act (GLBA) of 1999 cover data uses in narrow contexts, such as health and financial information respectively. Some privacy laws in the United States have been instituted as a reaction to current events, such as the Video Privacy Protection Act of 1998, which was enacted after contentious confirmation hearings for Supreme Court nominee Robert Bork. Others originated in states such as California and Texas, which have consistently legislated on an array of privacy laws, recently producing provisions on social media in schools and the confidentiality of personal information in mobile health apps.

Many key U.S. privacy laws draw upon concepts of individual control to regulate data collection. The Children's Online Privacy Protection Act (COPPA) and the Privacy Act are two examples of laws that require disclosure of data practices (notice) and give consumers the right to access

FIGURE 1 **Overview: Characteristics of data protection in the U.S., EU and Germany**
(based on Code of Fair Information Practices)



and correct personal data (redress).[8] But these laws offer only partial protection for personal information, and are often inconsistent in applying important components of the FIPs; for example, COPPA restricts data practices only for operators of websites and online services aimed at children under the age of 13.[9]

The Privacy Act of 1974 was the first law anywhere to meet the FIPs' standards. It covers the dissemination, maintenance, use and collection of personally identifiable information held by federal agencies.[10] However, the scope of the act is limited to personal information held in a system of records, which is a group of information retrieved through the use of a personal identifier (such as a Social Security number). This distinction has rendered the law outdated; the Privacy Act does not cover agency activities such as collecting and analyzing large data sets of personal information unless an individual identifier is used to get at the information. In modern computing, large data sets can be easily retrieved and organized without the use of such an identifier. In addition, agencies can access and use large data sets of personal information held by private entities without being subject to the act's requirements. These loopholes have eroded the power and undermined the intention of the FIPs in today's big-data world.

HIPAA is another example of a sector-based law in the U.S. that has struggled to harmonize its reliance on individual control with the emergence of new technologies and data

8 Other FIP-style privacy laws in the U.S. include: the Privacy Act of 1974, the Family Educational Rights and Privacy Act of 1974, the Right to Financial Privacy Act of 1978, the Cable Communications Policy Act of 1984, the Electronic Communications Privacy Act of 1986, the Employee Polygraph Protection Act of 1988, the Video Privacy Protection Act of 1988, the Telephone Consumer Protection Act of 1991, the Driver's Privacy Protection Act of 1994, the Children's Online Privacy Protection Act of 1998, the CAN-SPAM Act of 2003, and the Fair and Accurate Credit Transaction Act of 2003.

9 Federal Trade Commission Summary of Rule 16 CFR Part 312 COPPA: www.ftc.gov/enforcement/rules/rulemaking-regulatory-reform-proceedings/childrens-online-privacy-protection-rule

10 Privacy Act of 1974, U.S. Department of Justice. www.justice.gov/opcl/privacy-act-1974

uses. HIPAA applies to medical information held by covered entities, which include health-care providers, health-insurance plans, health-insurance exchanges and any business associates of these entities such as data processors and pharmacies.[11] HIPAA allows extensive sharing of personal health information if an individual provides consent, a process that typically consists of a person rapidly signing an unread form before a doctor's visit. Today's health-data ecosystem has rapidly outgrown HIPAA, as new devices and sensors have begun to collect and share a vast amount of sensitive health information, such as biometric and genetic information, outside the context of covered entities or other regulation.

The GLBA, which regulates privacy in the use of financial services, has also grappled with challenges to individual control deriving from the proliferation of big data. The law requires companies that offer consumers financial products or services, such as bank accounts, loans, investment advice or insurance, to provide notice of their data-collection and sharing practices. The GLBA also requires these entities to explain how customer information is used and secured, and requires individual consent for the sharing of such information with third parties.[12] In reality, these notices generally go unread by consumers and consent forms are hastily signed. Moreover, a growing number of companies that collect and share consumer financial information are not covered by the GLBA. Consumer-purchase histories are widely used to identify individuals' spending trends, target advertising to them and determine what prices to charge them. Such information is also used to prevent fraud and increase business efficiency. Generally, none of this collection and use is covered by the GLBA. The U.S. government has also found uses for consumer financial data, and is not itself subject to the GLBA. Documents released by journalists through Edward Snowden in 2012 uncovered "Follow the Money," a project run by the National Security Agency that secretly collects and analyzes worldwide financial data, mostly personal credit-card transactions.

The sector-based system in the United States is supplemented by guidelines issued by government agencies and industry organizations that function as best practices on data protection. While the law does not require these guidelines to be implemented, they create a unique self-regulatory framework that includes accountability and enforcement components, and are increasingly being used as a tool for enforcement by regulators such as the FTC.[13] The FTC is the U.S. consumer-protection agency responsible for policing privacy, and acts as a counter to the lack of comprehensive data-security and privacy laws in the country. However, the FTC has limited authority under Section 5 of the FTC Act, and must rely on an "unfair or deceptive" standard when it investigates commercial data practices. The agency's interpretation of this standard has centered on the company's intent to knowingly deceive or otherwise defraud customers, a focus that has led to strong emphasis on the issues of notice, choice and informed consent.

Although the FTC has had restricted ability to engage in robust enforcement,[14] it has arguably played an important role in shaping how we think about the role of the individual in the context of U.S. privacy regulation. For example, the agency has strongly advocated for commercial practices that facilitate individual control of personal information. In 2010, the FTC released a preliminary report on privacy that proposed a policy framework relying heavily on improved transparency, consumer education, and simplified settings and choices for data sharing. These principles were also at the heart of the agency's enforcement actions against Google and Facebook. The final version of this report, released in late 2012,[15] adopted a modified approach that placed greater emphasis on the context of the data transaction, implementing privacy by design and the need for further enforcement and accountability for commercial practices.

2. European Union: Rights and norms

The FIPs have been enshrined in law more broadly in the European Union than in the United States. Privacy and the protection of personal information have the status of distinct human rights, recognized by a multitude of legal provisions including the Treaty on the Functioning of the European Union (TFEU), the Charter of Fundamental

11 Data Protection in the United States, Thomson Reuters Practical Law. http://us.practicallaw.com/6-502-0467

12 Data Protection in the United States, Thomson Reuters Practical Law. http://us.practicallaw.com/6-502-0467

13 Thomson Reuters Practical Law. Data Protection in the United States. July 1, 2015. http://us.practicallaw.com/6-502-0467

14 Hartzog, Woodrow and Daniel J Solove. "The Scope and Potential of FTC Data Protection." George Washington Law Review 2230, November 1, 2015; GWU Law School Public Law Research Paper No. 2014-40; GWU Legal Studies Research Paper No. 2014-40. Available at SSRN: http://ssrn.com/abstract=2461096

15 Federal Trade Commission. "Protecting Consumers in an Era of Rapid Change." March 2012. www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf

Rights of the European Union (CFR),[16] and the 1995 EU Data Protection Directive (DPD), which created rights related to the "processing of personal data and … the free movement of such data." Most recently, these rights were reaffirmed in the General Data Protection Regulation (GDPR), which will come into force in 2018. The GDPR is a regulation, which unlike a directive, can be enacted at the EU level without the requirement that member states approve it via domestic legislation.

Regulators' attention in the EU has been focused on protecting broad individual rights as they relate to big-data issues such as data processing, data flows and the use of personal information. They also set normative standards such as the emphasis on "fair" and "legitimate" processing of data. Article 8 of the CFR, for example, establishes the principle that "personal data must be processed fairly and for specific purposes, based on the consent of the individual concerned or some other legitimate purpose laid down by law."[17] Article 8 also affirms the right of individuals to "access the data collected and the right to have it rectified, in case of inaccuracy or incompleteness."[18] In the EU, individuals have an absolute right to object to direct marketing, employers cannot read their employees' private email, and companies that process data must register with independent oversight agencies, the Data Protection Authorities (DPA). Authorities in the EU are also tasked with enforcing these broad protections in multiple jurisdictions. The TFEU, for example, grants every citizen the right to have his or her personal information be protected, and contains the specific legal basis for the "adoption of rules on data protection," while also "grant[ing] the authority to the EU bodies (Parliament and Council) to adopt rules concerning the processing of personal data by EU institutions, bodies, and member states, and [ensuring] that compliance with such rules is assigned to the control and review of independent authorities."[19] These rules have made it possible for the EU to regulate privacy rights across jurisdictions.

In 1995, the European Union passed the Data Protection Directive (DPD), a set of minimum standards for data protection that sought to regulate the processing of personal information. Among other protections, the directive safeguards an individual's personal information while regulating the flow of the data among EU member states, though its provisions are interpreted differently in the various member states.[20] At the time the directive was introduced, EU governing bodies were focused on promoting a common market for member states. The DPD made consumer protection a shared responsibility across member states, introducing incentives for the creation and enforcement of broadly applied data-protection rules aimed at reining in abuses that could wreak havoc on the fragile common market.

The role of individual control became more central to data-protection discussions in the EU after a 2012 recommendation from the Article 29 Working Party.[21] The Article 29 Working Party is composed of a representative of the supervisory authorities designated by each EU country, a representative of the authorities established for the EU institutions, and a representative of the European Commission. The Working Party's recommendation called for an increased emphasis on individual control while also advocating for broad-scope enforcement. Specific reforms recommended included: 1) increasing transparency by clarifying the data-minimization principle; 2) reinforcing a comprehensive scheme of responsibilities and liabilities for the controller – that is, the entity "determin[ing] the purposes and means of the processing of personal data"[22]; 3) requiring controllers and processors to implement a number of policies as well as technical and organizational measures to ensure data security; 4) requiring notification of the supervisory authority within 24 hours in the case of a personal-data security breach; 5) requiring that data subjects be notified if a breach could adversely affect individuals' privacy or personal data; and 6) imposing an obligation for controllers and processors to maintain documentation on all data-processing operations under

16 The Charter of Fundamental Rights of the European Union (CFR) recognizes a fundamental human right to privacy (in Article 7) as well as the right to protect one's personal data (Article 8).

