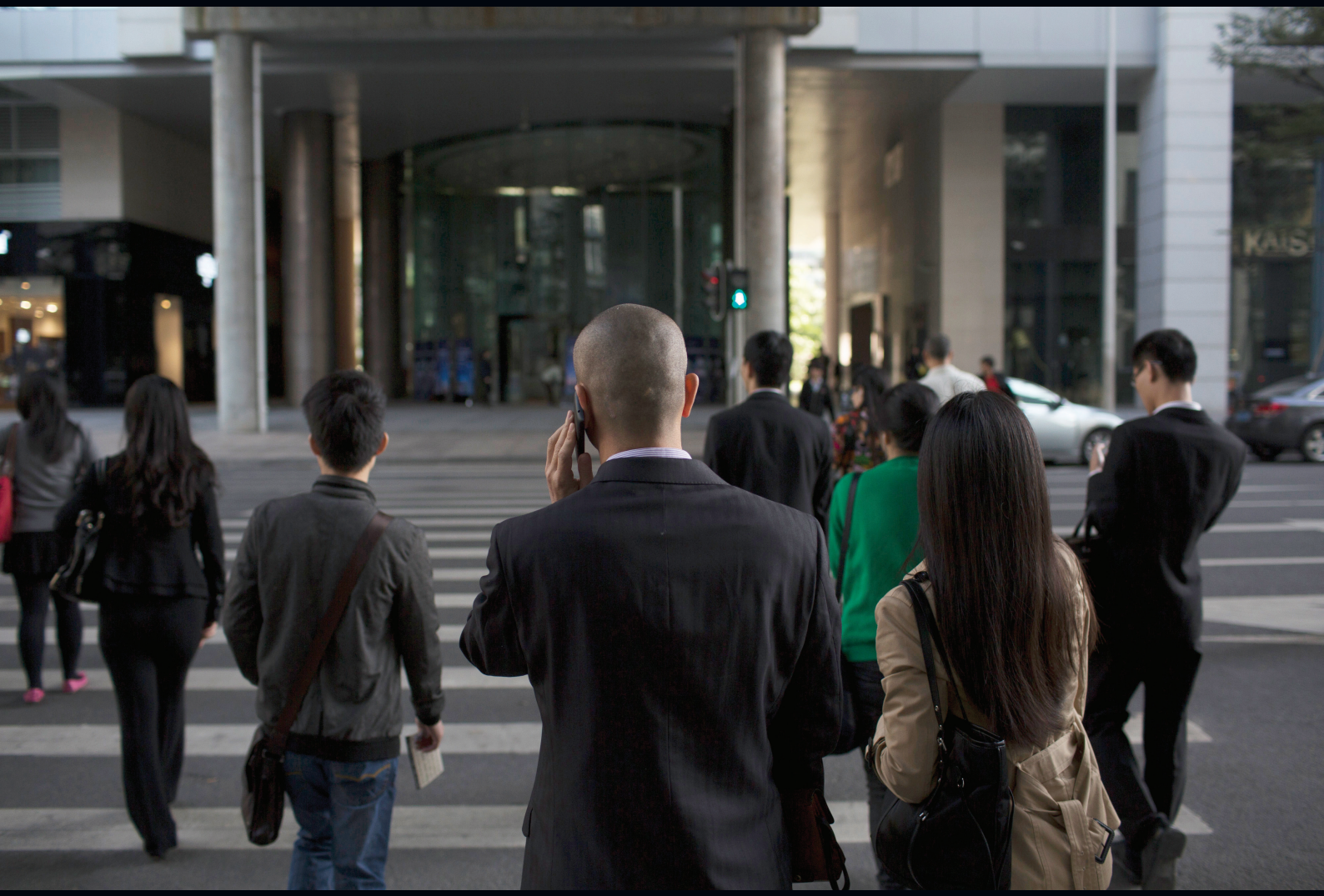
Data Brokers
IN AN OPEN SOCIETY



Upturn

OPEN SOCIETY FOUNDATIONS

Data Brokers
IN AN OPEN SOCIETY

An Upturn Report,
prepared for the Open Society Foundations

By Aaron Rieke, Harlan Yu,
David Robinson, and Joris von Hoboken

November 2016

Published by

Open Society Foundation–London
7th Floor Millbank Tower, 21–24 Millbank
London SW1P 4QP
United Kingdom
Phone: +44 207-031-0200

Cover photo: © Brent Lewin | Bloomberg | Getty
Design and layout: Judit Kovács | Createch Ltd.

About Upturn

UPTURN is a team of technologists based in Washington DC. Upturn works alongside social sector organizations to shape the impact of new technologies on people's lives.

About the Authors

Aaron RIEKE is a Principal at Upturn. Aaron works closely with major civil rights and consumer groups. Before joining Upturn, Aaron served as an attorney with Federal Trade Commission (FTC), where he focused on privacy and data security, and as a Ron Plesser Fellow with the Center for Democracy & Technology (CDT). Aaron earned his JD at Berkeley Law.

David ROBINSON is a Principal at Upturn. David leads the firm's work on automated decisions in the criminal justice system, and recently developed a coalition statement and research report on the civil rights impact of "predictive policing" products. His other areas of focus include broadband privacy, gig economy issues, and Internet Freedom. David holds a JD from Yale Law School and is a Visiting Fellow of its Information Society Project.

Harlan YU is a principal at Upturn. Recently, Harlan has been working closely with major civil rights organizations to examine law enforcement's use of body-worn cameras and other emerging police technologies. Harlan holds a Ph.D. in computer science from Princeton University and has extensive experience working at the intersection of technology and policy. He has worked at Google in both engineering and public policy roles, at the Electronic Frontier Foundation as a technologist, and at the U.S. Department of Labor.

Joris van HOBOKEN is an independent consultant, and a Postdoctoral Research Fellow in the Information Law Institute at New York University, and NYU's Department for Media Culture and Communication, a lecturer at CornellTech, a Visiting Scholar at the NYU Stern Center for Business & Human Rights, and an affiliate researcher at the Institute for Information Law (IViR) at the University of Amsterdam. His research addresses law and policy in the field of digital media, electronic communications and the internet, with a focus on the fundamental rights to privacy and freedom of expression. He has a PhD from the University of Amsterdam on search engines and freedom of expression (2012) and graduate degrees in Law and Theoretical Mathematics. Joris is the Chair of the Board of Directors of the Dutch digital rights organization Bits of Freedom.

Executive Summary

A sweeping "data broker" industry sells information about millions of people to corporate and governmental actors on both sides of the Atlantic. Data brokers, and the profiling techniques often at their core, are giving large institutions more visibility than ever before into people's lives. The industry has evolved rapidly in recent years, thanks to advances in information technology and to the industry's central role in enabling the marketing that underwrites much of today's Internet.

While some data broker products are beneficial or harmless, others threaten fundamental rights. Data brokers—and the information and inferences they supply—are playing central roles in key life decisions across a growing range of areas. Police in both the United States and Europe purchase corporate assistance to profile residents based on personal data. Credit bureaus, subject to special rules, play a critical data brokerage role in mediating access to finance. Political parties on both sides of the Atlantic are now targeting their digital outreach based on details of individual behavior. In the US, prospective employers routinely turn to data brokers to purchase criminal history reports regarding job candidates (reports that are notoriously error-prone).

Today, civil society struggles to identify the most concrete harms and risks presented by data brokers and products that they sell. In the US, policymakers have focused their attention on data brokers that sell products for use in marketing. These inquiries have reinforced general concerns about privacy and transparency, but have revealed little in terms of specific harms to rights and justice. In the European Union, robust-sounding legal principles are well established, but public authorities and civil society often struggle to apply them in concrete ways. On both sides of the Atlantic, there is not yet a clear regulatory agenda for data brokers.

The issues described in this report are central to a larger trend of automated, data-driven decision making by large institutions, when it comes to the key decisions that shape people's lives and impact their rights. The biggest open challenge in the field—as we detail below—is how to approach these technological changes in ways that best advance longstanding commitments to human rights and justice.

This report is organized in four sections:

- **What We Know** provides a high-level overview of the data brokerage industry. We define "data broker," review the different types of brokerage and profiling products sold by data brokers, identify leading companies and market segments, and summarize existing policy research.

- **The Legal Landscape** presents relevant laws and policy developments in both the United States and the European Union. In the United States, a sector-by-sector approach leaves vast swaths of consumer data largely unregulated. However, uses of data in key areas, such as credit, employment, insurance, and housing are subject to some restrictions. In the European Union, by contrast, privacy and data protection are treated as fundamental rights and the EU has a broad regulatory framework that attaches data protection safeguards to the processing of personal data by *any* entity in the private sector, including data brokers (as well as a separate regime for law enforcement uses of personal data). However, those rules are developed and debated in largely abstract terms, and all types of stakeholders in Europe—from regulators to civil society to companies—struggle to understand and apply the rules in concrete situations.

- **Data Brokers in Context** describes the impact of data brokers and profiling in three important domains of daily life: marketing, consumer credit, and policing. These examples illustrate both the scope of the data brokerage and profiling industry, and the range of ways in which public policy can, and sometimes does, specifically restrict problematic practices.

- **Strategic Directions and Open Questions** offers questions that the social sector should consider moving forward. We recommend an impact-driven, bottom-up approach to further investigation of data-driven profiling by data brokers and corporate profiling services.

Table of Contents

1. What We Know About Data Brokers and Profiling

This section offers an overview of the best available information about what data brokers do, how they do it, and how their products are used. We begin by defining key terms, and then describe the wide variety of industry practices, data sources, and product verticals that collectively comprise the data brokerage industry.

1.1 The definition of "data broker" is contested.

There is no authoritative definition of "data broker" on either side of the Atlantic. Neither United States nor European policy provides clear guidance.[1] For example, the US Census Bureau does not assign a business classification to data brokers, and relies instead on a range of overlapping categories (including "data processing and preparation," "credit reporting services," and "information retrieval services").[2] The International Organisation for Economic Co-operation and Development (OECD) has adopted a similar approach, remarking specifically that there is "no standardised classification of data brokers."[3]

Policy research in the US offers several sweeping definitions. The Federal Trade Commission (FTC) defines data brokers as "companies that collect consumers' personal information and resell or share that information with others."[4] Similar definitions have been adopted by the US Senate,[5] the US Government Accountability Office (GAO),[6] and the Office of the Privacy Commissioner of Canada.[7] Some industry studies resist the term data broker altogether, using terms like "data-driven marketing economy" to draw attention the complex interrelation of different kinds of data companies.[8]

In Europe, the term data broker is less common than in the US. European commentators use a variety of different terms to refer to data brokers, including "information resellers," "data vendors," "information brokers," "consumer data analytics," "data warehousing," "Datenhändler" (German), and "traders de données" (French). When the term "data broker" is used, it often reflects the definitions established by US discussants.[9] For example, in a recent study, the European Data Protection Supervisor (EDPS) defines data brokers as entities that "collect personal information about consumers and sell that information to other organisations."[10] The Norwegian Data Protection Authority uses the same definition as the US FTC.[11] And the OECD, in its study of the economics of personal data, defines data brokers as "firms that gather and merge aggregated information on individuals that is then sold for various uses . . . ."[12]

For the purposes of this report, we use the following definition of "data broker," which aligns closely with common definitions in the US and Europe. **A data broker is**:

A company or business unit	*Data brokers are often subsidiaries of larger companies.*
that earns its primary revenue	*Many companies earn* some *revenue by supplying data—but data brokers earn their* primary *source of revenue by supplying data.*
by supplying data or inferences about people	*Many data brokers sell not only data, but also inferences or predictions about people.*
gathered mainly from sources *other than* the data subjects themselves.	*Companies that collect most of their data directly from consumers are not usually considered to be data brokers.*

It is important to note that this definition, while broad, excludes a number of large, influential companies. For example, Google and Facebook interact directly and daily with most Internet users in the US and Europe. This gives these companies deep knowledge of people's behaviors and interests, and they can allow others to target ads on their platform to users most likely to respond. However, they are not "data brokers" under our definition—or in common parlance—precisely because their data comes primarily from consumers firsthand, rather than being sourced from other businesses. Similarly, Disney is not a data broker, despite the fact it shares its customers' data with a range of other companies, because such data is not the company's primary source of revenue.[13]

1.2 "Profiling" means making inferences about people.

The word "profiling" carries different meanings in different contexts. Broadly speaking, the term refers to any inference or decision that is based on someone's personal traits. However, the term is also commonly used to refer specifically to problematic or invidious inferences. For example, in the US civil rights context, "*racial* profiling" describes the highly controversial law enforcement practice of using race as a basis for investigative or other decisions.[14] And in Europe, profiling has long been considered a concerning practice, even when broadly defined.

US law does not provide a single definition of "profiling," and instead offers a patchwork of regulations that restrict decisionmaking in particular domains. The EU defines and regulates profiling explicitly in its General Data Protection Regulation (GDPR) as a form of automated decisionmaking. The GPDR defines profiling as "any form of automated processing of personal data consisting of using those data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements."[15]

In this report, **we use the term "profiling" in its broad sense, to refer to the creation or use of inferences about people**. Although we focus on inferences that either are made by data brokers or are informed by brokered data, it is important to note that profiling is not the sole province of data brokers. For example, Google and Facebook regularly engage in profiling to target advertisements, often without the assistance of data brokers. And some companies may specialize in profiling and analytics, without actually selling any data themselves. Finally, as we discuss in more detail below, law enforcement agencies may assemble data from disparate sources themselves (effectively internalizing a data broker function) while relying on corporate help to profile individuals using such data.

1.3 Data brokerage is an old practice, but new technologies are spurring growth in the industry.

Data brokerage predates the Internet. Companies have specialized in collecting and analyzing data for decades. For example, the practice of segmenting consumers for marketing purposes—often thought of as a very "modern" data broker practice—dates back to at least the 1970s, when a company called Claritas pitched a "lifestyle segmentation system" that promised to help marketers gain insight into their customers' preferences.[16] Primitive forms of credit scoring emerged in Europe a century ago. In Germany, the Berlin municipal electricity company started

assessing payment installment plans on the basis of electricity bill payments in the 1920s.[17] Credit bureaus began in the US in the 1950s as small, local data brokers that arose to help provide lenders with better information about prospective borrowers. And Fair Isaac Corporation (FICO)'s credit scoring technology, which is today used by more than half of the top 100 banks in the world, originated in 1981.[18]

What *has* changed in recent years is "the tremendous increase in the volume and quality of digitally recorded data—and the technological advances that have facilitated access to, storage, analysis, and sharing of this information."[19] New sources of data, new ways to access and store old sources of data (including government and commercial records that are now digitized), and powerful new analytical techniques are driving the data brokerage industry forward.

Today, as consumers increasingly use credit cards and turn to online purchases, commercial transactions result in structured digital records that can be aggregated to yield insights about individual people. This new and increasingly cashless economy, and the data it produces, is only growing. According to figures compiled by the Bank of International Settlements, all the US dollars in circulation in 2010 had a value of just 7% of the country's GDP.[20] Sweden is even further along this adoption curve, with cash in circulation equal to just 3% of GDP.[21] A recent study from Stockholm's Royal Institute of Technology argued that the country is on the brink of being effectively cashless.[22] Across the Eurozone, notes in circulation in 2010 totalled just under 10% of GDP.[23] And a growing alternative payments landscape creates "invisible payments" not reflected in the usual statistics, including "mobile money" apps like Venmo and "digital wallets" such as PayPal, which are now responsible for billions of transactions each year.[24]

The widespread and growing use of mobile devices provides another rich new source of revealing data. As of 2014, more than 60% of Europeans carried modern smart phones,[25] and 58% of Americans did (up from just 44% as recently as 2012);[26] another study claimed US smartphone penetration of 74.9% in 2015.[27] These modern mobile devices provide rich new data about people including their location, the apps they use, and their contacts. This information can not only be collected by mobile platform providers like Google and Apple, but also by app developers and the data brokers that provide developers with analytics and advertising.[28]

People's web browsing is another new source of data, on both mobile and desktop devices. A visit to a single website will often trigger interactions with dozens of other organizations involved in advertising or analytics, many of which either are data brokers or exchange data with brokers. The Wall Street Journal's extensive "What They Know" series, published between 2010 and 2012, found that a test computer, after visiting the top 50 web sites, was left with 2,224 cookies "installed by 131 companies, many of which are in the business of tracking Web users to create rich databases of consumer profiles that can be sold."[29]

In short, while data brokerage and data profiling are not new practices, the market is evolving. However, it is far from clear how this gold rush for data concerning people's digital activities will impact core social justice areas, including finance, criminal justice, and education.