17 Library of Congress. Online Privacy Law: European Union. Library of Congress. May 2014. www.loc.gov/law/help/online-privacy-law/eu.php

18 Library of Congress. Online Privacy Law: European Union, Library of Congress, May 2014. www.loc.gov/law/help/online-privacy-law/eu.php

19 Library of Congress. Online Privacy Law: European Union. Library of Congress, May 2014. www.loc.gov/law/help/online-privacy-law/eu.php

20 The directive has "resulted in diversity of implementation by the twenty-seven EU Members." Online Privacy Law: European Union, Library of Congress (May 2014). www.loc.gov/law/help/online-privacy-law/eu.php

21 Article 29 Data Protection Working Party. Opinion 05/2012 on Cloud Computing. May 2012. http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2012/wp196_en.pdf

22 Article 29 Data Protection Working Party. Opinion 1/2010 on the concepts of "controller" and "processor". Article 21 Data Protection Working Party. February 2010. http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2010/wp169_en.pdf

their responsibility.[23] Under each of these provisions, FIP-based individual-empowerment requirements are matched with liability for controllers of data.

In April 2016, the European Parliament, Council and Commission adopted the GDPR in an attempt to harmonize data-protection across EU member states and move toward plans for a Digital Single Market.[24] The GDPR will replace the 1995 DPD, and is designed to give individuals more control over their personal information through new data-processing requirements such as strengthened notice and consent provisions, improved transparency and access to data, deletion rights associated with the right to be forgotten, a new right to data portability, and new rules on law-enforcement access to citizen information. Many companies will be required to perform data-protection risk assessments, assign data-protection officers, and use "data protection by design" or "data protection by default" models when creating new products and services, though start-ups and small businesses may be exempt from some of these rules.

3. Germany: Informational self-determination

Germany was an early adopter of data-protection rules, passing one of the world's first data-protection laws, the Hesse Data Protection Act,[25] in 1970. The German data-protection regime places enormous importance on maintaining citizen confidentiality and ensuring the integrity of personal information, an approach influenced by privacy scholars in the United States.[26]

Laws in Germany are sector- and technology-specific, though the German Federal Data Protection Act (GFDPA) has separate provisions for data processing in the public and private sectors. The GFDPA addresses data processing within free and fee-based electronic information and communication services, while the German Telemedia Act (TMA) governs privacy in online services[27] such as search engines and social-media platforms. German laws generally follow EU directives on issues such as redress options, security requirements, restrictions on the retention of sensitive information and data minimization. Much as in the United States, privacy case law in Germany has shown variation across issues, reflecting the balance between commercial, government and individual interests; however, unlike most U.S. privacy law, this balance stems from a German statute that imbeds a principle of proportionality in gauging competing interests. Moreover, the country's laws do go beyond the EU's de minimis requirements at times. For example, German law transposes EU directives 95/46 on the Protection of Personal Data and 2002/58 on Privacy and Electronic Communications, strengthening transparency requirements and the ability for consumers to become aware of and exercise their privacy rights.[28] Some experts believe this complexity actually undermines the transparency requirement, as it keeps consumers from being aware of and successfully exercising their rights.[29]

While data policies in the United States and the European Union are founded on the principle of individual control, German data-protection law (while also conforming to EU law) is credited with inventing "the right of informational self-determination" or Recht auf Informationelle Selbstbestimmung.[30] Individual control and informational self-determination are different concepts, despite being rooted in a similar belief that privacy is a fundamental instrument for democracy. Individual control, as practiced in the United States and at the EU level, implies a constellation of separate legal rights regarding privacy and data protection, each operative within a particular data-type and data-usage context; by contrast, informational self-determination more broadly reflects the value of a person's ability to exercise control of her personal information. Informational self-determination laws are "assumed to protect human dignity and self-development."[31] German citizens have strong constitutional protections for their personal information that have been interpreted and are enforced by the Federal Constitutional Court (FCC). The German constitutional court has ruled that the right of informational self-determination permits

23 Article 29 Data Protection Working Party, Opinion 01/2012 on the Data Protection Reform Proposals, 00530/12/EN, WP 191 (Mar. 23, 2012). http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2012/wp191_en.pdf See Online Privacy Law: European Union, Library of Congress (May 2014). www.loc.gov/law/help/online-privacy-law/eu.php

24 http://ec.europa.eu/priorities/digital-single-market/

25 Hessisches Datenschutzgesetz (The Hesse Data Protection Act), Gesetz und Verordungsblatt I, 1970: 625.

26 The federal data-protection law in Germany drew heavily on the work of U.S. writers, particularly Alan Westin and Arthur Miller.

27 www.loc.gov/law/help/online-privacy-law/germany.php#_ftn1

28 Library of Congress. Online Privacy Law: Germany. Library of Congress, June 5, 2015. www.loc.gov/law/help/online-privacy-law/germany.php

29 Library of Congress. Online Privacy Law: Germany. Library of Congress, June 5, 2015. www.loc.gov/law/help/online-privacy-law/germany.php

30 Library of Congress. Online Privacy Law: Germany. Library of Congress, June 5, 2015. www.loc.gov/law/help/online-privacy-law/germany.php

31 Rouvroy, Antoinette and Yves Poullet. "The right to informational self-determination and the value of self-development. Reassessing the value of privacy for democracy." FUNDP.

the processing of personal data only if such activity is authorized by statute (while also recognizing that there are times when personal information must be used for the public interest). In 2008, the court expanded these principles by articulating a constitutional guarantee regarding the confidentiality and integrity of IT systems. In 2010, this court further struck down a transposition of the EU Data Retention Directive into German law, ruling that it violated the principle of proportionality and the individual's rights of personhood.

Advantages and shortcomings of approaches based on individual control

Each of these approaches to data protection – respectively in the United States, the European Union and Germany – has merits that have allowed them to stand the test of time, notwithstanding updates and changes in the law. The U.S. system provides protection for personal information in a large variety of settings, recognizes the differing sensitivity of data types, and has enabled active enforcement of privacy violations in recent years through the Federal Trade Commission. The EU has raised the bar for privacy and data-protection standards in such a way as to affect commercial entities worldwide, while Germany's legal precedents in the area of data protection have constructed a human-rights foundation for the concept of informational self-determination. However, each has faltered with the rise of big data and its impact on individuals' ability to control, see or use their personal information. As globalization has eroded the borders between geographic and industry sectors, the U.S. approach has become less tenable and enforceable, the EU's rights and normative standards have struggled to keep pace with the reality of large-scale data collection and use, and Germany's informational self-determination principle has been weakened by provisions giving authorities default access to personal information.

When the GDPR goes into full effect in May 2018, Germany, like all EU member states, will still retain key data-regulation responsibilities. There are 70 "opening clauses" in the GDPR that delegate certain data-protection responsibilities to the member states; these include the determination of legitimate grounds for public-sector data processing, the definition of rights for data subjects, and the incorporation of the freedoms of expression and information into law. Additionally, though the regulations are intended to harmonize laws across the EU, and include extensive details on consistency, they will be implemented by member states based on their own legal systems, and with reference to their own specific legal histories.

While legislative and enforcement structures relating to data protection differ between the United States, the European Union and Germany, there are fundamental commonalities between them that represent points of possible transatlantic engagement as new policies and laws are being developed.

III. The impact of big data on privacy self-management

The exponential growth and adoption of data-driven processes in all sectors has created huge digital catalogues of personal information that can be continually analyzed and categorized using machine-learning algorithms. This increase in data capture and processing – also known as big data – has given rise to advances in areas such as public health and city planning, but has also challenged the viability of economic paradigms such as the autonomy of individuals within the digital marketplace. The application of data analytics to large datasets has raised concerns regarding profiling, discrimination, economic exclusion and enhanced government surveillance, all of which threaten to undermine individual and group civil rights. Some of the most cutting-edge big-data technologies include: 1) predictive analytics software and/or hardware; 2) NoSQL databases, which are databases that offer quicker storage and retrieval of data by using different mechanisms than those employed by traditional relational databases; 3) stream-analytics software that can analyze and aggregate data from multiple live sources and across numerous formats; 4) the processing of ever-larger quantities of data thanks to the distribution of this data in random-access memory; and 5) systems that deliver information from various data sources, including big-data sources, to distributed data stores in real- or near-real time.

The combined rise of product and service personalization and the ubiquitous, interconnected devices of the internet of things (IoT) has further spurred the use of big data, and further undermined individuals' ability to manage their privacy effectively on their own.

1. Impact on individuals and groups

Individuals in the big-data marketplace are faced with enormous opportunities and obstacles in managing their personal information. The personalization of content, for example, has arguably been big data's greatest contribution to the commercial internet. Personalization is the process of targeting information and advertising to an individual based on personal characteristics such as demographics, location, purchase history and the device being used. This allows companies to design products and services that more closely suit a person's needs and interests, while giving individuals more leverage in comparing prices and products. At the same time, many people have a difficult time ascertaining when they are being tracked online and by which entities, what data is being collected and shared about them, and for what purposes the data is being collected. Notices about sharing have limited utility for most people, not only because they typically offer a "take it or leave it" approach to data sharing, but also because they tend to be written in a kind of legalese that is difficult for many to comprehend fully.

There is considerable debate as to whether individuals understand that their personal information is paying for the free services they are offered online. Behavioral scientists have argued that people commonly make decisions online under conditions describable as "bounded rationality"[32] – that is, rationality limited by numerous factors such as a lack of actionable information, the complexity of available choices and a lack of sufficient contextual cues to make a truly informed decision.[33] Each of these factors is exploited in the big-data environment. Faced with the decision to share or not share information with companies online, for instance, an individual must understand the sheer number of entities involved in such a transaction, as well as determine the consequences or benefits of sharing the personal information, without access to contextual information explaining exactly how this data will be merged or aggregated with other sources. These constraints make the management of personal information impractical, if not impossible, for most people.

32 Simon, Herbert A. Models of Man: Social and Rational. New York: John Wiley and Sons, Inc., 1957.

33 Acquisti and Grossklags, supra note 26, at 25-26.

The individual ability to manage personal data successfully is inexorably intertwined with the implementation of key FIPs such as purpose specification, a principle that restricts uses of data that are incompatible with the reasons for which it was originally collected. However, big-data environments are unwelcoming to this principle simply because there is often no defined reason or purpose for the collection to begin with other than tracking a person's online activities and preferences. Purpose specification has been interpreted as requiring entities to specify the purposes for which data is being collected up front, and as limiting future use to those specified purposes unless new consent is obtained. However, the analytic capabilities of big data often are "aimed precisely at… unanticipated secondary uses."[34]

In accordance with an individual-centric regulatory structure, modern users of digital technologies are presented with an illusion of control through consent notices and tools such as ad-preference options, although many people have become aware that there is a massive data-generation system at work behind the scenes of these apparent choices.[35] The disconnect between artificial and actual control has served to make many individuals in the United States and the European Union feel that they have lost control, with many simply becoming resigned to this condition.[36] We will discuss this feeling in more detail in the section on public opinion.

In general, the application of big-data analytics "can reproduce existing patterns of discrimination, inherit the prejudice of prior decision-makers, or simply reflect the widespread biases that persist in society."[37] Big-data analytics has thus incubated another potent weapon against individual autonomy and rational decision-making: the ability of a small group to be representative of larger populations. Analytics and algorithms used in processing employ modeling programs that match characteristics across multiple data sets, so the data collected on you tells a story that is not about you, but "someone like you." Inferences then become data classifications that are used to make assumptions about a person, creating a profile that may produce outcomes not necessarily in an individual's best interest. Because it is not necessary to obtain consent or establish an exact identity when making these assumptions, big data can create a "tyranny of the minority"[38] or an archetype wherein the small amount of individuals who choose to disclose detailed personal information implicate everyone else who happens to share observable correlating traits. Exacerbating this situation is the fact that hiding or masking one's true identity online has become increasingly difficult, even with de-identification techniques.[39] This type of representative analysis does not require a large amount of individually identified information in order to make inferences about groups of people, and can thus lead to a form of collective discrimination that has broad impact. Much of the use of algorithms and other forms of automated decision-making is opaque to individual users, making it difficult to assess bias or discrimination.

2. The rise of IoT devices

New sources for big data are rapidly developing as an increasing number of our devices are connected to the internet, complicating privacy self-management. Commonly referred to as the IoT, the increased connectivity of devices and sensors to the internet means that more and more information of varied kinds is being collected on individuals for commercial use from sensitive and perhaps unexpected places such as thermostats or refrigerators. IoT devices are often small and collect data in highly discreet ways. They are frequently generating and collecting data about individuals from multiple sources at once, often via continuous connections, evading ordinary constraints to data collection such as a person turning off a device.

34 Ira Rubenstein, Big Data: The End of Privacy or a New Beginning?, 3 Int'l Data Privacy Law 74 (2013). http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2157659

35 Hoofnagle, Chris Jay et al. "Privacy and Modern Advertising: Most US Internet Users Want 'Do Not Track' to Stop Collection of Data about their Online Activities." Amsterdam Privacy Conference, October 2012.

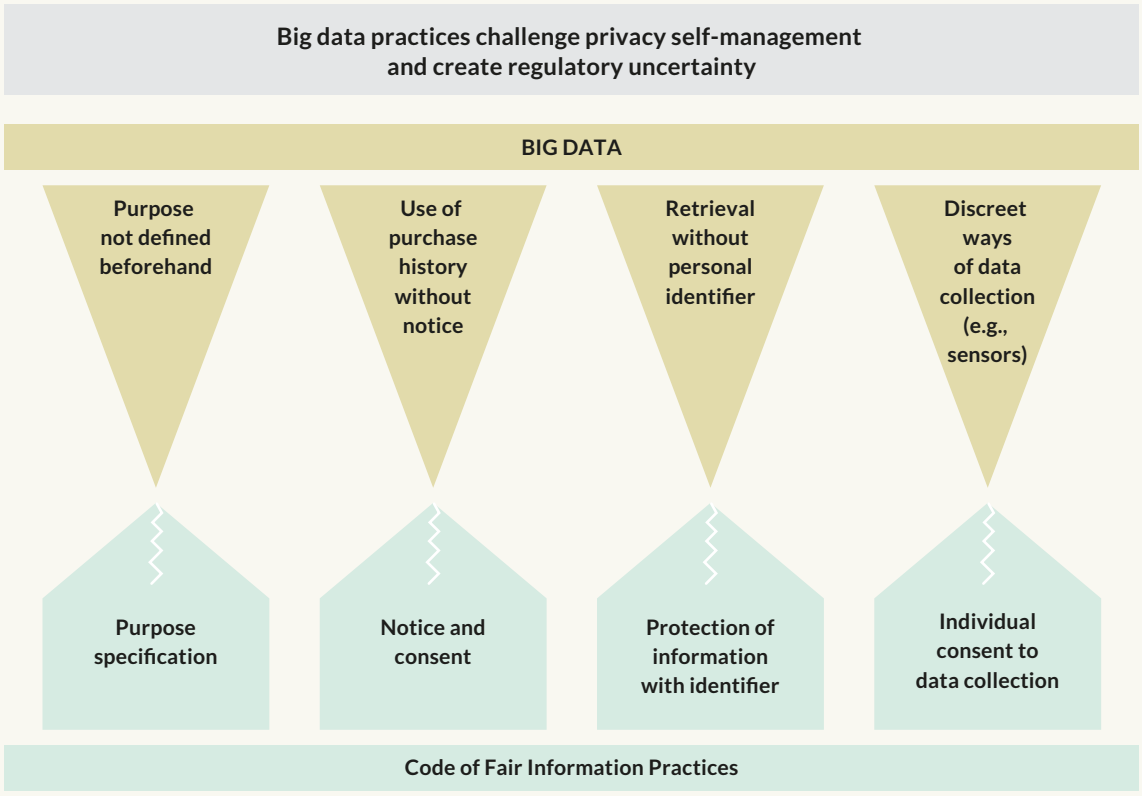
36 Draper, Nora, Michel Hennessy and Joseph Turow. "The Trade-off Fallacy: How Marketers are Misrepresenting American Consumers and Opening them up to Exploitation." Annenberg School of Communication at the University of Pennsylvania, June 2015: 3. www.asc.upenn.edu/sites/default/files/TradeoffFallacy_1.pdf

37 Federal Trade Commission. "Big Data: A Tool for Inclusion or Exclusion?" January 2016. www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf

38 Borocas, Solon and Helen Nissenbaum. Big Data's End Run Around Procedural Privacy Protections, Communications of the ACM, Vol. 57 No. 11, Pages 31–33 10.1145/2668897

39 See El Emam, Khaled and David Houlding. "Secure Analytics on the Cloud." Privacy Analytics, http://cdn.oreillystatic.com/en/assets/1/event/91/Secure%20Analytics%20on%20the%20Cloud%20Presentation.pdf; Catherine Tucker, Ethical and Legal Issues Arising from the Informed Consent Process in Fertility Treatments, 9 ABA Health eSource (March 2013); Narayanan, Arvid and Edward W. Felten, "No Silver Bullet: De-Identification Still Doesn't Work, Random Walker (July 9, 2014), http://randomwalker.info/publications/no-silver-bullet-de-identification.pdf; Omer Tene, People Like You, YALE JOURNAL OF LAW & TECHNOLOGY (Nov. 28, 2015), http://yjolt.org/blog/2015/11/28/people-you.

FIGURE 2 **Data practices create regulatory uncertainty and challenges for privacy self-management**



The novelty of the devices in this space also means that information about homes and daily lives are being collected by and shared with companies that are newly formed, and which may not have experience with safeguarding such data or offering privacy self-management options in the same way that more traditional and regulated entities such as utilities or health care providers might have. In addition, the variability of start-ups makes it uncertain whether a business may still be operating in a year or two. This associates all data collected by it with some elevated risk, whether from a security perspective or simply through the potential that the information may later be sold as a business asset to other entities. Furthermore, the constellation of devices that make up the IoT has rendered data collection less obvious to the common user due to limiting factors, such as smaller screen sizes, that make it difficult for the device to communicate useful data-management information to the user.