1.4 The data brokerage industry is sweeping in scope.

The data brokerage industry is vast, varied, and complex. Data brokers count among their customers advertisers, merchants, employers, bankers, insurers, police departments, schools, hospitals, and others. They seek to meet the varied needs of their customers by collecting data from many different sources, and selling different *types* of products, ranging from simple lists to scores produced by proprietary actuarial models.

It is important to note that the division between data brokers and their many customers varies. Data brokers *can* play a central role in profiling, but they are far from having a monopoly on data analysis. Small entities may rely almost totally on brokers to analyze the data and produce insights, while larger entities are likely to have at least some profiling capacities in-house. And, in the context of policing and intelligence work, governments (particularly in Europe) may internalize the process of assembling and storing personal data, while still using corporate services to profile people based on such data.

1.4.1 There are many markets for brokered data.

Data brokers sell products tailored for many different purposes to many different types of customers. Markets for brokered data include:

- **Advertising and marketing.** Data brokers help companies target advertisements, create marketing strategies, set corporate goals, and determine where to open new branches. A wide range of companies use data brokers' products. For example, in 2013, Acxiom's customer list included "47 *Fortune* 100 clients; 12 of the top 15 credit card issuers; seven of the top 10 retail banks; eight of the top 10 telecom/media companies; seven of the top 10 retailers; 11 of the top 14 automotive manufacturers; six of the top 10 brokerage firms; three of the top 10 pharmaceutical manufacturers; five of the top 10 life/health insurance providers; nine of the top 10 property and casualty insurers; eight of the top 10 lodging companies; two of the top three gaming companies; three of the top five domestic airlines; six of the top 10 US hotels."[30]

- **Credit and insurance.** Data brokers help lenders and insurers set prices for financial products, manage their risk, and comply with regulations. Virtually every major financial institution relies heavily on data brokers to supply data with which to underwrite their products. Working together with analytics firms, data brokers enable a high degree of automation in lending. For example, American Banker reported that, as of 1999, at large banks "no [human being] even looks at any [credit request] for $50,000 or less—the computer does it all."[31]

- **Identify verification and fraud detection.** Data brokers help entities verify people's identities and credentials, and detect fraudulent purchase patterns. These products and services are useful and widely used by lenders, retailers, telecommunications firms, and many other entities. For example, LexisNexis Risk Solutions—a market leader in this area—recently counted 38 of the top 50 US banks, as well as 90 percent of the Fortune 500 companies, as clients.[32] The US "risk information" industry was estimated to be about $7 billion in size in 2010.[33]

- **Health.** Data brokers help healthcare providers measure and improve their performance. For example, one major data broker boasts that it holds "over 85% of the world's prescriptions by sales," data that can help healthcare providers "run their organizations more efficiently and make better decisions to improve their operational and financial performance."[34]

- **Education.** Data brokers help schools and educational technology companies share access to student data, allowing teachers and software developers to evaluate and respond to student performance. For example, one United States data broker is used by more than 44,000 elementary and secondary schools—about one-third of the kindergarten through 12th-grade schools in the country.[35]

- **Government and Law Enforcement.** Although the public sector market for data broker products is small in financial terms, such uses may have an outsized impact on people's lives. For example, person-search tools make it easier for law enforcement in the US to locate potential suspects. In fact, one major person-search product is touted to be used by "over 4,000 federal, state and local law enforcement agencies across the country."[36] Private investigators are customers in the same market. But in a criminal justice system where certain communities are already overrepresented as suspects, easy search tools that can be used without careful justification may reinforce existing disparities. And minor inaccuracies can lead to dire consequences—from a wrongful arrest to the incorrect application of force. In addition, even when they do not buy brokered data, law enforcement authorities on both sides of the Atlantic rely on corporate help to profile individuals based on the extensive data that they hold.

- **Consumer services.** Data brokers can help consumers locate old friends or research their genealogy. For example, data broker Ancestry.com claims to be the "world's largest online resource for family history."[37]

1.4.2 The industry absorbs data from a wide range of sources.

Data brokers amass different types of data from different sources. For example, after studying nine data brokers that represented a cross-section of large, mid-sized, and small data brokers, the FTC summarized that:

> Data brokers collect and store a vast amount of data on almost every US household and commercial transaction. Of the nine data brokers, one data broker's database has information on 1.4 billion consumer transactions and over 700 billion aggregated data elements; another data broker's database covers one trillion dollars in consumer transactions; and yet another data broker adds three billion new records each month to its databases. Most importantly, data brokers hold a vast array of information on individual consumers. For example, one of the nine data brokers has 3000 data segments for nearly every US consumer.[38]

It is difficult, if not impossible, to pull apart these impressive-sounding figures for closer analysis. Data brokers are resistant to sharing details about their data sources, "citing confidentiality clauses in their contracts, and concerns about putting themselves at a competitive disadvantage," even when facing formal inquiries or confronted with individuals exercising their rights to transparency in Europe.[39] In any event, because data brokers so frequently buy and sell data from one another—including inferred or predicted data—it would be in many cases infeasible to fully account for how some data brokers obtain their data, even if they were willing to share all relevant details.[40]

Different regulatory regimes shape different data brokers' collection practices. For example, in the United States, data brokers are allowed to collect data permissively. In the EU, by contrast, the flow of data between sources, data brokers and their customers is theoretically limited, because every new transfer of data needs its own legal justification, sometimes requiring consent. These limits apply even to the collection of public records and publicly available data. However, the Norwegian Data Protection Authority concludes that increased international competition will embolden data brokers established in Europe to "look into the possibility of collecting and combining information from a wider array of sources than today."[41]

The quality and accuracy of data brokers' data varies depending not only on regulations, but also on the intended uses. For example, a company seeking to tailor its advertising efforts is

likely to tolerate some guesswork. In these and other marketing contexts, "contracts between data brokers and their clients include few provisions regarding the accuracy of their products."[42] However, in other contexts, such as lending, legal requirements and customers will demand more accurate data.

Not all data collected and used by data brokers identifies an individual person. For example, a marketer will happily target an online advertisement to a *device* known to be associated with certain browsing habits. A bank seeking to open a new branch may be satisfied to know about the demographics and wealth level of a particular neighborhood. In short, even data that is not "personal data" can still inform decisions that affect people's lives.

At a high level, data brokers can acquire data from the following categories, including:

- **Publicly available data.** Many data brokers collect and organize data that is available to the general public. Data brokers will commonly collect such data using "web crawlers" (software programs designed to automatically collect data from the Internet) or purchase it from other data brokers that specialize in digitizing particular types of records.[43] Publicly available data includes:
    - *Government records*, including property reports, court filings, criminal convictions, and professional licensures. Public entities can share a surprising amount of data. For example, in the US, state Departments of Motor Vehicles can sell data to private companies for identity verification purposes.[44] And some states allow voting records to be bought and sold.[45] In Europe, the availability of public records for further processing by data brokers varies. For instance, some Scandinavian countries are notoriously transparent with respect to people's tax records under freedom of information laws, while such data is not available to the public elsewhere.[46]
    - *Business listings*, including telephone books and classified advertisements.
    - *Media, social network and online data*, including public information from LinkedIn, Facebook, Twitter, and YouTube and discussion sites.

- **Nonpublic data obtained through private contract.** Data brokers obtain significant amounts of data from private entities, including other data brokers. The contracts governing these exchanges typically include a range of provisions, including some license to use the data (*e.g.*, use of data for a defined period of time, the right to resell, etc.) and warranties that the data was obtained legally.[47] Some data brokers facilitate "data cooperatives," where companies provide information about their customers "in exchange for information to enhance their existing customer lists or identify new customers."[48] Data brokers commonly collect data from:

- *Retailers* will frequently sell information about their customers' purchases, the frequency of those purchases, and how those purchases were made. For example, Datalogix boasts that it has "information on more than $1 trillion on consumer spending 'across 1400+ leading brands."[49]
- *Financial institutions*—like banks, credit unions, brokerage services, and insurers—often share detailed information with data brokers, sometimes as a condition to accessing credit reports and credit scores.
- *Employers.* In the US, one data broker sells detailed salary and pay stub information for almost 40% of employed Americans.[50]
- *Registration information from websites.* Acxiom claims that there are "over 250,000 websites who state in their privacy policy that they share data with other companies," including data brokers.[51] Data brokers commonly obtain lists of people who register for retail, news, and travel websites.
- *Data from other data brokers.* Data brokers frequently sell data to one another. The FTC's recent report on the field found that "[m]ost of the [studied data brokers'] commercially sourced data . . . comes from other data brokers outside this study."[52]

► **Online tracking data.** Data brokers obtain data by tracking a person's web browsing behavior, or their use of a mobile device. For example, popular websites will frequently result in the exchange of data with tens, or even hundreds, of behind-the-scenes trackers that record the websites a particular internet user has visited.[53]

1.4.3 Data brokerage products take many different forms.

Data brokers offer different kinds of products. For example, one data broker might specialize in selling basic lists of consumers, while another might specialize in analyzing a person's detailed financial history and producing a single credit score.

Broadly speaking, data brokers' products consist of both "actual" and "modeled" data. Actual data is factual information about people, such as their name, contact information, demographics and other behavioral data. Modeled data is the result of profiling (*i.e.*, inferences or guesses about people based on actual data). For example, a data broker might infer a person is a woman based on her shopping habits. Or, a data broker might infer that a person is likely to default on a loan based on her past financial behavior.

Common data broker products include:

- **Original lists.** Lists of people and contact information based on certain factual criteria, commonly used in marketing. For example, a data broker could sell a list of "men living in New York who have memberships at golf courses."

- **Segments.** Lists of consumers grouped by predicted characteristics or behaviors, commonly used in marketing. For example, one US data broker sold a segment called "Thrifty Elders," which includes "singles in their late 60s or 70s in 'one of the lowest income clusters'."[54] In the Netherlands, Experian offered data in a category of "Less Successful", meaning "people who are home a lot" or who live in "decayed houses."[55] Data brokers will often segment consumer by major life events, like getting married, buying a home, or sending a kid to college.[56]

- **Consumer reports.** Dossiers about particular, identifiable people, commonly used for credit, insurance, employment, or similar types of personalized assessment. For example, credit bureaus sell records of people's financial standing and past repayment behavior. Consumer reports often contain actual data, but will be used to make inferences (*e.g.*, in the form of credit scores).

- **Look-alike models.** Models that use known data about a person to predict the behavior or characteristics of that person based on the behavior of similarly-situated people about which the data broker already has data. These models are commonly used in marketing.

- **Scores.** Predictions about consumer behavior based on data about that person, used for a wide range of different purposes. For example, credit scores predict the likelihood that a person will default on a loan. Marketing scores can predict how likely a person is to buy a particular product. Lead scores might predict how likely a prospective consumer is to be a profitable customer. Fraud scores might predict how likely a particular transaction is to be fraudulent. And stress scores might help organizational customers manage healthcare costs and risks.[57]

- **Data "appends."** Services to customers that have some data about a person, but want to build a more complete record. Typically, a customer will provide a data broker with some identifying information about a person (such as an e-mail address) and the data broker provides additional information about that person (such as data about their recent purchase patterns).

1.5 The US market has clear leaders. The EU market structure is more fragmented.

Due to the vast scope of the data broker marketplace, it can be difficult to characterize its precise size. Both the United States and Europe lack comprehensive lists or registries of companies that resell personal information.[58] Several privacy groups maintain lists of data brokers, but none are exhaustive or up to date.[59] Trade group memberships offer a distorted picture, because trade groups often represent wide swaths of companies, many of whom are not data brokers.[60] For example, one news story reported that "[n]o one even knows how many companies there are trafficking in our data. But it's certainly in the thousands, and would include research firms, all sorts of Internet companies, advertisers, retailers and trade associations."[61]

However, despite the difficulty of precisely defining its boundaries, the US data broker market takes in significant revenue. US Senator John D. Rockefeller claimed that "[i]n 2012, the data broker industry generated $156 billion in revenues," a sum that is "more than twice the size of the entire intelligence budget of the United States Government."[62] Although there are numerous data brokers active in the Europe, the European data broker landscape is not comparable to the US market in terms of market size. The European revenues of large data brokers, such as Acxiom, LexisNexis, amount only to a fraction of their overall revenues.

The data broker marketplace includes companies of many different types, sizes, and service specialties. Many data brokers perform multiple functions. For example, Experian, a leading credit data broker in the US, also has a sizable marketing division. The company sells both highly-regulated credit reports and lightly-regulated lists that include "names of expectant parents and families with newborns."[63] Equifax, another large credit bureau, maintains tens of thousands of individual data elements for its marketing products, including "information as specific as whether a consumer purchased a particular soft drink or shampoo in the last six months . . . ."[64]

Data brokers are also diversified at an intra-organizational level. While some data brokers report large revenues, not all of this revenue was generated from data brokerage activity (as opposed to, say, other technological capacities). For example, while Acxiom made $1.098 billion in total revenue in 2014, only $676.9 million came from data brokerage activity.[65] Some data companies do not publicly disaggregate their revenue. For example, though Epsilon reported $1.5 billion in revenue in 2014, it is unclear what amount of revenue was generated by their ad agency services versus their marketing-based data brokerage activities.[66]

In the US, the market for "risk information"—a category that covers insurance, risk management, fraud detection, verification services, and credential authentication activities—appears to dwarf the market for marketing-oriented data broker products. In 2010, LexisNexis estimated that this "risk information" industry was about $7 billion in size.[67] For perspective, combining the 2014 revenues from the marketing activities of Acxiom, the three major credit bureaus, and Datalogix approximately equals $1.56 billion. That combined total is still less than the 2014 revenue from just the *singular* risk mitigation data broker market leader, LexisNexis Risk Solutions, with $1.58 billion.

There are also clear leaders in the marketing space. For example, in 2014, Experian recorded $433 million in revenue from marketing services (17 percent of Experian's North American total revenues),[68] while Equifax generated $197.8 million in revenue,[69] and TransUnion generated $134.5 million in revenue.[70] Acxiom generated $676.9 million in revenue from its marketing activities in the US.[71] Epsilon, another marketing-based data broker recorded $1.5 billion in revenue for 2014, though it is not clear how much of that revenue comes from its traditional data brokerage activities and from its growing ad agency services.[72] And Datalogix—recently acquired by Oracle Corp. for a reported $1.2 billion—reported approximately $125 million in revenues in 2014.[73]

The European landscape, by contrast, is highly fragmented across national European markets, complicating measurement. This fragmentation is the result of different national legal regimes and varying availability of data. In the context of marketing, there appears to be the least fragmentation, in particular in the market of online marketing. Data brokerage for direct marketing is still mostly a national affair, with strong national players connected to the media industry.[74]

Many major US data brokers are expanding into Europe. For example, Acxiom has activity across Europe and offices in several European countries. Acxiom did close an office in Spain due to the regulatory environment in 2007.[75] And Experian has a strong presence in Europe. Datalogix, recently acquired by Oracle, has a strong presence in the UK.

A look at LexisNexis illustrates how data brokers are reacting to European regulatory fragmentation. LexisNexis Risk Solutions offers its people search data products across Europe, but this offering is not similarly advertised everywhere. In the Netherlands, it offers LexisNexis Diligence for the "screening of persons and organizations," while in Germany this product is only discussed for the screening of "business partners" and in France, more vaguely, for "stakeholders." In the UK, this product is advertised most broadly, including for the screening of employees.[76] At least for its people search product, these differences may reflect differences in regulatory restrictions but may also hint at differences across Europe in terms of demand as well as incentives to stay below the radar.

2. The Legal Landscape

The US and EU take different approaches to privacy and data protection, the lenses through which data brokerage is most often framed. This section describes relevant US and EU policy frameworks, including some laws that may not, at first glance, seem relevant to data brokers themselves.