It is unclear what disclosures would give individuals using this type of device sufficient clarity regarding the information the device is collecting and what is to be done with the data collected.[40]

Thus, the rise of the internet of things adds further opacity and complexity to the nature of big data. Through these characteristics, along with the unpredictability regarding the ultimate purposes of data usage big data practices interfere with central principles of the FIPs and individual control, which form the basis of existent data protection regimes (Figure 2).

40 Sweeney, Latanya, et al. "Identifying Participants in the Personal Genome Project by Name." Harvard College. http://dataprivacylab.org/projects/pgp/1021-1.pdf

IV. Public opinions on big data, individual control and privacy

Public values and perceptions of data contribute to the development of legal and regulatory privacy frameworks, and individuals often have subjective considerations when it comes to their views on privacy. In many ways, an individual's valuation of privacy in general is difficult to measure, as it depends on the social value of the information to be disclosed versus the value of such information remaining private. But cultural and historical factors also play an important role in how privacy rights are understood, and in whether the individual or government agencies are regarded as the best guarantors of these rights. Cultural factors "become intertwined with and exert a significant influence over differing legal environments."[41] Across cultures and geographies, there is a shared belief about the importance of the internet, but there are also questions regarding the use of personal information in big-data applications, the importance of access to and use of the internet, as well as the advantages that accrue when big-data mechanisms are used responsibly. Many people share the view that they have lost control of their personal information in digital systems, and do not know how to regain such control.

1. Transatlantic commonalities

Views on data protection in the European Union, the United States and Germany are very similar, with individuals in all three places expressing a belief that the internet brings value to their lives, while also feeling trepidation regarding the collection and use of their personal information (Figure 3).[42]

The American public has long expressed deep concerns regarding the privacy of their information in automated systems. A 1973 report from the Electronic Privacy Information Center noted that "[Americans'] worries and anxieties about computers and personal privacy show up in the replies of about one-third of those interviewed."[43] Similarly, those from the European Union and Germany have indicated a mistrust of large-scale data processing for decades.[44]

Less than a third of European respondents in one survey believed that there were advantages to big data, while less than a quarter thought that companies respected the privacy of users' personal information.[45] In Germany, 56 percent of respondents said that they deliberately avoided including "personal information in emails and text messages, because they fear the privacy implications."[46] In both the United States and the European Union, individuals have expressed feeling defeated and resigned over their inability to control their personal information, as well as a strong desire to decide how their information is shared and used.[47]

In a Harvard Business Review survey that included interviews with individuals from the United States, the

41 Baumer, David, Julia B. Earp and J.C. Poindexter. "Internet Privacy Law: A Comparison Between the United States and the European Union." Computers & Security, Vol. 23, No. 5: 400–412, July 2004. http://ssrn.com/abstract=1823713

42 Dutta, Soumitra, William H. Dutton and Ginette Law. "The New Internet World: Perspective on Freedom of Expression, Privacy, Trust and Security." April 2011. http://ssrn.com/abstract=1916005

43 Secretary's Advisory Committee on Automated Personal Data Systems. Records, Computers, and the Rights of Citizens. Dept. of Health, Educ. and Welfare, July 1973, available at www.epic.org/privacy/hew1973report/

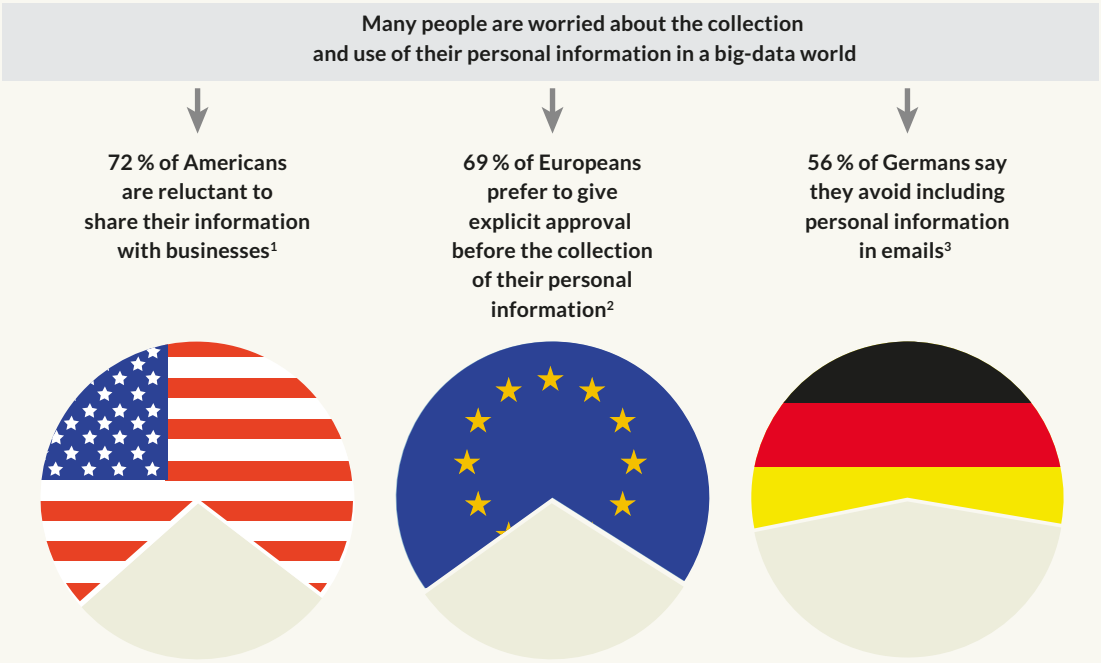
44 Alvar C.H. Freude & Trixy Freude, Echoes of History: Understanding German Data Protection, Bertelsmann Foundation Newpolitik

45 David Meyer. "Europeans Remain Far from Sold on the Benefits of big data." Fortune, Jan. 18, 2016. http://fortune.com/2016/01/18/europe-data/

46 David Meyer. "Europeans Remain Far from Sold on the Benefits of big data." Fortune, Jan. 18, 2016. http://fortune.com/2016/01/18/europe-data/

47 (Vodaphone, Turow, Pew) A 2015 survey of 1,506 Americans age 18 and older found that 91% disagreed with the statement: "If companies give me a discount, it is a fair exchange for them to collect information about me without my knowing." In the same survey, 71% disagreed that: "It's fair for an online or physical store to monitor what I'm doing online when I'm there, in exchange for letting me use the store's wireless internet, or Wi-Fi, without charge."

FIGURE 3 **Many people are worried about the collection and use of their personal information in a big-data world**



1) Forbath, Theodore "Theo", Timothy Morey and Allison Schoop. "Customer Data: Designing for Transparency and Trust." Harvard Business Review, May 2015. Available at: https://hbr.org/2015/05/customer-data-designing-for-transparency-and-trust.
2) Eurobarometer. "Data Protection." Special Eurobarometer 431, March 2015, http://ec.europa.eu/public_opinion/archives/ebs/ebs_431_en.pdf
3) David Meyer. "Europeans Remain Far from Sold on the Benefits of big data." Fortune, Jan. 18, 2016, http://fortune.com/2016/01/18/europe-data/.

BertelsmannStiftung

United Kingdom and Germany, 80 percent of Germans and 72 percent of Americans were reluctant to share their information with businesses because of a desire to maintain personal privacy.[48] The survey also found that Germans placed more value on their personal information than did British and Americans. Health and credit-card information, as well as the details of government-issued credentials, were the most highly valued category across the various countries, with location and demographic information among the least-valued categories. A total of 97 percent of the people surveyed expressed a concern that businesses and the government might misuse their data.

While legislative and enforcement structures for data protection in the United States and the European Union differ, both regions have recognized the growing importance of data and technology in everyday life, as well as the importance of thoughtful regulation that protects citizen interests. Likewise, the publics in the U.S. and the EU have expressed similar apprehension over privacy and the lack of individual control or transparency. A Eurobarometer survey in 2015[49] revealed that 69 percent of Europeans would prefer to give their explicit approval before personal information is collected and processed.

2. Transatlantic differences

Attitudes on privacy and individual control as captured by surveys of Americans and Europeans also reflect cultural narratives. Generally, Americans are more enthusiastic about the notion of an individual being in control of his or her personal information, according to several surveys performed by the Pew Research Center.[50] Perhaps this is related to the mythology of the American Dream, which centers on an ideal of a self-made person. Capitalist theory espouses reliance on individual initiative in an environment

48 Forbath, Theodore, Timothy Morey and Allison Schoop. "Customer Data: Designing for Transparency and Trust." Harvard Business Review, May 2015. https://hbr.org/2015/05/customer-data-designing-for-transparency-and-trust

49 Eurobarometer. "Data Protection." Special Eurobarometer 431, March 2015. http://ec.europa.eu/public_opinion/archives/ebs/ebs_431_en.pdf

50 Rainie, Lee. The State of Privacy in America, Pew Research Center, September 2016. www.pewresearch.org/fact-tank/2016/09/21/the-state-of-privacy-in-america/

characterized by competition and conflict resolution taking place between private entities free of government interference. Germans and those from the European Union typically espouse a more collective viewpoint - here, there is less suspicion regarding the role of the government in protecting privacy, as governments are generally viewed as the guarantor of the common good.[51]

In surveys, non-U.S. respondents were statistically more likely to be concerned about private organizations using personal information for customization and personalization purposes than about government data collection. Across the European Union, there is broad skepticism and mistrust regarding the use of personal information in big-data systems. In a 2016 Vodafone survey,[52] the majority of respondents strongly disapproved of personal information being passed on for commercial purposes, irrespective of the reason or type of data. By contrast, a 2016 Pew survey showed that many Americans believe that privacy is less a "condition of life," which would ostensibly point to the value of government-led data protection, and more "a commodity to be purchased," placing privacy in a more commercial framework.[53]

Concerns for personal privacy and security resonate strongly in Germany, with 70 percent of German mobile internet users agreeing that accessing the internet on the go creates a risk that their personal information could be accessed.[54] However, despite concerns about privacy, German citizens are keen users of social networks and the internet in general. Of those owning a smartphone (80 percent) or a tablet (46 percent), many use those mobile devices as their primary route to the internet (84 percent of smartphone users and 62 percent of tablet users). More than one-third (35 percent) of smartphone internet users spend more than one hour each day on the mobile web.[55]

51 Baumer, David L., Julia B. Earp and J.C. Poindexter. Internet Privacy Law: A Comparison between the United States and the European Union. College of Management, North Carolina State University, Raleigh, NC 27695-7229.

52 www.vodafone-institut.de/wp-content/uploads/2016/01/VodafoneInstitute-Survey-BigData-en.pdf

53 www.pewresearch.org/fact-tank/2016/01/20/the-state-of-privacy-in-america/

54 "Germany Posts Big Increase in Mobile Internet Use." eMarketer, Jan. 22, 2015. www.emarketer.com/Article/Germany-Posts-Big-Increase-Mobile-Internet-Use/1011884

55 "Smartphones and Tablets Drive Internet Use in Germany." eMarketer, Apr. 6, 2016. www.emarketer.com/Article/Smartphones-Tablets-Drive-Internet-Use-Germany/1013757

3. Opinions on specific issues

To further compare and contrast the views of the publics in the United States, Germany and the European Union, the following section offers a snapshot of public opinion on specific data-protection issues currently being debated.

Right to be forgotten

The European Union and the United States are in sync when it comes to the right to be forgotten (RTBF), though less so regarding the operationalization of this right. While a 2014 survey found that 61 percent of U.S. residents supported the RTBF in general,[56, 57] only 39 percent wanted a European-style blanket RTBF, without restrictions.[58] The idea of reputational harm resonated in both geographies, with nearly half the respondents expressing a concern that "irrelevant" search data could do damage to an individual's social standing.[59] According to one survey, many Americans, echoing common viewpoints in the European Union, felt that "the appeal of the [right to be forgotten] law is not... based on fears of the negative consequences of search results - but rather, is based on a belief in the individual's right to privacy."[60]

Online privacy and control of personal information

Individuals on both sides of the Atlantic and across the web feel they have lost control over the way their personal information is collected and used.[61] Many respondents in the United States feel that "there's not much we can do to find out which aspects of our personal lives are being bought and sold by data brokers."[62] Some 91 percent of

56 Attitudes on Data Protection and Electronic Identity in the European Union, European Commission 7 (June 2011). http://ec.europa.eu/public_opinion/archives/ebs/ebs_359_en.pdf

57 Humphries, Daniel "US Attitudes Toward the 'Right to Be Forgotten.'" Software Advice, Sept. 5, 2014. www.softwareadvice.com/security/industryview/right-to-be-forgotten-2014/.

58 Kemp, Cheryl. "61 Percent of Americans Support the Right to Be Forgotten as California Enacts New Law." The Whir (Sept. 29, 2014). www.thewhir.com/web-hosting-news/61-percent-americans-support-right-forgotten-california-enacts-new-law

59 Humphries, Daniel "US Attitudes Toward the 'Right to Be Forgotten.'" Software Advice, Sept. 5, 2014. www.softwareadvice.com/security/industryview/right-to-be-forgotten-2014/

60 Humphries, Daniel "US Attitudes Toward the 'Right to Be Forgotten.'" Software Advice, Sept. 5, 2014. www.softwareadvice.com/security/industryview/right-to-be-forgotten-2014/

61 Madden, Mary. Public Perceptions of Privacy and Security in the Post-Snowden Era. Pew Research Center, Nov. 12, 2014. www.pewinternet.org/2014/11/12/public-privacy-perceptions/

62 Lafrance, Adrienne. "Why can't Americans find out what big data knows about them?" The Atlantic (May 28, 2014). www.theatlantic.com/technology/archive/2014/05/why-americans-cant-find-out-what-big-data-knows-about-them/371758/

Americans did not feel that they had much control over the collection of their personal data and were not confident that their data will be handled as private data by companies and remain secure.[63] In the, EU 70 percent (along with 82 percent of online shoppers and 74 percent of social-network users) felt that they did not have complete control over their personal information,[64] that companies were not straightforward about their data practices, and that consumers had "only partial, if any, control of their own data."[65]

As much as 86 percent of users in the United States have taken steps to cover their digital footprints, with most individuals saying they want to do more to protect their data online, but lack the means to be anonymous online.[66] EU residents are also concerned about their online privacy, and were more likely to have used technical or procedural means to protect it, such as implementing tools and strategies to limit unwanted emails (42 percent), checking that an electronic transaction is protected on the site (40 percent), or using anti-spyware software (39 percent).[67] A total of 62 percent of EU respondents also said they provide only the minimum amount of information required online in order to protect their identity.[68]

Trustworthiness of governments and companies

Current events also play a part in shaping public opinions about privacy. For example, U.S. residents said in 2015 that they were less likely to trust their government's privacy and security practices and more likely to trust commercial entities,[69] perhaps reflecting fallout from revelations in 2013 that the U.S. government was secretly conducting surveillance on its residents. EU residents, on the other hand, showed higher rates of trust in government and less in companies, which could be attributed to increased attention to and enforcement of data-protection laws. Fully 55 percent of EU respondents indicated trust in the privacy and security practices of the European Commission and the European Parliament, a figure that greatly outweighed their confidence in similar practices by commercial entities.[70]

63 Madden, Mary and Lee Rainie. Americans' Attitudes About Privacy, Security and Surveillance. Pew Research Center, May 20, 2015. www.pewinternet.org/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/

64 European Commission. "Attitudes on Data Protection and Electronic Identity in the European Union." Special Eurobarometer 359, European Commission, June 2011. http://ec.europa.eu/public_opinion/archives/ebs/ebs_359_en.pdf

65 European Commission. "Data Protection: Europeans Share Data Online, but Privacy Concerns Remain — New Survey." European Commission press release, June 16, 2011. http://europa.eu/rapid/press-release_IP-11-742_en.htm

66 Madden, Mary. Public Perceptions of Privacy and Security in the Post-Snowden Era. Pew Research Center, Nov. 12, 2014. www.pewinternet.org/2014/11/12/public-privacy-perceptions/

67 European Commission. Attitudes on Data Protection and Electronic Identity in the European Union. Special Eurobarometer 359, European Commission, June 2011. http://ec.europa.eu/public_opinion/archives/ebs/ebs_359_en.pdf

68 European Commission. Attitudes on Data Protection and Electronic Identity in the European Union. Special Eurobarometer 359, European Commission, June 2011. http://ec.europa.eu/public_opinion/archives/ebs/ebs_359_en.pdf

69 Madden, Mary and Lee Raine. Americans' Attitudes About Privacy, Security and Surveillance. Pew Research Center, May 20, 2015. www.pewinternet.org/2015/05/20/americans-attitudes-about-privacy-security-and-surveillance/

70 European Commission. Attitudes on Data Protection and Electronic Identity in the European Union. Special Eurobarometer 359, European Commission, June 2011. http://ec.europa.eu/public_opinion/archives/ebs/ebs_359_en.pdf

V. Elements for regulating privacy through data sovereignty in the big-data era

Data is an incredibly valuable resource for both the public and private sectors, and is becoming increasingly so thanks to big-data technologies'[71] ability to process large amounts of data quickly and cheaply. The benefits to society of collecting and processing large amounts of personal information are numerous, including advances in public health, education, safety and welfare. But the generation of such vast amounts of personal information poses real risks, such as the revelation of information about individuals that they wish to keep private because it is harmful or embarrassing; the facilitation of discrimination and profiling not in the interest of the individual; identity theft; and extortion.

Big data poses unique challenges to notions of individual control as they currently exist in U.S. and EU law and policy. The legislative regimes built on concepts of individual control are ill-equipped to deal with the consumer-protection challenges of big data, automation and predictive analytics. Historically, U.S. and EU data-protection laws rely on an aspirational vision of consumers, imagining them as rational actors whose engagement in the political process and economic marketplace is centered on and will produce self-interested and desirable outcomes.[72] This vision assumes the individual will balance sometimes-competing interests and values. In the opaque and highly complex online system of big data, in which an individual's consent allows for the collection, use and processing of personal information, this vision seems unreasonable.

Current U.S. and EU data-protection regimes based on the FIPs have few effective mechanisms for limiting the unfettered production, collection and use of personal information. New technologies and uses of data have raised important data-protection concerns that have yet to be answered in policy or practice.