At an abstract level, the US and EU share some common conceptions of privacy. Both view privacy safeguards as important, and "expound a core set of broadly similar principles for the protection of personal data" that informs the development of more specific policies.[77] These similar principles—often called the Fair Information Practice Principles (FIPPs)—comprise the most widely-accepted privacy framework in the world.[78] They speak to both the collection and use of data, and recommend, among other things, that data collection be minimized where feasible and that data be used for limited and particular purposes. These principles leave enormous room for interpretation and varied application.[79]

However, in practice, the US and EU have very different approaches to privacy.[80] In the US, the starting assumption is that processing of data is permitted, whereas in the EU, all personal data processing needs a legal justification and is subject to a set of interlinked obligations for transparency, fairness and lawfulness. In contrast to the US, in the EU, data privacy enjoys a status as a fundamental right. The result is a broad and comprehensive legal posture. In the US, privacy law lacks a clear source of moral authority, except where the Constitution addresses privacy with respect to government actors.[81] The result is a body of US privacy law that has been characterized as "haphazard and riddled with gaps."[82]

2.1 The US regime: A patchwork of specific, entrenched rules.

In the US, there is no overarching federal law that governs the collection and sale of data by commercial entities, including data brokers.[83] Instead, a patchwork of sector-specific laws govern the collection and use of personal information in certain situations, in certain sectors, or by certain types of entities.[84] For example, different federal laws set different rules for credit reports,[85] education records,[86] bank records, [87] video rentals,[88] health information,[89] and information gathered from children.[90] State lawmakers have passed a range of their own laws requiring, for example, that websites post privacy policies and requiring employers to give notice before monitoring email.[91] As a result, most consumer data is not covered by any privacy law.[92]

The Fair Credit Reporting Act (FCRA) is a particularly important part of this legal patchwork. It regulates data brokers that collect data about consumers for the purpose of selling that data for use in certain, enumerated eligibility decisions, including credit and employment. The FCRA requires that covered data brokers act in ways that are "fair and equitable to the consumer, with regard to the confidentiality, accuracy, relevancy, and proper utilization of information."[93] However, the FCRA does not apply to data brokers that collect and sell data for purposes not specifically covered by the law (including, for example, marketing).

A separate category of law regulates various decisionmakers, such as creditors and landlords. Although these laws do not regulate data brokers *per se*, they affect the types of data that data brokers' customers are willing to purchase and use in their decisions. For example, the Equal Credit Opportunity Act (ECOA) is designed to stop creditors from unfairly denying credit opportunities to qualified borrowers on account of a "prohibited basis" such as a borrower's race or age.[94] It controls how credit scoring systems may be built and how they must be validated.[95] As a result, data brokers are cautious about the kinds of credit scores they sell. Similarly, the Fair Housing Act (FHA) protects people from discrimination when they are renting, buying, or securing financing for any housing.[96] Both laws incorporate the "disparate impact" doctrine, which allows for civil rights claims against practices that, even if well intentioned, have a disproportionate adverse impact on protected groups.

The FTC, chartered as a nationwide consumer protection agency, is the closest thing the US has to a general purpose data protection authority, but its powers are limited. The FTC is empowered to sue companies for deceptive and unfair practices that impact consumers or competition, and will often sue companies that misrepresent their privacy practices, or that have been negligent in providing data security.[97] It also has the authority to compel companies to answer interrogatories and file reports, which allow it to perform a general fact-finding role.[98] However, the FTC has very limited powers to police "unfair" practices beyond those resulting

in direct, economic injury to consumers.[99] The agency acknowledges that its current powers are insufficient to deter potentially harmful data practices. It has repeatedly urged Congress to grant it additional authority.[100]

On a legislative level, advocates have fought hard for an omnibus federal law governing commercial uses of personal data. This goal faces long odds. Although Congress has held numerous privacy-related hearings over the last ten years, these hearings tend to be mere political or exploratory exercises, rather than adjuncts to serious legislative efforts.[101] To date, Congress has shown nothing resembling the focus or political will that would be necessary to undertake a vast new imposition of government authority over private companies with respect to their data handling practices. Most recently, the Obama administration's proposal for a "Consumer Privacy Bill of Rights" was panned by industry participants and privacy advocates alike, offering little apparent common ground.[102] Looking ahead, the ascendance of technology industry lobbying and the public's increasing comfort with digital technologies suggests that such legislation remains unlikely in the foreseeable future.[103]

In sum, as characterized by privacy scholar Paul Schwartz, the US approach is one of "regulatory parsimony": "before the US legal system acts, the lawmaker will wait for strong evidence that demonstrates the need for a regulatory measure."[104] And it is a safe bet that little will change in the near future, absent new evidence of harm. It will fall on the public and social sectors to demonstrate the need for new regulatory and legal interventions.

2.1.1 US regulators have recently focused on marketing, with limited success.

In recent years, US policymakers have focused significant attention on the data brokers that specialize in marketing. Each inquiry has noted a lack of public information about the industry. "There is little publicly known information about the [data broker] industry as a whole," wrote the GAO in 2013.[105] The US Senate Committee on Commerce observed "gaps in public knowledge" regarding data broker practices that same year.[106] And as of 2014, the FTC noted that the practices of many data brokers still "remain opaque."[107]

There are several reasons for this focus on marketing. First, because the FCRA already regulates data brokers that sell data for credit, employment, and other important purposes, regulators have chosen to highlight data brokers that are less regulated. Second, marketing data brokers are making some of the most intense and visible use of new sources of data (*e.g.*, personalized advertisements as a result of online tracking). And the FTC, the closest thing the US has to a general purpose data protection authority, is a consumer protection agency, has a long history of

engagement in marketing and advertising issues. However, as described in subsequent sections, these explorations have yet to thoroughly substantiate or articulate harms to consumers apart from abstract erosion of privacy.

2.2 The EU regime has powerful principles, but uncertain impacts.

The EU has a comprehensive data privacy framework, the Data Protection Directive (DPD), that regulates the processing of personal data by any entity in the private sector and much of the public sector.[108] The DPD harmonized data privacy laws across the EU, but implementations varied considerably. Further harmonization will occur under the new General Data Protection Regulation (GDPR), which was officially adopted in early 2016, and will come into force and replace the DPD in early 2018.[109] Although the GDPR is similar to the DPD in many respects, implementation of the GDPR is likely to have a significant impact on data brokers.

The DPD and GDPR data privacy framework is the primary lens through which most EU policymakers address issues related to data brokers. This framework exists under the umbrella of the protection of information privacy as a fundamental right at the EU and the Council of Europe level.[110]

Below, when discussing the fundamental contents of European data privacy law, we are describing elements that are for the most part common between the existing DPD and the new GDPR. Where appropriate, we note differences and the likely impact of the changes in the GDPR for data brokers and profiling.

The DPD and GPDR broadly provide a set of interlinked data privacy safeguards on "data controllers" involved in the "processing" of "personal data." A "data controller" is an entity responsible for a particular personal data processing operation.[111] "Processing" includes any operation one could perform on data, including collection, storage, organization, and disclosure or transfer to third parties. "Personal data" is defined broadly as "any information relating to an identified or identifiable natural person ('data subject')".[112] Currently, the DPD applies to companies, including companies headquartered outside of the EU, that have an establishment or use equipment in a European country in their handling of personal data.[113] The GDPR more explicitly "applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not".[114]

The European data protection framework implies that if personal data is processed, then there must be a data controller that can be held responsible for complying with a set of data privacy safeguards. These safeguards include the following main principles (Article 5 GDPR):

- **lawfulness, fairness and transparency:** personal data should be processed lawfully, fairly and in a transparent manner in relation to the data subject. Lawfulness requires a legitimate ground for such processing. Legitimate grounds include the consent of the data subject, if the processing is necessary for a contract or provision of a service, or the fulfillment of a balancing test (GDPR, Article 6(a), (b) or (f) respectively).

- **purpose limitation:** personal data should only be collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes.

- **data minimization:** personal data should be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed.

- **accuracy of personal data:** personal data should be accurate and, where necessary, kept up to date.

- **storage limitation:** personal data should be stored or kept in a directly identifiable form for no longer than necessary.

- **integrity and confidentiality:** personal data should be protected against unauthorised or unlawful processing, accidental loss, destruction or damage.

- **accountability:** the data controller should have the ability to demonstrate compliance with the previous principles.

In addition, the DPD and GPDR provide for the following:

- the processing of special, sensitive personal data—and separately, of data relating to criminal convictions and offenses—is not allowed outside specific regulated circumstances or without explicit consent.[115]

- data controllers must be transparent about the processing of personal data, including the purposes for which data are being processed and the possibility of exercising data subject rights (transparency) and individuals have rights to access their data, to ask for correction of incorrect data and under some circumstances a right to object and deletion (data subject rights).[116]

- automated decisions based on the processing of personal data, including through profiling, trigger certain protections.[117]

EU data protection law is enforced by independent, national Data Protection Authorities, who coordinate their interpretation of the rules through the Article 29 Working Party. The GDRP replaces the Article 29 Working Party with a European Data Protection Board (EDPB) with enhanced powers. In recent years, the Article 29 Working party has adopted numerous opinions and working papers, addressing core definitions and principles such as the concept of personal data[118] and the legitimate interests provision,[119] as well as topics such as behavioral advertising[120] or the Internet of Things.[121]

In Europe, data privacy is recognized as a fundamental right, providing a favorable background to implementation of specific safeguards in the GDPR. The European Convention on Human Rights, adopted in 1950 by the Council of Europe (CoE), includes the right to private life in Article 8. This fundamental right generally applies to the collection and use of personal data, including by the private sector.[122] A specific data privacy convention (Convention 108) was adopted in 1981 and several non-binding recommendations on relevant issues have since been adopted in the CoE context.[123]

Since 2009, the EU has had its own binding "Bill of Rights," the Charter of Fundamental Rights.[124] Within the EU Charter, Article 8 provides a specific right to the protection of personal data. Article 8 implements the core elements of existing European data protection law: the obligations on data controllers, the rights of data subjects, as well as independent oversight.[125]

Three recent judgments by the Court of Justice of the EU (CJEU) show that the Charter has significantly enhanced data protection in Europe. In the last two years, relying on the Charter, the CJEU established a right to request delisting of search results,[126] struck down mandatory data retention,[127] and annulled the EU–US Safe Harbor agreement.[128] The first of these cases was brought by the Spanish DPA, while the other two cases were brought by digital rights organizations in civil society. The strong stance that the CJEU has taken on data privacy foreshadows a period of growing litigation as a strategy to increase protection, in a field where case law was notoriously sparse.

Beyond the DPD, there are a number of other laws and rules at the European level that are relevant to data brokers and their customers. First, there are sector-specific rules for electronic communications services, in the EU ePrivacy Directive.[129] This Directive requires informed consent when a service provider stores or accesses information, such as a cookie, on a user's device, also when no personal data are involved.[130] Second, there are several EU Directives on consumer protection, which are complementary to the rules on data protection.[131] In areas

such as consumer credit, where data brokers play a major role in certain European countries, consumer protection rules can safeguard consumers against unfair decision making, including those that are data-driven.[132] Third, the EU has adopted several non-discrimination laws that apply to employment, access to goods and services (including housing) and racial and gender equality.[133]

Finally, there are data protection rules for sectors that are not covered by the DPD and GDPR, such as the recently adopted Data Protection Directive for Law Enforcement (DPDLE), which seeks to harmonize national laws for the handling of personal data by the police.[134] Currently, data protection rules for law enforcement are a matter of national law, outside EU-level arrangements for the cross-border sharing of personal data and the fundamental rights safeguards established by the ECHR and Convention 108.[135] Ambitiously, the DPDLE goes beyond the scope of such arrangements, and will harmonize data protection standards among domestic law enforcement agencies.[136] These standards recognize "an increased awareness for the dangers resulting from profiling methods used in LE data processing."[137] Like in the situation of the GDPR, however, these safeguards are vague and provide for broad exceptions. It is unclear whether the DPDLE's harmonized standards, which are less precise than the safeguards in the GDPR, will result in higher levels of protection in practice for subjects of personal data processing by the police.[138] The DPDLE's apparent primary goal is to "make it easier for . . . criminal law enforcement authorities to work together in exchanging information."[139] Public debate about the DPDLE has been minimal.

2.2.1 The EU regime creates an ambiguous landscape for data brokers.

There is no authoritative report about how European data protection law applies to data brokers, and no coordinated enforcement action against data brokers at the EU level. There has also not been coordinated enforcement against data brokers comparable to the enforcement actions taken against Google and Facebook in recent years. At the EU level, the Article 29 Working Party has not specifically addressed the application of the rules to data brokers. And in debates about the GDPR, data brokers also did not play a significant role.

As previously mentioned, the GDPR will have a major impact on the EU data protection landscape—including on data brokers. In particular, the GDPR will further harmonize data protection law across the EU. But the possibility of exceptions, along with divergent interpretations, legal cultures, and other relevant national laws that lack harmonization, will remain. However, a more uniform set of rules could decrease fragmentation of the existing data broker market and lower regulatory barriers to entry and incentivize consolidation, even as limitations on data broker practices remain in place.

The most relevant changes introduced by the GDPR for data brokers include:

- Enhanced data subject rights, including a right to erasure, a right to restrict processing and stronger protections against the use of sensitive personal data in profiling.

- Regulatory recognition of pseudonymization, which is considered a protective measure that helps data controllers to fulfill requirements for security and data protection by design, and to avoid restrictions on processing for new incompatible purposes and data mining.

- Increased clarity about the limitations on the use of personal data for data mining and other statistical purposes, as well as for purposes beyond the purposes when the data was first collected.

- The introduction of a risk-based approach to compliance and enforcement and the codification of a variety of governance instruments, such as audits, privacy impact assessments and certification regimes.

- Significantly stronger enforcement powers of the DPAs (with fines up to 4% of global turnover) and better enforcement coordination among DPAs.

The application of the European data privacy framework to data brokers involves three categories of legal questions, each with their own ambiguities. First, what data broker practices are covered by data protection rules? Second, what procedural safeguards and requirements apply? And third, what hard prohibitions and limitations apply, if any?

The threshold questions—about what data broker practices are even covered—can be a significant challenge to enforcement of the law. Data brokers may claim to deal with anonymous data, or deny being a data controller, or structure their operations in order to avoid EU jurisdiction.[140] The significantly increased non-compliance risk introduced by the GDPR is likely to have a positive effect on the resolution of these questions that has dominated many enforcement discussions until now. Recent CJEU case law on the territorial scope of EU data privacy law has already clarified some of these questions in ways that simplify enforcement.[141] A CJEU judgment on the definition of personal data is forthcoming.[142]

When a data processing operation involving a data broker is covered, transparency and data subject rights apply. However, their precise application as well as the application of other procedural guarantees, such as whether consent is required, can be challenging in practice.

Key questions about the extent to which a data broker can rely on the balancing provision for the lawfulness of processing, or whether it is allowed to use personal data for new purposes or include them in data mining operations, are vague and subject to debate.[143] The GDPR does mention fraud prevention, direct marketing and network security as examples of legitimate interest purposes, but this only clarifies half of the balancing exercise in those cases.[144] Outside of the purposes for which data brokers can claim a legitimate interest ground, data broker activity is restricted quite significantly by the GDPR's clarified consent requirements.[145] These requirements include a presumption that omnibus consent is not valid,[146] which could limit the number of data broker sources that rely on consent.

The DPD and GDPR do place some hard limits on data brokers' activities. For example, both require that data brokers obtain explicit consent when processing sensitive categories of data.[147] There are clear opt-out provisions in the GDPR for direct marketing, and profiling for direct marketing.[148] Finally, there is a restriction on profiling and automated decisions which appears to set a hard limit.[149] However, upon closer examination, this restriction is limited to cases that have a "significant impact" on the data subject, and even then, there are further exceptions that imply the protection is mostly procedural: a right to transparency, a right to intervention by a human decisionmaker, and the ability to express one's point of view and to contest the decision.[150]

In sum, the EU data privacy framework provides a broad set of data privacy guarantees that limits the activity of data brokers and provides for relevant rights and safeguards for individuals. In particular, the purpose limitation requirement poses a significant challenge for data brokers that would like to serve customers across different markets and purposes. Purpose specification and limitation requires that data are collected for specific purposes. The collection and resale of personal data for a yet-to-be-determined purpose, is generally not permitted under the EU legal framework. However, despite the seemingly significant legal differences between the EU and the US, for some data brokers, especially those in the field of marketing and credit that are aggressive in their interpretation of the law, operating in the EU is possible and the differences may not ultimately be that material today, save for a substantially higher overhead compliance cost on the EU side.

3. Data Brokers in Context

In the discussion that follows, we describe the impact of data brokers and profiling in three important domains of daily life: marketing, consumer credit, and policing. Our goal is to illustrate both the variety of ways in which data brokers can impact people's daily lives, and also the ways that public policy does, or does not, constrain their activities.

3.1 Marketing is a driving force behind emerging data brokerage and profiling.

Many data brokers thrive by providing products for sales and marketing purposes: data and predictions about consumers that help businesses optimize their commercial offerings. These data brokers often work hand-in-hand with large online advertising platforms, such as Facebook and Google, to help target advertisements. They help fuel individualized treatment of consumers across a variety of different channels, including email, social networks, mobile apps, postal mailers, and in-store purchases.

Marketing is a key commercial rationale for today's online digital environment, where many valuable online services are provided at no direct financial cost to users. Personal information and human attention have become a quasi-currency. Security and privacy expert Bruce Schneier has written that surveillance has become the "business model of the Internet."[151] He observes that data brokers and tech firms have created a "shockingly extensive, robust, and profitable surveillance architecture," and notes that people are "being tracked pretty much everywhere you go on the Internet, by many companies and data brokers: ten different companies on one site, a dozen on another."[152]

Data brokers bring these new data sources together to help businesses understand and target consumers. At their best, data brokers enable more efficient commerce, lower search and transactions costs, personalize product offerings, and support free Internet services. However, at their worst, these practices expose vulnerable individuals and communities to new risks, exacerbate inequalities, and erode people's privacy.

3.1.1 Data brokers are making innovative uses of new data, sometimes hand-in-hand with major Internet platforms.

Online data collection practices have evolved quickly in recent years. Today, data brokers observe people's behavior across many websites, making sophisticated use of browser cookies and other technologies.[153] Other data companies have worked with Internet Service Providers (ISPs) to inspect customer traffic in real-time and obtain "access to all or substantially all of an individual's Web traffic as it traverses the ISP's infrastructure, including traffic to all political, religious, and other non-commercial sites."[154] The techniques that are used to identify and combine the behaviors of specific individuals across different channels are getting more sophisticated. For instance, one company emits high-pitched "audio beacons" from television commercials and online ads. These beacons are then recognized by other devices in the room—say, a user's phone or computer—which allows advertisers to know "which ads the user saw, how long the user watched the ad before changing the channel, which kind of smart devices the individual uses, along with other information that adds to the profile of each user that is linked across devices."[155]

Data brokers can combine offline data with online data to produce a wide range of predictions about consumers' behaviors and likely interests.[156] Ad networks (and their advertisers) can then act on these predictions to target consumers both online and off. For example, a data broker might sell the ability to target luxury car owners online. Or, a data broker could help a business (like an insurance broker) to analyze its existing customers to identify new, similarly situated customers.[157]

Large online platforms, such as Facebook, Twitter or Google, work hand-in-hand with data brokers such as Acxiom, Datalogix, and Epsilon to target advertisements and optimize the effectiveness of online ads.[158] Data brokers allow advertisers to use consumers' purchase histories, as well as other online and offline behavior to target ads.[159] For example, an advertiser could target "children's cereal buyers" (relying on data collected and analyzed by third-party data providers) who live in Washington, D.C. (relying data that a user has provided directly to Facebook or Twitter). Moreover, using both on-site and off-site data, Facebook and Twitter help marketers create "lookalike audiences," which allow marketers to show ads to people who are similar to their current customers.[160]

3.1.2 Studies of data-driven marketing practices speculate about a range of harms, but most have struggled to substantiate these concerns.

There have been several major reports from US policymakers expressing concern about data broker practices in the marketing context. In Europe, online tracking and behavioral targeting have informed legislative agendas and outcomes and shaped the enforcement activities of regulators. However, on the whole, policymakers, civil society groups, academics, and journalists have struggled to articulate concrete harms related to emerging data broker practices, aside from an abstract erosion of privacy and a lack of awareness of individuals about what is going on. These difficulties are likely to persist.

One leading concern is that data brokers expose vulnerable populations to those offering predatory products. The US Senate report warned that the data sold by some brokers is "likely to appeal to companies that sell-high cost loans and other financially risky products," and the FTC observed that many would find it "disconcerting," to know that products can easily be targeted at disadvantaged people.[161] Data brokers sell marketing lists with titles like "'Rural and Barely Making It,' 'Ethnic Second-City Strugglers,' 'Retiring on Empty: Singles,' 'Tough Start: Young Single Parents,' and 'Credit Crunched: City Families.'"[162] The FTC's report also highlighted segments focused on minority communities and low-income individuals, including one called the "Urban Scramble."[163] However, there are have been few studies explaining how these products are actually used. And in some cases, they may be useful to entities with laudable intentions, such as political organizers or non-profit community lending organizations.

Another concern is that companies will use data about people to conduct "differential pricing."[164] There are at least two kinds of differential pricing. The first, "risk-based pricing," occurs when a business prices a product based on the cost of providing it to different groups of buyers. The second, "value-based" pricing, occurs when a business prices a product based on buyers' willingness to pay. Risk-based pricing is common in the insurance and credit markets, and has been for many years. Evidence of value-based pricing has, by contrast, been very limited. For example, in the US, a box store website was found to price its staplers differently based on where the company thought the customer was located.[165] In another case, the travel website Orbitz showed pricier hotels to users who used Mac computers.[166] Although illustrative of value-based pricing, neither of these cases were particularly consequential, and neither involved data brokers.

Yet another concern is that new data will be used to discriminate unfairly against some consumers. For example, the White House released a report in May of 2014 that expressed concern that detailed consumer profiles might lead to race or income-based discrimination.[167] It cited recent research by Latanya Sweeney, a computer science professor at Harvard, who described how

Google ads differ based on the name of the person searched.[168] Sweeney explained that a greater percentage ads with "arrest" in their text appeared for black-identifying names than for white-identifying names, to an extent that could not plausibly be explained by chance. This happened because Google's software automatically learns which ad combinations are most effective (and most profitable) by tracking how often users click on each ad. These user behaviors, in aggregate, reflect the biases that currently exist across society. However, although Dr. Sweeney's research helped show that racism can be perpetuated by complex online systems, even when the companies that create these systems do not intend to discriminate, this was not an example of how data brokers, collecting personal information, might themselves worsen the problem.

Finally, there is a category of so-called "mission creep" concerns, where marketing data is put to non-marketing uses. For example, some commentators highlight that government surveillance efforts can "piggyback" on data brokers' collection activities.[169] And credit scores, traditionally used for underwriting loans, are now used by nearly half of US employers to screen job applicants.[170]

In short, the harms of modern, data-driven marketing include a lack of transparency, and might include discriminatory profiling and chilling effects on expression and commerce. However, despite sustained attention, substantiating these harms has proven difficult.

3.1.3 US regulators have few powers with which to act, and EU regulators are struggling to enforce existing rules effectively.

In the US, there has been widespread concern from regulators and calls for greater transparency and choice when it comes to data broker practices in marketing. The GAO concluded that the current US statutory framework "does not fully address new technologies—such as the tracking of online behavior or mobile devices—and the vastly increased marketplace for personal information."[171] A Senate Committee report stressed that "it is important for policymakers to continue vigorous oversight to assess the potential harms and benefits of evolving industry practices . . . ."[172] And the FTC has urged Congress to require marketing data brokers "to provide consumers access to their data, including sensitive data held about them, at a reasonable level of detail, and the ability to opt out of having it shared for marketing purposes."[173] However, despite these calls, even public-private standards to curtail online data collection, such as "Do Not Track" have also been unsuccessful. Regulators continue to study emerging marketing practices.[174] But, so far, a mandate for greater regulation has been elusive.

In the EU, regulators are still struggling to enforce the DPD effectively. The large consumer-facing online platforms, including Google and Facebook, have received most attention. Most

recently, the Norwegian Data Protection Authority has conducted a study on issues related to online marketing, which also discusses the role of data brokers to some extent. The report concludes that still more information is needed.

Thus, on the whole, while the role of data brokers in marketing is a central focus of policy discussions in the US and Europe, it does not appear to be a likely area for high impact intervention in the near term.

3.2 In consumer credit, brokers follow clear rules, and provide useful data.

On both sides of the Atlantic, lenders pool credit information about consumers into centralized databases managed by third-party data brokers known as credit bureaus.[175] For many years, these credit bureaus have played a central role in allowing lenders to evaluate consumers for financial products and services. Credit bureaus, together with analytics firms, are part of a larger "credit reporting" industry that has fueled large-scale automation of consumer credit decisions.

Few dispute the value of credit reporting in the abstract. Access to credit is an important part of building wealth. Credit is often necessary to buy a car or a home, to build a business, and to send one's children to college.[176] And creditors need timely and accurate data about people to make sound lending decisions. This well-regulated, and reasonably well-understood corner of the data broker landscape has an important role to play in people's daily lives. However, there will always be room for improvement.

Although credit bureaus control access to credit, policy frameworks can impact the decisions bureaus reach.[177] For example, in the US and much of Europe, the collection and redistribution of credit data is a for-profit enterprise.[178] Private credit bureaus often prioritize products that enhance efficiency and profitability of lenders. However, some European countries operate state-run credit bureaus with the primary purpose of protecting borrowers from assuming too much debt.[179]

In contrast to data brokers in other contexts, the bureaus that broker credit-related data are part of a longstanding and well-understood industry. In the US, beginning in the 1950s, a network of local, private "bureaus" arose to provide lenders with more data about consumers. These bureaus initially gathered data through interviews and other labor intensive techniques.[180] In the decades following, new regulations and the digitization of bureaus' files drove a rapid consolidation of the industry—resulting in the "big three" private credit bureaus the dominate

the US credit reporting market today.[181] The European picture is more complex and fragmented, with different-sized credit bureaus (many of whom predated the European Union) serving different countries under different regulatory regimes.[182]

Credit bureaus provide data that ultimately helps lenders make decisions about whether and under what terms to offer consumers financial products. In many cases, these decisions are entirely automated.[183] This close relationship between lenders, credit bureaus, and analytics providers can make it seem as if credit decisions have been completely delegated by the lender, despite the fact that the lender has the "final say" (and, of course, has her choice of underwriting methods). For many consumers, this results in efficient and fair decisions. However, people can be harmed if the credit reporting industry lacks data about them, if their data is inaccurate, or if their data is used in unfair ways.

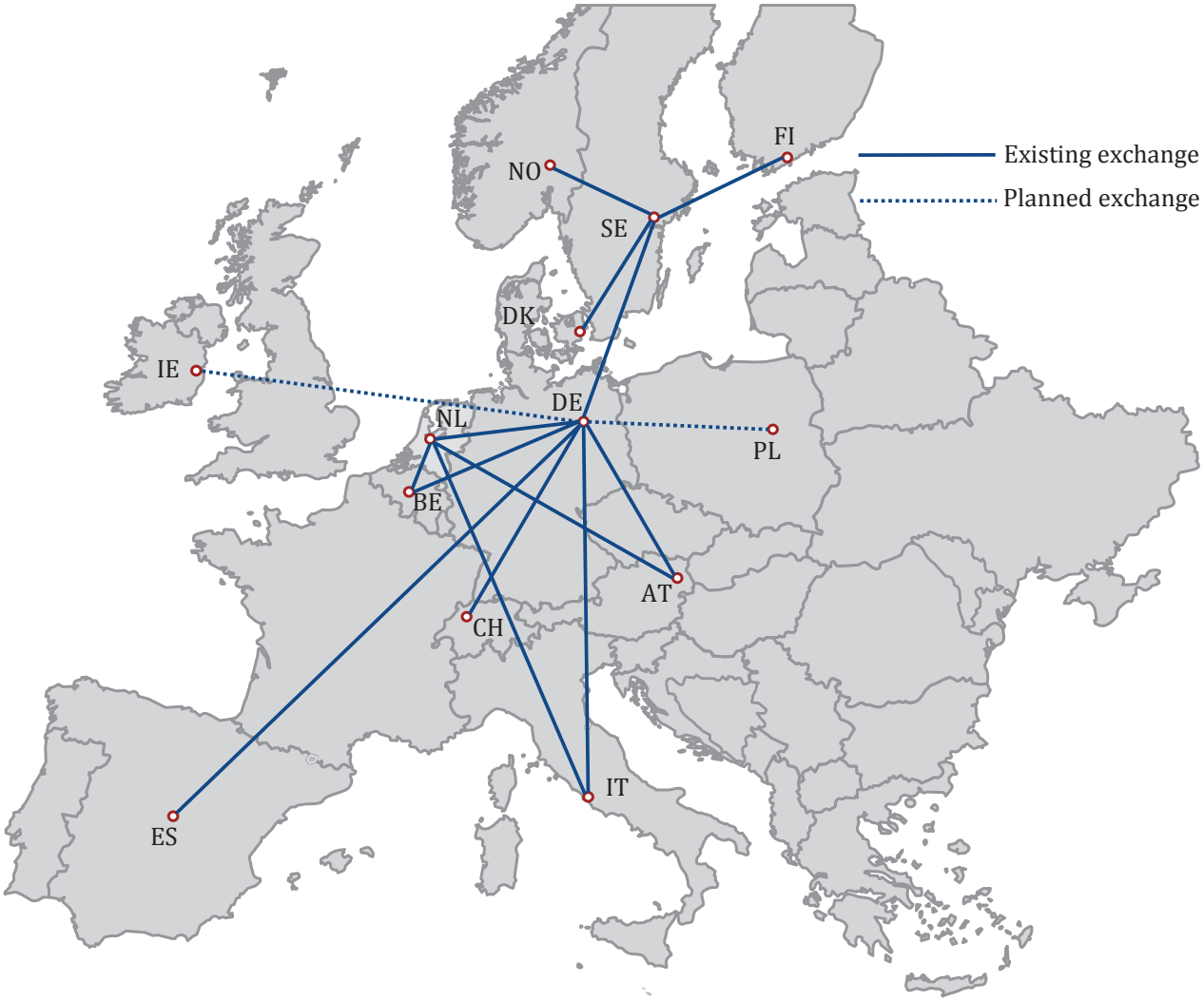
3.2.1 Credit bureaus, and their reports, shape individual access to credit.

In the US, three national credit bureaus—Equifax, Experian, and TransUnion—dominate the industry.[184] Fair Isaac Corporation (FICO), an analytics company, helps US credit bureaus convert the data they have about people into three-digit credit scores.[185] In Europe, the picture is more varied: European credit bureaus operate nationally with limited cross-border initiatives and use different procedures to develop credit scores.[186] In some markets, including Belgium and France, only public credit bureaus operate.[187] However, in most countries, including Cyprus, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Malta, the Netherlands, Poland, Sweden, and the United Kingdom, credit reporting is left to free-market forces.[188] Other countries, including Austria, Bulgaria, Germany, and Italy, have both private and public credit bureaus.[189] Cross-border credit reporting is still in its infancy: As of 2010, a minority of credit bureaus surveyed by an industry study accessed credit data across borders.[190] (See Figure 1.)