To address these issues, we present three concepts of privacy regulation below, and consider how they might address the challenges of privacy self-management in the big-data ecosystem. We initially discuss these ideas individually in order to assess their value in bolstering data-protection regimes in today's big-data world, though each has an intrinsic interdependency with the core concept of data sovereignty and portability. The concepts are:

1) **Individual empowerment through education and data portability;**
2) **Corporate accountability through industry self-regulation; and**
3) **Collective accountability through the use of legally mandated impact assessments.**

Our analysis of these policy concepts is not meant to promulgate one approach or another; rather, it is a way to consider how each might improve individuals' ability to control their personal information, while reducing the potential for individual harm resulting from big-data processing. The idea of data sovereignty is at the core of this analysis, replacing outdated notions of individual control, but this cannot be effectively realized without also considering how corporate and collective accountability might function under this kind of regime. In this way, the three concepts intertwine and supplement each other (Figure 4). Each concept's weakness can be compensated for by another's strength. The concept of education and data portability empowers the individual but also demands substantial individual effort. This burden is lessened when responsibility for data protection is additionally placed in the hands of companies through a system of industry self-regulation. Yet companies' assessments are often not as

71 According to "TechRadar: Big Data," a Forrester Research report released in Q1 2016, some of the latest big-data technology includes predictive analytics software and hardware. www.forrester.com/report/TechRadar+Big+Data+Q1+2016/-/E-RES121460

72 A survey conducted by Turow et al. concludes that the image of an "informed consumer" cannot be supported, because consumers lack basic knowledge regarding the functioning of data aggregation and profiling.

FIGURE 4 **Data sovereignty in the big-data era**

BertelsmannStiftung

accountable or forceful as required. At this point, collective responsibility comes into play, specifically through the performance of impact assessments by independent third parties; these assessments would be mandated by the state, would serve to advocate for users' interests, and would create more transparency for users. These third-party authorities too would relieve individual users of some responsibility; however, this could lead to patronization of users. In order to avoid this outcome, individuals need education that allows them to understand the information created through assessments and use it in a self-determined way. Thereby the circle is complete.

Ultimately we seek to create an effective policy framework that helps both to seize the opportunities provided by data technologies and to protect privacy and civil liberties, thereby increasing public trust in online-data systems. Surely an overall legal framework and governmental regulation are also needed. But since technology develops more swiftly than laws, legislative and policy-based approaches can only be put into practice effectively in an environment of corporate and collective accountability.

1. Individual empowerment: Education and data portability

Problems with individual control in the context of big data are numerous, and many stem from a fundamental disconnect between the information accessible to individuals regarding the likely uses of their personal information and the actions they can take to protect such information. Concepts of data sovereignty and data portability offer levels of individual empowerment that

could close this disconnect. Data sovereignty,[73] for the purposes of this paper, refers to the legal right of an individual to maintain control over the possession, use and deletion of their personal information, subject to the laws of the jurisdiction in which the individual resides. Moreover, it implies empowering individuals to exert this control in practice through education initiatives that teach basic technology and data-management skills. Data portability[74] is the right of an individual to move his or her personal information between online locations without loss or distortion (we do not include jurisdictional and data-flow questions in this definition).

Benefits

Implemented in conjunction with one another, data sovereignty and data portability would ideally facilitate increased engagement with data-management tasks, allowing people to determine how, when and for what purposes their data are used. It would give individuals authoritative legal rights over the data, with these rights traveling as the information moved. Data sovereignty also refers to the ownership of and responsibility for information. Proponents of this concept believe it offers a way to give people a power of self-determination with regard to their information in big-data systems, leveling the playing field between individuals and the commercial and noncommercial entities that capture and share their information. In this way, data sovereignty mirrors some of the concepts in individual control, because it implies the ability to "access, create, modify, package, derive benefit from, sell or remove data, but also the right to assign these access privileges to others."[75] Overall, the ideas of data authority and portability are appealing to many people who envision a system in which they have complete control over the use or removal of their personal information.

Data sovereignty does appear to address the concerns of the individuals who consistently state in public-opinion polls that they feel powerless and resigned to the ubiquitous collection and use of their data. Ownership is a formidable way to empower individuals. Offline, however, property laws typically do not confer absolute rights of control on owners; instead, they typically impose both limitations and duties. Ownership as interpreted by this concept of data sovereignty would encompass a set of rights and responsibilities aimed at protecting the interests of individuals and other entities with a stake in ownership. To achieve this standard of ownership, data sovereignty would have to empower individuals to manage their own data, while also granting businesses some ability to "lease" or "rent" data.

Likening data ownership to traditional property ownership is well-trodden ground. However, there are many alternative, arguably more nuanced, models to consider in determining how to implement data sovereignty. In practice, for example, data ownership might resemble co-ownership, in which multiple parties have rights of access and use. Data sovereignty in practice could also resemble the nonexclusive rights held by riparian owners with regard to the river next to their land - that is, a right to use the river, but without the right to interfere with others' simultaneous uses such as fishing and navigation. Alternately, it could work like a copyright that expires after a period of time and allows fair use by others during the period of protection. Though the European Union and Germany have data-ownership laws, the United States does not. Complicating any implementation of data sovereignty in the United States would be the fact that most property law in the country is set at the state level, except under limited circumstances.

Data portability would allow individuals to move their personal information at will, and thus is complimentary to ownership regimes as a method of creating a more level playing field between individuals and businesses in a big-data world. The global move toward cloud computing is part of a computing-industry progression in which the locus of data storage has shifted from centralized mainframes to personal computers and finally to a "cloud" made up of remote, often geographically distributed servers connected to the internet, and thus accessible by personal computers. Each step in this progression has increased pressure on governments to create effective data-portability policy.

Regulations on data portability are currently focused on creating seamless transfers and facilitating data storage based on an individual's wishes. For example, Article 20 of the European Union's Data Protection Regulation (DPR) created the right to data portability. The "rectification and erasure" section of the DPR is a part of Chapter III, "Rights of the Data Subject," encompassing Article 16, the right to rectification; Article 17, the rights to be forgotten and to

73 A fundamental right and an individual's ability to maintain transparency and control over the possession, use, or deletion of one's personal data, subject to the laws of the jurisdiction in which the individual resides.

74 The ability for people to move their data across interoperable applications and to control their identity, media and other forms of personal data.

75 Loshin, D. "Knowledge Integrity: Data Ownership." 2002. http://ori.dhhs.gov/education/products/n_illinois_u/datamanagement/dotopic.html

erasure; and Article 18, the right to data portability. Article 18(1) states that the data subject has "the right, where personal data are processed by electronic means and in a structured and commonly used format, to obtain from the controller a copy of data undergoing processing in an electronic and structured form which is commonly used and allows for further use by the data subject."

In 1983, the German Federal Constitutional Court ruled[76] that whoever "cannot survey with sufficient assurance the information concerning himself known in certain areas of his social surroundings, and whoever is not in a position to assess more or less the knowledge of possible partners in communication, can be essentially obstructed in his freedom to make plans or decisions on the basis of his own self-determination." In this decision, the court identified elements related to the principle of self-determination, such as the context in which an individual accesses his or her data, or the restrictions imposed on the purpose or use of the data, which it deemed essential if the concept was to be meaningful. Some experts have argued that data portability and data-sovereignty rights must be combined in order to avoid governmental paternalism; under this model, a policy framework might seek to treat individuals as truly sovereign with respect to their own data, while also providing for portability.

The creation of a data marketplace in which individuals have a relationship of sovereignty to their personal data, with the ability to move data at will, is one way these concepts might work in practice. The data-marketplace idea has been proposed numerous times over the years. One scholar has dubbed this a "National Information Market" (NIM)[77]; it follows an economic model under which individuals would sell personal information only if they were offered an acceptable price (one equal or greater than the value of not releasing the information). Under the NIM scenario, individual considerations and valuations of personal privacy would function as a limiting factor on the market, as buyers would also be determining whether the social value of the access to the information they hope to purchase was worthwhile.[78]

Challenges

The viability of the data-sovereignty and data-portability concepts depends on whether concerns regarding the ability of an individual to engage rationally in such a marketplace can be addressed. Any framework that emphasizes these two concepts must build or require the creation of technical tools for data management (such as centralized data depositories that are regulated by industry or government, as well as improved user-interface designs), implement consistent education and outreach programs aimed at improving individuals' capacities to navigate data choices, and finally create policy levers that allow individuals to negotiate fair terms for the use of their data.

However, data sovereignty and portability are not panaceas for every challenge that stems from big data, and could end up creating new problems for individuals and businesses. Granting people data property and movement rights could lead to a removal of information that might otherwise benefit the public, as well as end up generating higher costs for consumer transactions and services online. Implementing policies that create data sovereignty would require governments to commit resources to educating individuals on data-management issues and the rules of ownership. Complicating matters further, it is likely that most people would be reluctant to devote themselves to taking the time and learning the skills required to manage data effectively. As Jonathan Obar writes in his 2015 paper "Walter Lippmann and the Fallacy of Data Privacy Self-Management," what people "desire is the freedom to pursue the ends of digital production, without being inhibited by the means." A peripheral but important component to data portability is the establishment of security protections for the data as it moves or rests in its various repositories. Even with strong security protocols, assigning lifetime data-portability rights to individuals raises the stakes for concerns such as identity theft. Moreover, the sovereignty and portability approach does not address the hugely important issue of data that refers to a collective rather than solely to individuals, such as data in a social graph that describes multiple people. Big-data analytical tools raise this question with particular urgency, as their ability to carry out statistical analyses enables conclusions to be drawn about members of large groups who have not given consent to data collection, based on analysis of a smaller group's behavior.

Additionally, it is difficult to imagine how data-portability and data-sovereignty laws would function in all of today's current legal settings; for example, how would portability

76 29 Bundesverfassungsgericht, judgement from Dec. 15, 1983. Decisions of the German Federal Constitutional Court 65, 1, 41.