Despite these many variations, credit bureaus perform a fundamentally similar role: building "credit reports" about people. Credit reports typically contain a limited set of credit history data provided by financial institutions. Credit history data includes how a person has banked, borrowed, repaid their debts, and paid their bills. (Financial institutions provide credit bureaus with this data because they benefit from credit reporting and because bureaus often require that they do so in exchange for access to the credit bureaus' databases.) Credit history data can also include a variety of legal judgments, including liens (claim against property for debt owed) and bankruptcy. And, increasingly, the credit reporting industry is exploring the feasibility and usefulness of collecting other types of consumer payment behavior, such as rental records, utility bills, and telecom bills.[191]

Figure 1

Formal cross-border data exchange agreements in Europe among credit bureaus



Source: Association of Consumer Credit Information Suppliers, ACCIS survey 2010

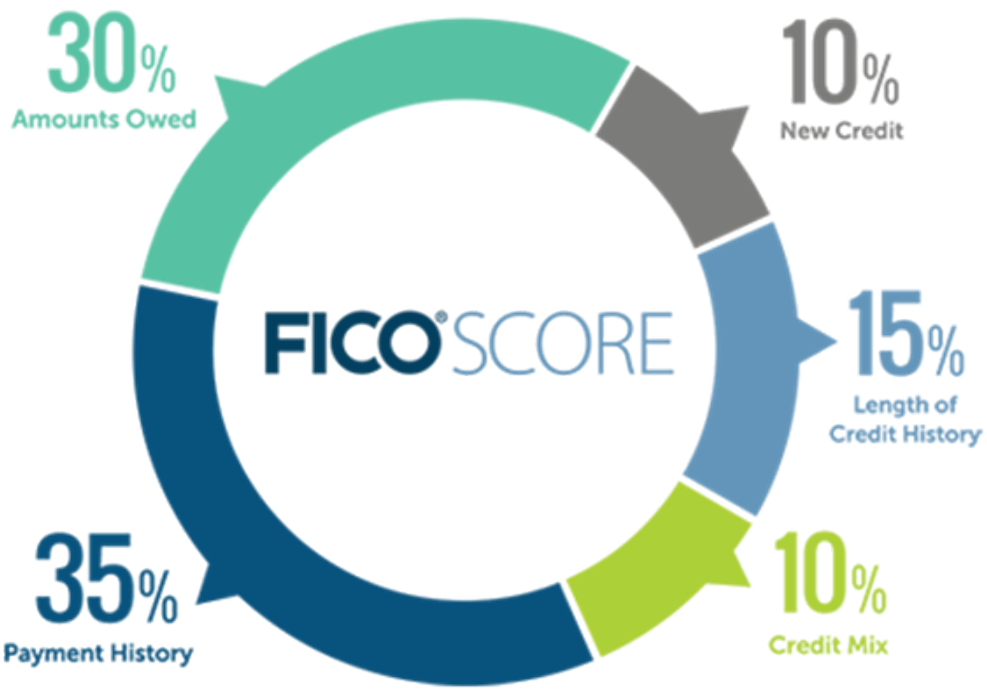
Credit bureaus amass credit history data because it is the most useful data to predict whether a person is likely to pay back a loan. For example, repayment history data, standing alone, can be used to identify 77% of the defaulting population in the US.[192] Other kinds of data can be helpful, but they deliver sharply diminishing returns. For example, a person's preferences about where to shop for certain consumer goods delivers almost no additional predictive value when combined with credit history data.[193] In short, for the purpose of predicting creditworthiness, credit bureaus don't want just any data: they want the right *kinds* of data.

To understand a consumer's credit report more quickly and automatically, lenders typically turn to credit scores: a summary of a person's apparent creditworthiness derived from the person's credit history via a statistical model. Most credit scoring models are proprietary, but their basic

workings have been publicly documented and are reportedly similar across the industry.[194] Companies build credit models by comparing snapshots of data from the same group of individuals at different moments in time. They then isolate characteristics that correlate with default between the two snapshots.

Figure 2

The basic recipe for a FICO credit score, which is commonly used in the United States



Lenders use credit scores as an important factor—and often the only factor—in making lending decisions. For example, in the US, credit scores are almost always a number that ranges between 300 and 800. This one number, delivered by a credit bureau (sometimes with the assistance of third-party algorithms), can effectively decide whether and under what terms a consumer receives a loan. A good credit score can mean access to a wide range of credit products at better rates, while a bad credit score can lead to greatly reduced access to credit and much higher costs.[195]

3.2.2 Modern credit reporting, though imperfect, often benefits consumers.

The best available evidence suggests that modern credit reporting, including the brokerage of credit data, is helpful to consumers. The European Commission's Expert Group on Credit Histories recognized that "[c]redit data sharing between creditors is considered an essential

element of the financial infrastructure that facilitates access to finance for consumers . . . [and] assists creditors in complying with responsible lending obligations."[196] The World Bank claims that a wide range of stakeholders are recognizing that credit reporting can increase "financial supervision and financial sector stability" as well as "enhance access to credit."[197] It noted that to accomplish public policy objectives, credit bureaus must have "relevant, accurate, timely and sufficient data . . . collected on a systematic basis . . . ."[198] And the US Consumer Financial Protection Bureau (CFPB) observed that consumers whose data is not captured by the credit reporting industry "face significant challenges in accessing most credit markets."[199]

Credit scoring, the act of profiling individuals based on their credit history, has its own benefits. The US Federal Reserve, in a thorough study of the issue, concluded that credit scoring based on credit history data "has increased the availability and affordability of credit."[200] It claimed that credit scores allowed creditors to offer expanded access to previously credit-constrained populations.[201] It also found that these credit scores do not overestimate credit risk among minority groups, nor do the data typically contained in US credit reports allow sensitive factors (such as race) to influence the score.[202]

Despite these benefits, some advocates claim there is a need to fundamentally rethink how credit reports are structured and how creditworthiness is judged. For example, some groups believe that today's credit data and credit scoring techniques too harshly penalize those who have simply fallen on temporary hard times. As the US-based National Consumer Law Center argues, "We need a system that can distinguish between consumers who are truly irresponsible and those who simply fell on hard times. We need a system that can take into account both economic factors and extraordinary life circumstances particular to an individual consumer."[203]

3.2.3 Errors and missing data are primary risks to individuals.

Given the entrenched and central role of credit reporting to the modern credit marketplace, people can suffer when the system lacks data about them. In the US, consumers with limited data in the three largest credit bureaus face "significant challenges in accessing most credit markets."[204] In 2010, 26 million American consumers were so-called "credit invisibles." An additional 19 million consumers do not have enough data to be scored by commercially available scoring methodologies. In total, almost 20% of Americans thus face challenges in accessing mainstream credit.[205] These harms are disproportionately borne by Blacks, Hispanics, and lower-income consumers.[206] EU commentators have similarly noted that "accessibility to full credit and other noncredit data may affect the inclusion, exclusion, or sorting within different economic spheres of the consumers."[207]

Inaccurate data can also hurt people's prospects of obtaining fair and affordable credit. In a 2012 study of US credit reporting, the FTC found that "one in five consumers had an error that was corrected by a [credit bureau] after it was disputed on at least one of their three credit reports."[208] Sometimes, these errors "resulted in a decrease in their credit risk tier, making them more likely to be offered a lower auto loan interest rate."[209]

Looking to the future, credit scores may one day be commonly generated using more exotic kinds of data, which would raise new concerns. For example, some commentators have observed that a person's professional contacts are "especially revealing of an applicant's 'character and capacity' to repay."[210] Facebook recently patented, but has not publicly implemented, a method for gauging a person's creditworthiness based on the creditworthiness of their friends.[211] Even data such as "how many times a person says 'wasted' in their profile, it has some value in predicting whether they're going to repay their debt," acknowledged a FICO employee.[212] However, methods such as these are not widely deployed in the US or the EU, due in part to regulatory restrictions that favor methods already demonstrated to predict repayment and already scrutinized for potential racial or other biases.

3.2.4 Domain-specific laws play an important role in protecting individual rights.

Both the US and the EU have well-developed legal frameworks that help to protect individual rights in the context of credit reporting. In the US, the FCRA, a law passed to address the emergence of credit bureaus, helps prevent oversharing of consumers' data and provides some guarantees of access and accuracy. A complimentary law, the ECOA, prohibits decisionmakers from using certain sensitive data in credit decisions. Together, these laws have regulated both credit bureaus and lenders for decades.

In the EU, Directive 87/102 resulted in the fragmentation and segmentation of credit markets into separate, national entities.[213] Despite this lack of harmonization, the general provisions of comprehensive law (*e.g.*, Directive 95/46) still apply with force. Many credit bureaus use the "legitimate interests" ground to justify their activities, and typically provide some level of accuracy and access.[214] Other EU-level laws prohibit discrimination, although "debates over the impact of credit scoring on communities of color and other protected groups, including minorities, are almost absent in the EU."[215]

On the whole, data brokerage in the credit context is well regulated with tailored, domain-specific controls.

3.3 Policing: Brokered data can add bias and noise to criminal justice decisions.

Law enforcement agencies have a long history of using data to investigate and solve crimes, and to keep communities safe. Before computers, agencies that "did crime mapping relied on primitive techniques such as sticking thousands of pins into large maps attached to the wall."[216] Agencies use data to determine where and when to patrol, who to approach, question and investigate, and who to cite and arrest. For the most part, the data that agencies collect on their own—to the extent that they actually do—is primarily local, based on their own investigatory and policing activities, and based on what their members learn on the street.

With modern technologies, law enforcement agencies have access to public and private records far beyond the agencies' traditional reach. Agencies share data with each other and have many avenues to gain access to data from the private sector. Agencies in the US can subscribe to commercial data broker products that are purpose-built for law enforcement. In Europe, however, there is no visible market for similar data broker services aimed at law enforcement. Law enforcement agencies in Europe may in fact be using commercial "people search" or other products built on brokered data, but we did not find any evidence of this in our investigation (for example, we did not find any marketing material that specifically targets brokered data to the European law enforcement market.)

In both the US and Europe, however, law enforcement relies heavily on corporate vendors for profiling and analysis tools—systems that analyze personal data held by the government, instead of or in addition to analyzing data held by the vendor. For example, in Europe, the Thales Group helps law enforcement establish new capabilities derived from data that the government itself gathers and holds.[217] And Palantir, a leading US-based vendor to law enforcement and intelligence agencies, sells data analytics tools, rather than offering new data to its clients.

One of the leading data broker products on the US market is Accurint for Law Enforcement Plus, a person-search product provided by LexisNexis. The company markets the tool to law enforcement as a way to easily "locate suspects, witnesses and fugitives," "quickly uncover assets," and "discover links between people, businesses, assets and locations."[218] Using Accurint, agencies can purportedly "maximize budget and resources, enhance officer safety, solve cases faster, [and] reduce crime rates."[219] Similarly, a competing product from TransUnion, called TLOxp for Law Enforcement, "offers the largest, most powerful online database of public and proprietary records available providing information about people, businesses, assets, and locations" backed by a "trillion-record database."[220] These products essentially tailor existing commercial search products, and all of the data behind them, specifically for law enforcement purposes.

Like commercial products, brokerage and profiling products for law enforcement use vast collections of both public and private data. Much of the data that is targeted to law enforcement appear to be derived from public records, like court records (to collect criminal history, bankruptcies, and foreclosures), birth and death records (to establish familial relationships), driver's and business licenses, and other asset and property records. Data brokers in the US often station "stringers" at courthouses and other government offices to copy and collect such publicly available information nationwide. Even if these products only contained public information, they would already be immensely useful to police departments, who could now instantly access up-to-date, organized information from public sources far beyond their own jurisdiction. But, of course, through their business arrangements with private sources and by accessing publicly accessible data, data brokers also include large amounts of privately held data—like e-mail addresses, social media, news reports, vehicle sightings, and other location information[221]—to create custom-built search and mapping tools for the police in the US.

Such investigatory tools are already widely used in the US and elsewhere. LexisNexis claims that its Accurint tool is used "by over 4,000 federal, state and local law enforcement agencies" in the US—representing nearly a quarter of the 18,000 agencies in the country. TransUnion boasts that TLOxp is used by "over 100,000 law enforcement officers across the country." Despite wide adoption, revenue from law enforcement and intelligence agencies represent a relatively small portion of a data broker's overall revenue: for instance, in 2010, less than 2% of LexisNexis Risk Solutions' revenue came from US law enforcement clients.[222] But while these products may only be a small part of the overall data broker market, the use of these tools by law enforcement can have an outsized—and potentially life-changing—impact on people's lives.

3.3.1 Police use of brokered data and corporate profiling threatens fundamental rights.

When law enforcement relies on data broker search tools, even minor inaccuracies in the data could lead to dire consequences. A simple case of mistaken identity could lead to a wrongful arrest, or in a worse case could lead to officers using significant force against the wrong person. Even in less critical situations, officers could place intense scrutiny on a misidentified individual, causing long-term reputational and emotional harms.

Errors in individual profiles could arise in any number of different ways. People often have similar names, and their records could have accidentally been combined. When paper documents are scanned and digitized, the automated character recognition program could have misrecognized a digit of an ID number, linking to a different individual. Data could also easily fall out of date: as LexisNexis' own marketing material ominously suggests, "We don't throw away records.

We keep the oldies and the goodies."[223] Missing data can also paint an inaccurate picture: a record that indicates a pending felony charge has different consequences than one with a felony charge followed by a "not guilty" disposition.

There are no known studies on the accuracy of these tools and the data that support them. But it's safe to assume that these databases are riddled with errors. Even in the credit context—where the accuracy of credit reports is financially important to banks, and where the collection and use of data is highly regulated—inaccuracies are extremely common. A 2012 FTC study found that approximately 26% of consumers reported "potentially material errors on at least one credit report."[224] The data sources that drive these law enforcement tools are far more diverse and far less regulated than in credit, so it's easy to imagine how prevalent errors might be.

But even when the data is accurate, the use of data broker tools can exacerbate existing biases in policing and the broader criminal justice system, particularly for poor communities and communities of color.

3.3.2 Social media data can drive bias.

Person-search tools make it easier for law enforcement to locate potential suspects. But in a criminal justice system where certain communities are already overrepresented as suspects, making suspects easier to find further intensifies existing disparities. With broad officer discretion and cheap search tools, any inkling of suspicion can now more easily lead to an arrest, reinforcing the biases that already exist.

A number of small niche data broker products are designed specifically for police use in the US. For example, a product called BlueJay is advertised as a "law enforcement Twitter crime scanner"[225] that allows an officer to easily monitor selected users, keywords or geographic areas (such as the location of a protest). These tools may focus on information that users have chosen to make public, such as published Tweets, so they do not raise "privacy" concerns in the traditional sense[226]—but they nonetheless provide police with a new, easy and low-cost way to surveil people. The low cost of these tools makes them easier to use on a whim and creates new risks that officers will use them in biased ways.[227] For example, the NYPD have surveilled the activities of school-age children suspected of gang activity (including children who do not have a criminal record), and have used potentially innocent actions such as "liking" a photograph on Facebook as the basis for arrest.[228]

Biases in the databases themselves, based on how data are collected, may also lead to disparate outcomes. Those who have already been involved in the justice system, and thus are identified

in court records, will have more negative information about them in these databases. Although such tools do not yet exist in Europe, they may yet emerge under the GDPR's companion law enforcement directive.

3.3.3 Brokered license plate histories provide disparate visibility into heavily surveilled neighborhoods.

TransUnion's TLOxp product provides data on vehicle sightings, an ever-growing history of vehicle date, time and locations collected by automated license plate reader (ALPR) scans across the US.[229] By subscribing, agencies gain access to "a massive database of more than a billion vehicle sightings and the addition of up to 50 million sightings added monthly."[230] But what biases may exist in their database?

> "While the coverage is nationwide, certainly there will be areas with more expansive coverage than others," said James Reilly, TLO's senior vice president of sales and business development. "Variables such as the amount of time the vehicles are stationed in inaccessible areas (i.e. secured lots at places of employment, gated communities, etc.) could certainly affect the number of opportunities for 'sighting.'"[231]

It's not exactly clear where TransUnion purchases its vehicle sighting data, but it's well known that such databases are often powered by repossession companies, who send out "spotter cars" into certain neighborhoods to locate and impound vehicles that are identified as either stolen or in default:

> "Honestly, we've found random apartment complexes and shopping plazas that are sweet spots" where the company can impound multiple vehicles, explains Sousa, the president of New England Associates Inc. in Bridgewater. . . . Two repossession companies also told BetaBoston that they focus on low-income housing developments, since a significant number of residents are delinquent on their car payments.[232]

Vehicles from low-income communities are overrepresented in these databases compared to those from rich, gated communities. This makes it far easier for law enforcement to track the whereabouts of low-income individuals, simply because of how data brokers buy and assemble data.