77 Laudon, K. C. "Markets and Privacy." Communications of the ACM 39 (9), 1996: 92–104.

78 Mungan, Murat. "Conditional Privacy Rights." April 16, 2016.

square with antitrust law in the United States? Large internet companies like Facebook, which have already established a brand and hold data relating to as many as a billion people worldwide, would be less affected by users' ability to move their information from place to place than would small business operators. This in turn could result in less competition and diminished choice for individuals. Also, in the United States, physical possession is a primary criteria for a considerable amount of ownership law; this would become greatly complicated when applied to data stored in a centralized repository or in the cloud. In this case, ownership rights would depend not on possession, but on the ability to restrict or deny access, which could end up being duplicative to existing law. In the European Union, the contents of databases are afforded strong legal protections akin to ownership rights, though caveats in the law give third parties some access to personal information by default (though not a "substantial" amount). To be workable in this environment, data sovereignty and portability would have to amend laws or be complemented by policies that enable a person to determine how much of their personal information a data controller can see, both in specific circumstances and by default.

Another consideration is that giving individuals more control over their personal information could indirectly impact the fairness of data analytics, resulting in "cumulative disadvantage" due to the narrowing of possible categories or results.[79] In this case, it is irrelevant whether the information used were to be obtained with consent or not. Nor would giving a person more ability to control their personal information avoid "filter bubbles,"[80] or data distortions that stem from self-selected uses of data. These bubbles create feedback loops that reinforce stereotypes and existing worldviews, making the world smaller rather than opening it up to the possibilities of big data.

Although big data drastically limits self-determination at the moment at which the data is collected, the fundamental right of any individual to make decision regarding his or her own personal information cannot be eliminated. Individuals should have the right to be informed when their data is to be subject to processing, as well as the right not to take part in this.[81] Any commercial application of this argument would therefore imply the creation of technical controls enabling individuals to control uses of their personal information that were much more robust than those currently existing. Public surveys have made it clear that that companies which are transparent about the data they gather, give individuals control over their data, and offer fair value in return for the use of such data will be trusted and rewarded to ongoing and even expanded access to customer information.

In a policy framework centered on data ownership and portability, the government would need to implement education programs providing individuals and businesses with the appropriate tools to maneuver in the new data landscape. For example, policymakers would have to ensure fair competition by helping businesses build the technical capacity to evolve from a model under which user data is "locked in" to one in which user data is held for a time and then released. As referenced above, many internet companies today rely on individuals keeping their information in one place, using fees or other deterrents to make moving the data difficult. Government-run education programs could help individuals and companies understand how to port their data to other sites, what security risks and protocols to consider, and what compatibility issues might occur when moving data.

Helping the public understand data-processing practices and data-ownership rights, as well as their implications, should be in part the responsibility of the government, perhaps in partnership with commercial or nonprofit entities with communications expertise.

2. Corporate accountability: Industry self-regulation

One of the major problems with data policies hinging on individual control – and potentially a leading contributor to some of the resignation expressed by the public in both the United States and the European Union – is the burden such policies place on individuals. It is incredibly difficult for the average person to understand how data is collected, shared and used in the vast online ecosystem, and many regulatory systems ask individuals to make decisions despite this void of understanding. As one scholar puts it, "[T]he role of [the] user's self-determination in situations in which consumers are not able to understand deeply data processing and its

79 Oscar H. Gandy Jr., Engaging Rational Discrimination: Exploring Reasons for Placing Regulatory Constraints on Decision Support Systems, 12 Ethics & Info. Tech. 29, 37-39 (2010).

80 Eli Pariser, The Filter Bubble: How The New Personalized Web Is Changing What We Read and how We Think (2011).

81 Mantelero, Alessandro. "The Future of Consumer Data Protection in the E.U.: Rethinking the 'Notice and Consent' Paradigm in the New Era of Predictive Analytics." Computer Law & Security Report 30, Nov. 2014: 643, 655.

purposes,[82] or are not in the position to decide"[83] render individual control useless[84] and create resentment against the forces that produce this helplessness. Companies, fearing liability yet subject to enormous pressures to get their products to market quickly, often end up offering their customers minimal choice and more notice, rather than spending the time to implement thoughtful data practices and policies.

Benefits

The use of impact assessments in a self-regulatory scheme is one approach that could potentially provide more clarity and actionable information for individuals, while balancing companies' legitimate business interests. Under this model, companies themselves would produce assessments giving individuals a better understanding of how or when their personal information might be used in ways that are potentially beneficial or detrimental to them. A self-regulatory assessment system could also prompt companies to review their own data practices more rigorously, increasing transparency without increasing their liability. To be effective, voluntary risk assessments would have to address data processing and its subsequent uses, including variables such as "the relationship between the purposes, the context of collection, the reasonable expectations of the data subjects, the nature of the personal information and the impact [of its collection and use] on the data subjects."[85] These self-assessments could explore normative and ethical standards and thus move beyond pure "privacy by design" mandates.

One of the most important considerations in performing self-regulatory assessments would be determining when they should take place - that is, either before or after data is collected. Pre-collection assessment probably provides the most protection for consumers, as any such procedure would likely limit the scope and amount of data obtained. Post-collection assessments are also potentially useful as a way to explain the details of data processing and use to consumers, as well as serving as a form of accountability regarding actual practices, but could also end up functioning as simply as another box to check as companies rush to launch a product.

Companies might be provided with the incentive to adopt impact assessments both by the voluntary nature of the enforcement and by the potential market differentiation that could give companies using assessments a competitive advantage over those that did not. The use of assessments could shift some of the responsibility for data protection away from a pure reliance on individual initiative, while still preserving core privacy rights. One of the most convincing aspects of the argument for self-regulatory assessments is that it would reduce the need for a "notice and consent" regime. If certain impacts deriving from data processing were prohibited, for example, notices could be far more simplified and less burdensome for individuals to evaluate.

Another argument for the use of voluntary self-assessments is that they would place the assessment in the hands of experts – companies know their own data practices best – and would not require the revelation of important business secrets.

Challenges

Self-regulatory schemes are typically greatly limited by a lack of transparency and enforcement. Indeed, it was in part due to the failure of self-regulation in the first place that data-protection regulations were created in Europe (though the primary goal of the DPD was harmonization of standards across the European Union). The same is true in the United States, albeit to a lesser extent. It is possible that a legally mandated requirement to conduct impact assessments at the company level could obviate this. However, and most significantly, this approach would fail to address the cumulative and distributive social impacts of data practices, just as today's policy regime does.

A 2015 report from the OECD states that the success of industry self-regulation "depends on a number of factors, including: 1) the strength of the commitments made by participants; 2) the industry coverage of the [self-regulation]; 3) the extent to which participants adhere to the commitments; and 4) the consequences of not adhering to the commitments."[86] Creating a framework for data

82 The Boston Consulting Group. "The Value of Our Digital Identity." 2012: 4. www.libertyglobal.com/PDF/public-policy/The-Value-of-Our-Digital-Identity.pdf

83 Art. 7 (4), PGDPR ("Consent shall not provide a legal basis for processing, where there is a significant imbalance between the position of the data subject and the controller"). In 2013, the Committee on Civil Liberties, Justice and Home Affairs (LIBE) of the European Parliament dropped Art. 7 (4), see Art 7 PGDPR-LIBE.

84 Mantelero, Alessandro. "The Future of Consumer Data Protection in the E.U.: Rethinking the 'Notice and Consent' Paradigm in the New Era of Predictive Analytics." Computer Law & Security Report 30, Nov. 2014: 643, 655.

85 Article 29 Working Party, Opinion 03/2013 on purpose limitation. http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf

86 OECD. Industry Self-Regulation: Role and Use in Supporting Consumer Interests. Organization for Economic Cooperation and Development, March 2015.

usage that could draw support across a variety of business models would be difficult, and could mean watering down requirements. However, weak commitment from companies would not be likely to gain the support or trust of the government or public. Moreover, businesses that did not participate in the program might leverage the lack of government scrutiny to continue or increase unsanctioned data-usage activities.

Finally, self-regulation commitments could create market barriers for existing small businesses that could not afford to implement the requirements; these costs could end up being passed along to consumers. The power of big-business interests might also mean that the scheme could wind up being less favorable to smaller business needs. One step toward the achievement of balance in a self-regulatory framework might be increasing the participation of stakeholders such as governments and civil-society or consumer organizations. The role of such organizations in self-regulatory structures has traditionally been limited.

As an alternative to legal action by the government, businesses could produce standardized commitments that are reviewed and enforced by trade associations, neutral non-governmental organizations (NGO) or other noncommercial third parties. Third parties could also be enlisted to enforce commitments and/or administer assessments, reducing some of the burdens on individuals and increasing industry accountability without adding a layer of government bureaucracy.

3. Collective accountability: Legally mandated assessments

In the consideration of policy alternatives able to address big-data risks effectively, other areas in which risk is regulated, such as the automobile, pharmaceutical and environmental sectors, may offer some general guidance. In these sectors, the public is not expected to understand the details of how regulated products work or what side effects they may produce. We do not expect individuals to perform their own assessments of risk in these areas; instead we rely on entities, created by government mandate, that have the expertise to evaluate the efficacy and safety of products or industrial practices. In the same way, data -usage regimes could require a rigorous assessment[87] of the impact of any big-data processing, performed before such processing takes place, which would consider the impact and ethical considerations of the data use for individuals as well as for society as a whole.