3.3.4 The legal and policy tools to restrict police from using brokered data and corporate profiling are limited.

In the US, there are few legal limitations on government access to commercially available databases. The Fourth Amendment prohibits unreasonable searches and seizures. But the Supreme Court decided in the 1970s that individuals do not have a reasonable expectation of privacy in information held by third parties, like banks and telephone companies.[233] While the "third-party doctrine" might have made sense in decades past, the recent proliferation of data collection by private industry, including data brokers, about the intimate details of people's lives has given law enforcement unprecedented and easy access to personal information.

In 1974, the US Congress passed the Privacy Act "in response to concerns about the creation of large, centralized governmental databanks of personal information."[234] It is "the closest analogue to a European data protection law."[235] The Act established comprehensive rules on the government's collection, use, and management of personal information—but only for *federal agencies*, not for commercial entities, nor for state and local agencies, like local police departments. In addition, the "protections apply only where the government is creating a 'system of records'."[236] When it comes to government use of commercially available data products, "searches and data analysis can be conducted in such a way that the data never leaves private hands,"[237] meaning that the Privacy Act protections wouldn't apply.

As a result, aside from limited sector-specific protections, US agencies at all levels can simply purchase and use extensive commercial data broker products for law enforcement purposes, without any need for a subpoena, warrant or other legal process.

In Europe, the relevant legal frameworks for the collection and use of data by police agencies are predominantly national. At the European level, the relevant protections in the European Convention on Human Rights apply as well as Convention 108 and the Cybercrime Convention. Along with the GDPR, the EU has adopted new rules to harmonize data protection across law enforcement agencies.[238] These rules will apply to the handling of personal data (including data collected from private data brokers) by law enforcement. In addition, data brokers that cater to law enforcement are still covered by the DPD and future GDPR rules that govern their private sector services. As discussed above, these rules limit a data broker's ability to develop law enforcement products comparable to those offered in the US.

While there are no data broker products for law enforcement similar to those that exist in the US, there is an active market in Europe (as in the US) to help law enforcement agencies collect and use new types of data. The use of social media and other publicly accessible data on the Internet by police agencies is one example. Under European law, systematic monitoring and

registration of publicly accessible, online behavior by the police is considered an interference with subject's rights, but may nonetheless be authorized if it fulfils the requirements of Article 8(2) ECHR: it must be provided for by law, pursue a legitimate aim and be necessary in a democratic society.[239]

Law enforcement agencies' intensive processing of personal data raises significant concerns for rights and justice. But the corporate role in profiling and analysis, when it comes to law enforcement, is at least as important—and perhaps more important—than the corporate role in providing the data itself.

4. Strategic Next Steps and Open Questions

Powerful institutions are rapidly adopting new ways to gather, analyze, share and use data about the people with whom they interact. In order for societies to remain open, the social sector—and openness-enhancing offices of the government, including courts and market regulators—must continue to hold key institutions accountable, even as these new and technologically mediated modes of decisionmaking become widespread.

In order to ensure that this accountability happens, public interest stakeholders will need to pay careful attention to the rapidly evolving flows and uses of data, and will need to reflect on what this changing landscape means for longstanding social justice goals.

Data brokers are one important type of actor in this new ecosystem. This report aims to provide a baseline level of understanding of how these actors operate.

However, a core strategic finding of this report is that, **although there is ample reason to be concerned that new flows and uses of personal data, including some facilitated by data brokers, may threaten rights and justice, it would not be productive for researchers and advocates to focus their efforts primarily on the activities of "data brokers" as such**. There are several reasons for this:

- **Many activities of data brokers pose minimal risk for rights or justice.** The marketing and customer relationship management activities that are a central focus for large segments of the data broker industry are generally much less important for rights and justice than are major data-driven decisions regarding criminal justice, health, education, jobs, housing, and other key civil rights areas.

To the extent that brokers, and the data and inferences they supply, *do* play important roles in these key life decisions, it is indeed important to for activists and regulators to investigate, understand, and potentially regulate or constrain their activities. But when such efforts are warranted, they are justified by the importance or sensitivity of the particular decision being reached, rather than by the fact that the decision is reached with brokered data.

- **The division of labor between data brokers and other stakeholders is flexible and constantly changing.** As described above, brokers provide a broad spectrum of different services to their clients. Sometimes they provide additional details about an individual whose identifying personal data is already held by the client; sometimes they allow for targeting or decisions about a certain group or individual without knowing exactly who is in the group or what the person's identity may be; and sometimes they provide finished inferences or judgments whose sensitive predicates are held back from the client.

  The social sector's concern extends across the entire chain of events and stakeholders connecting an individual's personal data and subsequent decisions, actions or risks that matter to the person's rights. Efforts defined by a focus on data brokers will risk missing the mark because the commercial arrangements of the personal data ecosystem (and hence the role of brokers being regulated, or targeted in campaigns) can so easily be changed. Data brokers do not always exist as free-standing business units. Already, brokers' clients sometimes find it more convenient, for regulatory or other reasons, to exchange data directly with one another.

- **Government offices that combine and share people's personal information, which are functionally analogous to data brokers, may pose an equal or greater threat to rights and justice than the commercial broker industry.** In the US, "fusion centers" combine personal information from federal, state and local police and intelligence sources—including information that these entities have gathered themselves, information that has been routinely disclosed to them, and data they have purchased from brokers. In the Netherlands, news reports show that the detailed records on each citizen collected by the Dutch Tax Authority have been repurposed as a general resource for law enforcement and intelligence, and the Authority has embraced data mining on these records to establish behavioral change in society.[240] In these instances and many others, government offices are playing a role that would count as "data brokerage" if it took place in the private market.

  Moreover, non-broker contractors who serve government offices deserve more scrutiny. Palantir, an analytics firm and government contractor that is helping governments around the world mine insights from personal data, is careful to emphasize that it can use data from a wide range of the customer's existing sources, but its speciality is providing the tools of analysis.[241] In France, the analytics firm Thales boasts that it "help[s] France's Gendarmerie Nationale conduct operations and process operational intelligence data,"

by building a system that "stores information about operations in the field, recording details of every incident and every member of the public concerned" while achieving "full compliance with France's law on information technology and civil liberties."[242]

A clear understanding of activities like these—and rules that govern such activity effectively—are vitally important to ensure that rights and justice are protected.

- **Major online platforms such as Google and Facebook are not brokers, but their activities raise substantial concerns for rights and justice.** For many active users of Google and Facebook, these companies may know more about their lives than any other institution, corporate or governmental. Rather than sharing the rich data they possess about their users, these platforms instead have engineered a range of ways of allowing customers to use personal data without gaining full access to the data. As a byproduct of their intensive focus on personalized advertising, the major platforms have developed an extremely advanced capability to shape the online experience of their users. Facebook recently experimented, for example, with using social messaging to drive voter turnout. Encouraging voter participation is generally a good thing, but the exercise was also a striking illustration of a capacity that could also be used for troubling ends.[243]

Data-driven decisionmaking can pose significant risks for rights and justice. Some harms, such as chilling effects, may happen solely from the wrong actor possessing certain data. But the great majority of risks involving data profiling come to bad fruition only when a decision or action is taken based upon the data—that is, as a result of how the data is used rather than of who holds it. At one time, stopping data from being gathered or shared may have been easier than constraining its use in decisions. However, we live today in a world of pervasive digital tracking and monitoring, where organizations of all kinds possess a large and constantly growing range of ways of finding things out.

For example, a person with a chronic and potentially stigmatizing health condition (such as sexually transmitted disease) may generate digital traces that make it feasible for many different actors to infer her health status—not only advertisers and data brokers, but also, perhaps, her employer, her health insurer, and various government offices.

Under these circumstances, meaningful governance of data profiling requires regulating a type of problematic activity—whatever profiling needs to be constrained—rather than specifically tailoring rules for the organizations defined as "data brokers" in this report.

In other words, one might argue that an ideal frame for addressing the concerns behind this report would be to scrutinize data-driven decisions, writ large. But data driven decisionmaking

is not a complete focal lens. Data is everywhere, and stakeholders in the social and public sectors need a clear sense of what is most important, to prioritize their limited resources.

We have identified three open questions, below, for further investigation:

1. How can competencies related to data-driven profiling best be applied to existing rights and justice efforts?

The challenge of choosing domains in which data-driven profiling raises the greatest risks and is most urgently in need of scrutiny and advocacy is not a job solely for those who focus on data and technology. Instead, it requires close collaboration between technology experts and those with domain expertise in the core priority areas of social justice.

An occupational hazard of becoming a technology expert is that one risks losing touch with the animating concerns of social justice. For example, much of the advocacy work on privacy on both sides of the Atlantic concerns practices whose impact on people's daily lives is attenuated at best. As we describe above, advocates and regulators have struggled to connect their work to questions that other people—those not working on data and technology issues—would readily recognize as high-stakes questions of justice or human rights.

One strategy that we believe might help with this problem would be to build new bridges between those with technology-related expertise and subject matter experts in each of the high priority areas in which social justice organizations are working to protect rights and advance justice.

2. Which innovations or areas of work related to data-driven profiling would produce cross-cutting gains for social goals across the field?

Even as public interest stakeholders work more closely with subject area experts in various domains of rights and justice, we believe that there may well also be fertile opportunities for work on how data is handled and data-driven decisions are rendered and regulated, that do span the gamut of social justice concerns.

Technology

There are a number of technical efforts under way that aim to enhance the options available for governing potentially opaque data-driven decisions. One area of work is in "black box testing" of systems, where researchers probe opaque systems in order to learn more about how the systems work; another is in redesigning automated decisionmaking systems to ensure that their outcomes comport with pre-specified criteria of fairness or to make them more open to scrutiny. Concrete examples of both types of effort have been inventoried at the "FAT ML" series of academic workshops—Fairness, Accountability, and Transparency in Machine Learning.[244] An earlier, 2012 IEEE workshop on "Discrimination and Privacy-Aware Data Mining" explored similar themes.[245]

Specifically, on the black box testing front, Anupam Datta at Carnegie Mellon and his colleagues have conducted "information flow experiments" designed to infer the algorithmic personalization steps used in some online advertising systems;[246] Roxana Geambasu and colleagues have worked on XRay, "a system that predicts what data—such as emails or searches—is used to target ads in gmail, which prices in Amazon, etc";[247] and a number of other relevant efforts were presented at earlier stages of development. The US National Science Foundation is actively funding work in this area, including a three-year project (begun in late 2015) that aims to test the data broker ecosystem by carefully adding data at one point in the ecosystem and seeing how it is reflected at other points.[248]

Another area of effort involves new technologies and tools for analyzing and describing data-driven decisionmaking, whether facilitated by data brokers or by others. Some pioneering early work in this area took place in Europe, including the 2013 anthology, "Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases."[249] Another landmark work, "Fairness Through Awareness," proposed a computational method for protecting against discrimination by taking explicit account of sensitive categories such as race and gender.[250] These efforts have dealt with simplified, stylized versions of the discrimination risks present in real fielded systems, but may offer promising ways forward.

Policy design

In the US, some sectoral privacy rules operate by specifically excluding certain personal data from consideration. For example, creditors are generally barred from considering race or gender,[251] and under the US health privacy regime, "personal health information" (generally gathered in a medical setting) is subject to heightened protection.[252] But, in a rapidly changing landscape with an ever-growing supply of digital traces that reflect personal information, it is increasingly possible to use other (unregulated) signals to reach similarly sensitive findings. For example,

many consumer behaviors are closely correlated with race and with gender, and a growing amount of "quantified self" data (such as the data streams from personal fitness devices) allows for health-related inferences. For example, FICO now markets a "medication adherence score," based on publicly available data (rather than on the "personal health information" that the US HIPAA health privacy statute protects).[253]

These trends augur for policy approaches that focus on the nature or consequences of an inference, rather than on the data used to reach it. The new GDPR regime has some language suggesting this approach, and it is also allowed under the controversial and limited US doctrine of "disparate impact."[254]

Efforts to strengthen these policy tools might have a beneficial impact across the full gamut of data-driven decisions.

3. How can the work of data protection stakeholders, including data protection authorities and privacy compliance personnel, be channeled toward concrete gains for rights and justice?

As described above, regulatory authorities in the US and in Europe have struggled to convert their authority into highly visible or consequential victories for data protection.

The growing complexity of data protection rules and fast-changing corporate practices has also sparked a large and very fast-growing community of corporate executives, professional staff and lawyers whose working lives are devoted entirely to privacy and data issues. The "International Association of Privacy Professionals" or IAPP, founded in 2000, is an umbrella group catering to corporate employees who work full time on privacy policy, compliance and risk management. The group now has more than 20,000 members worldwide,[255] and runs several certification programs for people "who serve the data protection, information auditing, information security, legal compliance and/or risk management needs of their organizations."[256] The group's president has also served as the trade association head for an association of online advertising firms.[257]

It is far from clear how much benefit to the public or enhancement to the protection of fundamental rights actually emerges from these substantial corporate investments.

In Europe, there is a broad regulatory framework in place that limits the data broker landscape and puts safeguards in place with respect to the collection, analysis and use of personal data. The broad nature of this framework, both in terms of its overall reach and the variety of interlinked safeguards it contains, is a mixed blessing. On the one hand, social justice actors can point to the

trump card of an acknowledged fundamental right to data protection in litigation and the public debate, and can find a large number of plausible premises on which to pursue litigation—even if the actual scope of protection, and hence the outcome of litigation, is difficult to predict.

At the same time, the enormous variety of potential grounds for litigation and campaigning under European data protection law creates a challenge of priorities—there are many potentially low impact battles into which resources could be directed. Resources are limited everywhere: in civil society, of regulators, legislatures and courts, and of industry compliance professionals. It is vitally important to focus resources on issues that matter for real people in their daily lives—both because such issues matter in their own right, and because they best demonstrate the broader value of the data protection framework.

Social justice actors can easily find themselves playing a game of whack-a-mole, instead of focusing and winning battles on the underlying issues that truly matter most. For example, the social justice community might win a battle to define a robust affirmative consent requirement under the GDPR—but if such efforts succeed, relevant industry players may simply move to another legal ground for their data processing. Or, if the social sector wins strong protections against the further processing of personal data, the industry may migrate its activities toward more statistical data mining on pseudonymous data. Such battles may play out without much change to the practical realities of data-driven decisions in people's lives. A more explicit and sometimes more narrow articulation of what ultimately matters could have significant value and should be an important part of the question how to spend resources.

There is no doubt that the fundamental European right to data protection is broad and establishes some valuable safeguards. At the same time, there is a lack of clarity about what the right really means—what it stands for in practice. It is often interpreted as implying the control over one's personal data. While this appears to have wide appeal and was a pillar of the GDPR proposals, there are reasons to be skeptical about control over data as a long-term ideal. First, the exceptions in the legal framework make room for a daily reality in which there is little meaningful control, in the law as well as in practice. Second, it frames issues in terms of individuals in relation to "their" data, which is a way of looking at things that is less and less valuable as a starting point for protection. As Nissenbaum and Barocas explain:

> When analysts can draw rules from the data of a small cohort of consenting individuals that generalize to an entire population, consent loses its practical import. In fact, the value of a particular individual's withheld consent diminishes the more effectively a company can draw inferences from the set of people that do consent as it approaches a representative sample.[258]

When thinking about the substantive core of data privacy, different, sometimes incompatible, goals are competing: a right to be let alone, which may be worth articulating strongly in cases in which people should really be free of surveillance concerns, and what could be called fairness in data-driven decisionmaking. In addition, data privacy does help to further our society's ability to understand and scrutinize complex technical systems involving personal data. Which of these goals deserves to be furthered in specific contexts is non-trivial question that does deserve attention.

An impact-driven approach, that picks data protection battles based on the impact that particular profiling decisions have on people's lives (and particularly on the lives of people at the margins of society), may ultimately prove most useful to the social sector.

5. Conclusion

Data brokerage and profiling are increasingly important to the decisions that matter most for the rights of individuals, throughout the modern world. This trend will only continue.

We suggest focusing future efforts on situations that have dual reason for concern—situations where the stakes involve not only data protection and privacy, but also some other fundamental right like equality or due process. By focusing on such cases, it is possible both to highlight the strongest reasons for general privacy rules, and at the same time to make progress on other, longstanding social justice goals.

Endnotes

1. United States Senate Committee Commerce, Science, and Transportation, Office of Oversight and Investigations, Majority Staff, *A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes,* December 18, 2013, 1, *available at* http://educationnewyork.com/files/rockefeller_databroker.pdf.

2. United States Government Accountability Office, *Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace*, Report to the Chairman, Committee on Commerce, Science, and Transportation, US Senate, September 2013, 4, *available at* http://www.gao.gov/assets/660/658151.pdf.

3. Organisation for Economic Co-operation and Development, *Exploring the Economics of Personal Data*, 2013, 201, *available at* http://www.oecd-ilibrary.org/science-and-technology/exploring-the-economics-of-personal-data_5k486qtxldmq-en.

4. Federal Trade Commission, *Data Brokers: A Call for Transparency and Accountability,* May 2014, *available at* https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

5. United States Senate Committee Commerce, Science, and Transportation, Office of Oversight and Investigations, Majority Staff, *A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes,* December 18, 2013, *available at* http://educationnewyork.com/files/rockefeller_databroker.pdf.

6. United States Government Accountability Office, *Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace*, Report to the Chairman, Committee on Commerce, Science, and Transportation, US Senate, September 2013, 1, *available at* http://www.gao.gov/assets/660/658151.pdf. *(*GAO defines as a company "with a primary line of business of collecting, aggregating, and selling personal information to third parties.")

7. Office of the Privacy Commissioner of Canada, *Data Brokers: A Look at the Canadian and American Landscape*, September 2014, *available at* https://www.priv.gc.ca/information/research-recherche/2014/db_201409_e.pdf.

8. *See generally,* John Deighton and Peter Johnson, *The Value of Data*: *Consequences for Insight, Innovation & Efficiency in the US Economy*, October 18, 2013, *available at* https://thedma.org/wp-content/uploads/The_Value_of_Data_Consequences_for_Insight_Innovation_and_Efficiency_in_the_US_Economy_WEB1.pdf.

9. *See, e.g.*, D-Cent, Research on Identity Ecosystem, FP7 CAPS 2013, 2015, available at http://dcentproject.eu/wp-content/uploads/2015/08/D3.3-Research-on-Identity-Ecosystem_part1.pdf.

10. European Data Protection Supervisor, *Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy*, March 2014, *available at* https://secure.edps.europa.eu/EDPSWEB/webdav/shared/Documents/Consultation/Opinions/2014/14-03-26_competitition_law_big_data_EN.pdf.

11. Datatilsynet, *The Great Data Race: How commercial utilisation of personal data challenges privacy*, November 2015, *available at* http://www.datatilsynet.no/Global/04_analyser_utredninger/2015/engelsk-kommersialisering-november-2015.pdf.

12. OECD, Exploring the Economics of Personal Data, *available at* http://www.oecd-ilibrary.org/science-and-technology/exploring-the-economics-of-personal-data_5k486qtxldmq-en.

13. Lois Beckett, "Everything We Know About What Data Brokers Know About You," ProPublica, June 13, 2014, *available at* https://www.propublica.org/article/everything-we-know-about-what-data-brokers-know-about-you.

14. *See, e.g.*, ACLU, "Racial Profiling: Definition," *available at* https://www.aclu.org/racial-profiling-definition; *also see* US Department of Homeland Security, "The Department of Homeland Security's Commitment to Nondiscriminatory Law Enforcement and Screening Activities," April 26, 2013, *available at* https://www.dhs.gov/sites/default/files/publications/secretary-memo-race-neutrality-2013_0.pdf.

15. *See generally*, General Data Protection Regulation, Article 4, Section 3aa.

16. United States Senate Committee Commerce, Science, and Transportation, *A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes,* December 18, 2013, 34, footnote 117, *available at* http://educationnewyork.com/files/rockefeller_databroker.pdf.

17. Geschichte der SCHUFA, *available at* https://www.schufa.de/de/ueber-uns/unternehmen/geschichte-schufa/.

18. FICO, "About Us," *available at* http://www.fico.com/en/about-us#at_glance.

19. US Senate, *A Review of the Data Broker Industry*, 1, *available at* http://educationnewyork.com/files/rockefeller_databroker.pdf.

20. The Economist, "Cash in hand," April 2, 2012, *available at* http://www.economist.com/blogs/graphicdetail/2012/04/focus; *see also,* Knowledge@Wharton, "Going Cashless: What's Good for Banks May Not Be Best for You," June 6, 2012, *available at* http://knowledge.wharton.upenn.edu/article/going-cashless-whats-good-for-banks-may-not-be-best-for-you/.

21. Knowledge@Wharton, "Going Cashless: What's Good for Banks May Not Be Best for You," June 6, 2012, *available at* http://knowledge.wharton.upenn.edu/article/going-cashless-whats-good-for-banks-may-not-be-best-for-you/.

22. Katarina Ahlfort, "Cashless future for Sweden? Popularity of Swish hastens digitalisation of country's money," KTH Royal Institute of Technology, October 14, 2015, *available at* https://www.kth.se/en/forskning/artiklar/cashless-future-for-sweden-1.597792.

23. The Economist, "Cash in hand," April 2, 2012, *available at* http://www.economist.com/blogs/graphicdetail/2012/04/focus.

24. *See generally,* Capgemini and Royal Bank of Scotland, *World Payments Report 2015*, 16, *available at* https://www.worldpaymentsreport.com/download.

25. Google, Consumer Barometer, "The Online & Multiscreen World: Types of Devices Used," *available at* https://www.consumerbarometer.com/en/graph-builder/?question=M1&filter=country*:austria,belgium,bulgaria,croatia,czech_republic,denmark,estonia,finland,france,germany,greece,hungary,ireland,italy,latvia,lithuania,netherlands,norway,poland,portugal,romania,russia,serbia,slovakia,slovenia,spain,sweden,switzerland,united_kingdom,ukraine.

26. Google, Consumer Barometer, "Trended Data: USA, All," *available at* https://www.consumerbarometer.com/en/trending/?countryCode=US&category=TRN-NOFILTER-ALL.

27. Greg Sterling, "Report: US Smartphone Penetration Now At 75 Percent," Marketing Land, February 9, 2015, *available at* http://marketingland.com/report-us-smartphone-penetration-now-75-percent-117746.

28. Tom Warren, "Millions of Android users 'deceived' by flashlight app that shares location with advertisers," The Verge, December 6, 2013, *available at* http://www.theverge.com/2013/12/6/5181472/brightest-flashlight-free-ftc-location-data-settlement.

29. Julia Angwin, "The Web's New Gold Mine: Your Secrets," The Wall Street Journal, July 30, 2010, *available at* http://www.wsj.com/articles/SB10001424052748703940904575395073512989404.

30. US Senate, *A Review of the Data Broker Industry*, 29, *available at* http://educationnewyork.com/files/rockefeller_databroker.pdf.

31. Paul Nadler, *Weekly Adviser: Horror at Credit Scoring Is Not Just Foot-Dragging*, 164 Am. Banker, no. 211, November 2, 1999, at 9.

32. LexisNexis, "LexisNexis Risk Solutions Investor Seminar," May 10, 2011, 11, *available at* http://www.relx.com/investorcentre/Documents/presentations/lexis-nexis-risk-solutions-investor-seminar-100511.pdf.

33. Ibid.

34. Marketwatch, "10-Q: IMS HEALTH HOLDINGS, INC.," November 4, 2015, *available at* http://www.marketwatch.com/story/10-q-ims-health-holdings-inc-2015-11-04.

35. Logan Koepke, "Clever helps 44,000 schools share student data with tech firms. What comes next?" Equal Future, September 23, 2015, *available at* https://www.equalfuture.us/2015/09/23/clever-student-data-schools-tech-firms/.

36. *See* LexisNexis, "LexisNexis® Accurint ® for Law Enforcement" (promotional brochure), *available at* http://www.lexisnexis.com/government/solutions/literature/accurintle_ss.pdf ("Currently used by over 4,000 federal, state and local law enforcement agencies across the country, LexisNexis ® Accurint for Law Enforcement is a proven and effective tool.").

37. Ancestry.com, "About Ancestry," *available at* http://corporate.ancestry.com/about-ancestry/.

38. FTC, *Data Brokers: A Call for Transparency and Accountability,* May 2014, iv, *available at* https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

39. US Senate, *A Review of the Data Broker Industry*, 15, *available at* http://educationnewyork.com/files/rockefeller_databroker.pdf. Even though data brokers have transparency obligations about sources in Europe, reports show they resist giving such information with appeals to business secrecy. Goslinga, 2015, https://decorrespondent.nl/3472/Zo-houden-datahandelaren-ons-in-de-gaten-maar-wie-controleert-hen-/180128214112-3b51067e.

40. FTC, *Data Brokers: A Call for Transparency and Accountability*, iv, available at https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

41. Datatilsynet,*The Great Data Race: How commercial utilisation of personal data challenges privacy*, November 2015, *available at* http://www.datatilsynet.no/Global/04_analyser_utredninger/2015/engelsk-kommersialisering-november-2015.pdf.

42. FTC, *Data Brokers: A Call for Transparency and Accountability,* May 2014, 41, *available at* https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

43. Ibid., 12, 17.

44. Lois Beckett, "Everything We Know About What Data Brokers Know About You," ProPublica, June 13, 2014, *available at* https://www.propublica.org/article/everything-we-know-about-what-data-brokers-know-about-you.

45. Ibid.

46. Anna Bernasek, Should Tax Bills Be Public Information?, NYTimes, Feb 13, 2010, *available at* http://www.nytimes.com/2010/02/14/business/yourtaxes/14disclose.html. The republishing of tax records in newspapers has produced an important line of cases about the balance of data protection and freedom of expression. *See* Wouter Hins, Case C-73/07, Tietosuojavaltuutettu v. Satakunnan Markkinapörssi Oy and Satamedia Oy, judgment of the Grand Chamber of 16 December 2008, Common Market Review, 2010. See also Dirk Voorhoof, ECtHR accepts strict application of data protection law and narrow interpretation of journalistic activity in Finland, Strasbourg Observers, 2015, *available at* http://strasbourgobservers.com/2015/08/12/ecthr-accepts-strict-application-of-data-protection-law-and-narrow-interpretation-of-journalistic-activity-in-finland/.

47. FTC, *Data Brokers: A Call for Transparency and Accountability*, 16, *available at* https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

48. US Senate, *A Review of the Data Broker Industry*, 16, *available at*: http://educationnewyork.com/files/rockefeller_databroker.pdf.

49. Lois Beckett, "Everything We Know About What Data Brokers Know About You," ProPublica, June 13, 2014, *available at* https://www.propublica.org/article/everything-we-know-about-what-data-brokers-know-about-you.

50. Lois Beckett, "Everything We Know About What Data Brokers Know About You," ProPublica, June 13, 2014, *available at* https://www.propublica.org/article/everything-we-know-about-what-data-brokers-know-about-you. (Equifax says this is only sold to customers who been been verified through a detailed credentialing process.)

51. US Senate, *A Review of the Data Broker Industry*, 29, *available at*: http://educationnewyork.com/files/rockefeller_databroker.pdf.

52. FTC, *Data Brokers: A Call for Transparency and Accountability,* 14, *available at* https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

53. *See generally,* The Wall Street Journal, "What They Know," *available at* http://www.wsj.com/public/page/what-they-know-digital-privacy.html.

54. FTC, *Data Brokers: A Call for Transparency and Accountability*, 20, *available at* https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

55. Maaike Golsinga, 'Zo houden datahandelaren ons in de gaten (maar wie controleert hen?)' [This is how data brokers are tracking us. (But who controls them?'], De Correspondent, available at https://decorrespondent.nl/3472/Zo-houden-datahandelaren-ons-in-de-gaten-maar-wie-controleert-hen-/377813634352-70e0c6f6.

56. Lois Beckett, "Everything We Know About What Data Brokers Know About You," ProPublica, June 13, 2014, *available at* https://www.propublica.org/article/everything-we-know-about-what-data-brokers-know-about-you.

57 Lexis Nexis, "Stress Index Executive Summary," *available at* http://www.lexisnexis.com/risk/downloads/assets/stress-index-executive-summary.pdf.

58. United States Government Accountability Office, *Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace*, September 2013, 5, *available at* http://www.gao.gov/assets/660/658151.pdf.

59. *See, e.g.,* Privacy Rights Clearinghouse, "Online Data Vendors: How Consumers Can Opt Out of Directory Services and Other Information Brokers," March 2013, *available at* https://www.privacyrights.org/online-information-brokers-list.

60. See, e.g., United States Government Accountability Office, Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace, September 2013, 5, *available at* http://www.gao.gov/assets/660/658151.pdf. (Mentioning DMA's 2,500 figure.)

61. Steve Kroft, "The Data Brokers Selling Your Personal Information," CBS News: 60 Minuts, March 9, 2014, *available at* http://www.cbsnews.com/news/the-data-brokers-selling-your-personal-information/.

62. Senator John D. Rockefeller IV, "What Information Do Data Brokers Have on Consumers, and How Do They Use It?" December 18, 2013, *available at* https://www.commerce.senate.gov/public/index.cfm/hearings?Id=a5c3a62c-68a6-4735-9d18-916bdbbadf01&Statement_id=A47C081A-D653-4272-8D12-D6EDC1E04DC6.

63. Lois Beckett, "Everything We Know About What Data Brokers Know About You," ProPublica, June 13, 2014, *available at* https://www.propublica.org/article/everything-we-know-about-what-data-brokers-know-about-you.

64. US Senate, *A Review* of *the Data Broker Industry*, 14, footnote 67, *available at* http://educationnewyork.com/files/rockefeller_databroker.pdf.

65. Acxiom Corporation, *Annual Report 2014*, 6, *available at* http://files.shareholder.com/downloads/ACXM/0x0x763250/A1DBFBD8-E136-4701-B0F2-3DC695E5ED08/acxiom2014_Annual_Report_FINAL_RRD_PDF_.pdf.

66. Alliance Data, *2014 Annual Report*, 41, *available at* http://phx.corporate-ir.net/External.File?item=UGFyZW50SUQ9MjgwOTIyfENoaWxkSUQ9LTF8VHlwZT0z&t=1.

67. LexisNexis, "LexisNexis Risk Solutions Investor Seminar," May 10, 2011, 11, *available at* http://www.relx.com/investorcentre/Documents/presentations/lexis-nexis-risk-solutions-investor-seminar-100511.pdf.

68. Experian, Annual Report, 2015, 30, *available at* https://www.experianplc.com/media/2607/experian-ar15.pdf.

69. Equifax, Annual Report, 2014, 16, *available at* http://www.equifax.com/pdfs/corp/Equifax_2014_Annual_Report.pdf.

70. TransUnion, Form 10-K filed with the US Securities Exchange Commission, 45, *available at* http://s2.q4cdn.com/601093094/files/10-K-as-filed-(2).pdf.

71. Acxiom Corporation, Annual Report, 2014, F-5, *available at* http://files.shareholder.com/downloads/ACXM/0x0x763250/A1DBFBD8-E136-4701-B0F2-3DC695E5ED08/acxiom2014_Annual_Report_FINAL_RRD_PDF_.pdf. ("US MDS revenue increased $17.6 million, or 2.6%. → 17,600,000 / .026 = $676,923,076.92.)

72. Alliance Data, 10-K filed with the Securities Exchange Commission, 41, *available at* http://phx.corporate-ir.net/External.File?item=UGFyZW50SUQ9MjgwOTIyfENoaWxkSUQ9LTF8VHlwZT0z&t=1.

73. Trefis Team, "Here's Why Oracle Paid Over $1.2 Billion for Acquiring Datalogix," Forbes, February 20, 2015, *available at* http://www.forbes.com/sites/greatspeculations/2015/02/20/heres-why-oracle-paid-over-1-2-billion-for-acquiring-datalogix/.