Benefits

Agencies could begin this assessment process by determining the range of risk for products and services deemed acceptable both for individuals and for society as a whole. After those agencies have done their work, it would be published accompanied by publicity intended to help the information reach the target group. Individuals would still make the final decision as to whether or not to participate in the trade of their personal information.

Challenges

Education is a crucial component to this approach as well, even though legally mandated collective risk assessments would decrease the data-management burden on the individual in the short run. To avoid overly paternalistic regulation and a continued disempowerment of the public, policymakers would need to increase transparency and accountability by publishing assessments along with contextual information describing how the public interest might be adversely or positively affected as a result of the data processing or use. Government regulators might also be responsible for determining what should be considered normative standards for big-data processing beyond the principles of the FIPs. FIPs focus on process, a type of guidance that works well in a "privacy by design" approach; however, they do not set normative or ethical standards. Of all the frameworks analyzed in this report, this option is offers the most potential to create such standards, and is thus perhaps the most future-proofed.

A number of logistical questions would have to be answered regarding these assessments. For example, what method would be used to perform the assessments? How would they be altered or standardized across sectors and countries? What parties would be responsible for performing the assessments, and how frequently would they be required?

This collective approach to the use of assessments beyond a self-regulatory scheme would necessitate the creation of legal mandates for data controllers - that is, the entities deciding on the objectives and methods of the processing of personal data. Moreover, at the risk of being paternalistic, it would restrict the role played by individual control in order to increase the influence of independent authorities

87 Mantelero, Alessandro. "Data protection in a big data society: Ideas for a future regulation." Digital Investigation, November 2015.

acting on behalf of the common good.[88] In this scenario, data-protection authorities rather than individuals would be viewed as holding the technological knowledge necessary to evaluate collective risks associated with data processing, and would adopt the appropriate legal remedies and oversight mechanisms to address them. These authorities would have the perspective and position to balance stakeholder interests, including the interests of those groups who are disproportionately impacted under the current regime,[89] particularly with regard to projects requiring extensive collection and mining of public data. Rather than entirely reshaping traditional models of data protection in the United States and the European Union, this option would be responsive to the power asymmetry between data subjects and controllers created by the big-data environment. "Data protection by design" and "data protection by default" are principles essential to the 2016 EU DPR. These provisions require that data-protection safeguards be built into products and services from the earliest stage of development, and that these products and services provide privacy settings by default. Thus, EU rules are already today aimed at strengthening individuals' rights in a practical way through the mechanism of collective protection.

4. General aspects of a framework addressing big-data issues

In order to meet the challenges of big-data technologies, governments, companies and individuals each need to participate in data governance and make decisions about data as a shared resource. Incorporating data sovereignty and data portability into impact assessments will support individuals in determining if, how and with whom they want to share their data. At the same time, this approach allows the individual to do so both within the legal parameters set by governments and the technological and structural constraints defined by companies. Combining the concepts as presented here entails establishing data ownership that is shared among commercial and individual stakeholders.

What might shared data ownership look like in practice, and what would the components of such a framework include? Academics Elinor Ostrom and Edella Schlager[90] propose that governments seeking to develop an ownership framework for data protection craft policies that reflect privileges already given to people in other contexts, such as shared property rights related to natural resources. In this viewpoint, data is seen as a collective resource that should be owned and shared collectively, similar to a community park. Existing rights for the shared use natural resources include rights for access, use, and withdrawal; the same could be true for data rights. This approach could include "collective-choice" rights, such as the ability to manage group uses of personal information (through adding, deleting, or editing data), a right of exclusion of personal information from datasets, and a right of alienation, or the ability to transfer the above rights to other people. Collective choice rights would address the issue of "ours" data, or data belonging to multiple individuals, which could prove incredibly valuable for advancing the interests of big data.

In shared ownership regimes, the individual does not typically enjoy "sole and despotic dominion,"[91] such as an inviolable right to consent to or reject specific resource uses. Rather, the individual has a voice in a collective decision-making process, as well as a right to exit the collective. In the case of shared natural resources, for example, an individual has the ability to visit a local park at will, use it within reasonable limitations, and to register complaints with local or national governments about the condition of the park; she also has the choice to leave and simply visit a different park.

Some of the difficulty in designing assessments or other legal instruments able to address the challenges of a big-data world comes in figuring out how to determine the risk, benefit and harm of data processes that are complex and unseen. A June 2016 paper[92] commissioned by the United Nations Office for the Coordination of Humanitarian Affairs (UNOCHA) recommended that the following elements be included in any responsible-data-use framework:

88 Bygrave. Data Protection Law. Approaching its Rationale, Logic and Limits (n 32) 86 ("the monitoring and enforcement regimes set up by data protection laws are also a mixture of paternalistic and participatory control forms").

89 U.S. Department of Commerce National Oceanic and Atmospheric Administration National Marine Fisheries Service. Guidelines and Principles For Social Impact Assessment. May 1994. www.nmfs.noaa.gov/sfa/social_impact_guide.htm#sectIII

90 Schlager, Edella and Elinor Ostrom. "Property-Rights Regimes and Natural Resources: A Conceptual Analysis." Land Economics, Vol. 68, No. 3, University of Wisconsin Press, Aug. 1992: 249-262. DOI: 10.2307/3146375. www.jstor.org/stable/3146375

91 Diamond, Michael. The Meaning and Nature of Property: Homeownership and Shared Equity in the Context of Poverty. Georgetown University Law Center, 2009. http://scholarship.law.georgetown.edu/facpub/423/

92 Berens, Os, Ulrich Mans and Stefaan Verhulst. Mapping and Comparing Responsible Data Approaches. June 2016.

1) A description of the scope of the policy (what kind of data is covered, what specifically is being collected, and for what uses) so that individuals and others can determine if and how a policy applies to them.
2) Information on the value proposition associated with use of the individual's or group's data, or a detailed description of the purpose and anticipated benefits.
3) A description of the data used and handled by the entity collecting the data.
4) A risk assessment, automated or otherwise, describing any risk that use of the data may generate for an individual or group.
5) A risk-mitigation policy or other response, including activities and protocols intended to mitigate risks to users such as aggregation and de-identification.
6) A description of the relevant value chain or data policies that identifies how users' data will be used and processed, from the point of collection to destruction.

This framework offers a balance between the core principles of the FIPs and the new realities of data collection and use. While grounded in respect for individual control and privacy, it employs technical and policy measures that follow the data throughout its life cycle, provide reasonable oversight and accountability, and allow for big-data innovations.

VI. Conclusion

Regulatory bodies across the globe are tasked with finding a policy framework that fits today's fast-moving technological world. Much of today's regulation of data processes has remained inspired by outdated notions of individual control that are ill suited to the current big-data environment. Like so many technological advances, big data necessitates a forward-thinking policy framework that moves away from a focus on pure individual control and toward the idea of collective impact.

An ideal policy solution would combine the strengths of each framework discussed here. Empowerment, ownership, portability, corporate accountability and collective assessment work well in conjunction with one another, and would benefit from the inclusion of key FIP principles such as access, transparency, purpose and use limitations, data minimization, and data retention. The goal for such a framework would be to enhance individual power in commercial transactions through ownership and portability, while allowing the government to play a role in assisting individuals and businesses by providing technical expertise on the issue of privacy management. Corporate risk assessments would impose limits on data production and collection, while government assessments of collective impact would create standards for normative and ethical data practices. A framework of shared ownership and responsibility, facilitated by the government, might offer the right balance between individual control, regulatory intervention and business innovation. This might include requiring data operators to join a self-regulatory scheme in order to obtain a government-issued license.

The United States, the European Union and Germany represent three different jurisdictions, but each faces similar challenges in addressing big data's problems, as well as in the court of public opinion. Populations across jurisdictions and cultures have expressed a sense of hopelessness and concern regarding the loss of control over personal information; this must be taken into account as future policy is created. Much of today's regulation of data collection and processing is focused on individual rights, with governments and companies each subject to some requirements. Legislators seeking to update these systems must determine how best to manage the risks to public values created by the internet and its culture. The ideals upon which the DPD and FIPs were founded, such as purpose specification, collection limitation, transparency, data integrity, information security, access and correction should remain important components of future data-usage frameworks. However, they should also work in conjunction with limited ownership and portability rights, mandated risk assessments, and accountability for algorithmic decision-making, with governments providing enforcement and helping to improve the public's data-management skills through education programs.

Imprint

Bertelsmann Stiftung
Carl-Bertelsmann-Straße 256
33311 Gütersloh
Germany
Phone +49 5241 81-0
www.bertelsmann-stiftung.de

Responsible
Ralph Müller-Eiselt

Author
Michelle de Mooy, Director, Privacy & Data Project,
Center for Democracy and Technology

Copy Editor
Barbara Serfozo, Berlin

Graphic design
Nicole Meyerholz, Bielefeld

Photo
Artem Sapegin | unsplash.com – CC0, Public Domain
https://creativecommons.org/licenses/by/1.0/deed.de

DOI 10.11586/2017009

Address | Contact

Bertelsmann Stiftung
Carl-Bertelsmann-Straße 256
33311 Gütersloh
Germany
Phone +49 5241 81-0

Dr. Sarah Fischer
Junior Professional
Phone +49 5241 81-81148
sarah.fischer@bertelsmann-stiftung.de

Ralph Müller-Eiselt
Senior Expert Taskforce Digitisation
Phone +49 5241 81-81456
ralph.mueller-eiselt@bertelsmann-stiftung.de

www.bertelsmann-stiftung.de

BertelsmannStiftung