74. A German overview and discussion can be found in Christl, 2014. Dominant firms are all German: BIS GmbH, Acxiom Deutschland GmbH, AZ Direkt GmbH, Deutsche Post Direkt GmbH, EOS Holding GmbH und der Schober Information Group Deutschland GmbH.

75. Daniel Farey-Jones, "Acxiom accepts $3bn takeover offer from private equity," Campaign, May 17, 2007, *available at* http://www.campaignlive.co.uk/article/658319/acxiom-accepts-3bn-takeover-offer-private-equity.

76. LexisNexis, "Lexis Diligence," *available at* http://bis.lexisnexis.co.uk/products/lexis-diligence.

77. Lee Andrew Bygrave, *Data Privacy Law: An International Perspective*, (Oxford University Press, 2014), 108.

78. These principles were drawn up in the early 1970s by expert committees working contemporaneously yet independently of each other on both sides of the Atlantic. The British Parliament was investigating potential privacy issues with computerized personal data records and the US Department of Health, Education and Welfare developed "remarkable similar" principles. Bygrave, Aims, 108. *See also* Younger Committee, *Report of the Committee on privacy*, Cmnd 5012 (HSMO 1972); HEW, *Records, Computers and the Rights of Citizens*, 41. Today, the FIPPs are enshrined in privacy laws across the world, including the United States's Privacy Act of 1974 (which covers government agencies) and the European Union's Data Protection Directive and the new General Data Protection Regulation.

79. Some have argued that privacy practitioners should dedicate more attention to preventing and curtailing harmful *uses* of data, and less to the *collection* of data itself. For example, Craig Mundie has argued that "'[b]ig data' has rendered obsolete the current approach to protecting privacy and civil liberties," which he sees as fixated on "controlling the collection and retention of personal data." Craig Mundie, "Privacy Pragmatism: Focus on Data Use, Not Data Collection," Foreign Affairs, (March/April 2014), *available at* http://www.foreignaffairs.com/articles/140741/craig-mundie/privacy-pragmatism (emphasis added).

80. Bygrave, *Data Privacy Law*, 107.

81. *See* US Const. Amend. IV.

82. Bygrave, *Data Privacy Law*, 110.

83. United States Government Accountability Office, *Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace*, September 2013, *available at* http://www.gao.gov/assets/660/658151.pdf. "There are no current laws requiring data brokers to maintain the privacy of consumer data unless they use that data for credit, employment, insurance, housing, or similar purposes," summarizes the Federal Trade Commission (FTC). Federal Trade Commission, "FTC to Study Data Broker Industry's Collection and Use of Consumer Data," December 18, 2012, *available at* https://www.ftc.gov/news-events/press-releases/2012/12/ftc-study-data-broker-industrys-collection-use-consumer-data.

84. S*ee generally*, Gina Stevens, *Privacy Protections for Personal Information Online*, Congressional Research Service, April 6, 2011, *available at* https://www.fas.org/sgp/crs/misc/R41756.pdf.

85. 15 USC 1681-81t. The Fair Credit Reporting Act of 1970 (FCRA) as amended by the Fair and Accurate Credit Transactions Act of 2003, 15 USC. 1681-81t.

86. 20 USC. 1232g. The Family Educational Rights and Privacy Act of 1974 (FERPA).

87. 12 USC. 3401. The Right to Financial Privacy Act of 1978.

88. 18 USC. 2710. The Video Privacy Protection Act of 1988.

89. 42 USC. 1320d note. The Health Insurance Portability and Accountability Act of 1996 (HIPAA).

90. 47 USC. 222. Communications Act of 1934, as amended.

91. National Conference of State Legislatures, "State Laws Related to Internet Privacy," February 24, 2015, *available at* http://www.ncsl.org/research/telecommunications-and-information-technology/state-laws-related-to-internet-privacy.aspx.

92. Comments of the Center for Democracy and Technology, In the Matter of: Information Privacy and Innovation in the Internet Economy, United States Department of Commerce, National Telecommunications and Information Administration, available at https://www.cdt.org/files/pdfs/20100613_doc_privacy_noi.pdf.

93. 15 USC. § 1681(b) (2012).

94. 15 USC. § 1691 (2012).

95. *See, e.g.,* 12 C.F.R. § 1002.15(b)(1)(i) (2014).

96. 42 USC. 3601-3619

97. FTC Act, Section 5.

98. *See, e.g.* FTC, *Data Brokers: A Call for Transparency and Accountability*, *available at* https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf. (For which it used 15 USC § 46.)

99. See, *e.g.,* In the Matter of LabMD, Inc., *available at* http://www.databreaches.net/wp-content/uploads/Ruling.pdf.

100. Federal Trade Commission, *Protecting Consumer Privacy in an Era of Rapid Change*, March 2012, *available at* https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf; *also see* FTC, *Data Brokers: A Call for Transparency and Accountability, available at* https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

101. Stevens, *Privacy Protections for Personal Information Online*, Congressional Research Service, *available at* https://www.fas.org/sgp/crs/misc/R41756.pdf.

102. Logan Koepke,"White House's Draft Consumer Privacy Bill Draws Criticsim from Companies, Privacy Advocates Alike," Equal Future, March 4, 2015, *available at* https://www.equalfuture.us/2015/03/04/white-houses-draft-consumer-privacy-bill-draws-criticsim-from-companies-privacy-advocates-alike/. (Concluding that "The White House's effort underscores the fundamental challenge of broadly regulating the powerful and diverse array of companies handling personal data" in the US.).

103. For example, in 2015, Google eclipsed the telecommunications behemoths in term of spending on lobbying. *See, e.g.,* T.C. Sottek "It's official: Washington D.C. is now Google's town," The Verge, January 21, 2015, *available at* http://www.theverge.com/2015/1/21/7867975/mr-google-goes-to-washington.

104. Paul Schwarz, *The EU-US Privacy Collision: A Turn to institutions and Procedures*, 126 Harvard Law Review 1966 (2013), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2290261.

105. GAO, *Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the* Marketplace, 4, *available at* http://www.gao.gov/assets/660/658151.pdf.

106. US Senate, *A Review of the Data Broker Industry*, 10, *available at*: http://educationnewyork.com/files/rockefeller_databroker.pdf.

107. FTC, *Data Brokers: A Call for Transparency and Accountability*, i, *available at* https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

108. European Parliament & Council Directive 95/46/EC, Data Protection Directive, http://ec.europa.eu/justice/data-protection/.

109. *See generally*, GPDR.

110. Charter on Fundamental rights of the European Union, December 18, 2000, OJ 2000/C 364/01 (EU Charter). Council of Europe, Convention for the Protection of Human Rights and Fundamental Freedoms, November 4, 1950, Europ. T.S. No. 5. (ECHR).

111. The entity "which alone or jointly with others determines the purposes and means of the processing of personal data." GDPR, Article 4(1).

112. Where an "identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person." GDPR, Article 4 (1).

113. For transfers of personal data outside of the EU, a set of rules apply that aim to ensure the protection afforded is not circumvented.

114. Article 3(2), GDPR.

115. Article 8 DPD, Article 9, 9a, GDPR.

116. DPD Chapter IV and V, GDPR, Chapter III.

117. Article 15 DPD, Article 20 GDPR.

118. Article 29 WP, 136 (2007).

119. Article 29 WP, 217 (2014).

120. Article 29 WP, 171 (2010).

121. Article 29 WP, 223, (2014).

122. *See* Fundamental Rights Agency, Data Protection Handbook (2014).

123. Council of Europe, Convention 108.

124. EU Charter of Fundamental Rights.

125. Gloria González Fuster, Beyond the GDPR, above the GDPR, Internet Policy Review, 30 November 2015, *available at* http://policyreview.info/articles/news/beyond-gdpr-above-gdpr/385.

126. CJEU 13 May 2014, C-131/12 (*Google Spain*).

127. CJEU 8 April 2014, Joined cases C-293/12 and C-594/12 (*Digital Rights Ireland*).

128. CJEU 6 October 2015, C-362/14 (*Schrems*).

129. Council Directive 2002/58/EC, as amended by Council Directive 2006/24/EC and Council Directive 2009/136/EC.

130. Article 5 (3) of the e-Privacy Directive.

131. *See e.g.*, the Directive on Consumer Rights (2011/83/EC) or the Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts.

132. Council Consumer Credit Directive (2008/48/EC) of 23 April 2008.

133. European Union Agency for Fundamental Rights, European Court of Human Rights, and Council of Europe, *Handbook on European Non-Discrimination Law*, July 2010, *available at* http://fra.europa.eu/sites/default/files/fra_uploads/1510-FRA-CASE-LAW-HANDBOOK_EN.pdf.

134. Data Protection Directive for Law Enforcement 2016/XX, consolidated text available at http://www.emeeting.europarl.europa.eu/committees/agenda/201512/LIBE/LIBE%282015%291217_1/sitt-1739884.

135. Boehm, A Comparison Between US and EU Data Protection Legislation for Law Enforcement Purposes, Study for the European Parliament, 2015, *available at* http://www.europarl.europa.eu/thinktank/en/document.html?reference=IPOL_STU%282015%29536459.

136. EDPS, Opinion, 6/2015, *available at* http://statewatch.org/news/2015/oct/eu-dir-dp-leas-edps-recommendations.pdf.

137. Boehm, 40.

138. EDRi, Everything you need to know about the Data Protection Directive for Law Enforcement, *available at* http://policingprivacy.eu/topics/. *See also* Anna Fielder, "New EU data protection laws: ok, but a tremendous missed opportunity with possible threats looming", 17 December 2015, https://www.privacyinternational.org/node/689.

139. European Commission, 'Questions and Answers—Data Protection Reform', Brussels, 21 December 2015, *available at* http://europa.eu/rapid/press-release_MEMO-15-6385_en.htm.

140. *See* Datatilsynet, 2015.

141. CJEU, Google Spain.

142. CJEU, Referral of BGH Germany, 17 December 2014, case C-582/14.

143. Frederik J. Zuiderveen Borgesius, 'Personal data processing for behavioural targeting: which legal basis?', International Data Privacy Law, Vol. 5, No. 3, 2015.

144. GDPR, Recital 38 and 39.

145. GDPR, Article 7.

146. Ruth Boardman et al., 'Agreement on general data protection regulation', Bird & Bird, 2015.

147. DPD, Article 8, GDPR, Article 9. In 2009, the Dutch DPA enforced these strict rules against the Dutch data broker Advance BV, https://cbpweb.nl/sites/default/files/downloads/pb/pb_20091218_advance_bevindingen.pdf.

148. GDPR, Article 19(2).

149. DPD, Article 15; GDPR, Article 20.

150. GDPR. Article 20(1b). For a discussion of the hardly ever invoked Article 15 DPD, see Bygrave, L., 'Minding the machine: Article 15 of the EC Data Protection Directive and Automated Profiling', Computer Law & Security Rep., 2001, pp. 17–24, *available at* http://www.austlii.edu.au/au/journals/PLPR/2000/40.html. See also Koops, B.J. (2013), 'On Decision Transparency, or How to Enhance Data Protection after the Computational Turn', in: M. Hildebrandt & K. De Vries (eds), Privacy, Due Process and the Computational Turn, Abingdon: Routledge, pp. 196–220, 2013.

151. Bruce Schneier, *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*, (W.W. Norton & Company, 2015), 49.

152. Ibid.

153. Nurie Mohamed, "You Deleted Your Cookies? Think Again" Wired, August 10, 2009, *available at* http://www.wired.com/2009/08/you-deleted-your-cookies-think-again/.

154. Center for Democracy and Technology, Consumer Action, and Privacy Activism, "In regards to the FTC Staff Statement, 'Online Behavioral Advertising: Moving the Discussion Forward to Possible Self-Regulatory Principles,'" April 11, 2008, 16, *available at* https://www.cdt.org/files/privacy/20080411bt_comments.pdf.

155. Center for Democracy and Technology, "Comments for November 2015 Workshop on Cross-Device Tracking," October 16, 2015, 5, *available at* https://cdt.org/files/2015/10/10.16.15-CDT-Cross-Device-Comments.pdf.

156. Julia Angwin, "Why Online Tracking Is Getting Creepier," ProPublica, 2014, *available at* http://www.propublica.org/article/why-online-tracking-is-getting-creepier.

157. S*ee, e.g.*, TruSignal, Specialized Audience Buyer's Guide, 2015, *available at* http://www.tru-signal.com/wp-content/uploads/2014/11/TruSignal_interactive_audience_guide.040615.pdf.

158. Facebook, "Partner Categories, a New Self-Serve Targeting Feature," April 10, 2013, *available at* https://www.facebook-studio.com/news/item/partner-categories-a-new-self-serve-targeting-feature; *see also* Kyle Boston, "Introducing partner audiences," Twitter, March 5, 2015, *available at* https://blog.twitter.com/2015/introducing-partner-audiences. On the measurement of advertisement performance, see Facebook, Relevant Ads that Protect Your Privacy, 2012, *available at* https://www.facebook.com/notes/facebook-and-privacy/relevant-ads-that-protect-your-privacy/457827624267125.

159. Facebook Help Center, "Ad Set Audiences," *available at* https://www.facebook.com/help/433385333434831/; *see also* Kyle Boston, "Introducing partner audiences," Twitter, available at https://blog.twitter.com/2015/introducing-partner-audiences.

160. Facebook Help Center, "Lookalike Audiences," *available at* https://www.facebook.com/help/231114077092092/; *see also* Kyle Boston, "Introducing partner audiences," Twitter, *available at* https://blog.twitter.com/2015/introducing-partner-audiences.

161. FTC, *Data Brokers: A Call for Transparency and Accountability, available at* https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

162. Ibid.

163. Ibid.

164. *See generally,* Aaron Rieke, "How to Think About Differential Pricing and Big Data," Equal Future, February 11, 2015, *available at* https://www.equalfuture.us/2015/02/11/differential-pricing-and-big-data/.

165. Jennifer Valentino-Devries, Jeremy Singer-Vine, and Ashkan Soltani, "Websites Vary Prices, Deals Based on Users' Information," The Wall Street Journal, December 24, 2012, *available at*http://www.wsj.com/articles/SB10001424127887323777204578189391813881534.

166. Dana Mattioli, "On Orbitz, Mac Users Steered to Pricier Hotels," The WAll Street Journal, August 23, 2012, *available at* http://www.wsj.com/articles/SB10001424052702304458604577488822667325882.

167. The White House, Executive Office of the President, *Big Data: Seizing Opportunities, Preserving Values,* May 2014, *available at*https://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf.

168. Latanya Sweeney, *Discrimination in Online Ad Delivery*, 11 Queue 10 (2013), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2208240.

169. Joseph Turow, Michael Hennesy, and Nora Draper, *The Tradeoff Fallacy: How Marketers Are Misrepresenting American Consumers and Opening Them up to Exploitation*, June 2015, *available at* https://www.asc.upenn.edu/sites/default/files/TradeoffFallacy_1.pdf. ("A broad public fear about what companies can do with the data, portends serious difficulties not just for individuals but also—over time—for the institution of consumer commerce.").

170. *See* Upturn, "Knowing the Score" at 6.

171. United States Government Accountability Office, *Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace*, September 2013, i, *available at* http://www.gao.gov/assets/660/658151.pdf.

172. United States Senate, *A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes,* December 18, 2013, 36, *available at* http://educationnewyork.com/files/rockefeller_databroker.pdf.

173. FTC, *Data Brokers: A Call for Transparency and Accountability,* May 2014, viii, *available at* https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

174. For example, the FTC has recently held workshops on "Cross-Device Tracking" and "Alternative Scoring." *See* https://www.ftc.gov/news-events/events-calendar/2015/11/cross-device-tracking; https://www.ftc.gov/news-events/events-calendar/2014/03/spring-privacy-series-alternative-scoring-products

175. Federico Ferretti, "The Legal Framework of Consumer Credit Bureaus and Credit Scoring in the European Union: Pitfalls and Challenges—Overindebtedness, Responsible Lending, Market Integration, and Fundamental Rights," Suffolk University Law Review, Vol. XLVI:791, 794, *available at* http://suffolklawreview.org/wp-content/uploads/2014/01/Ferretti_Lead.pdf.

176. Ashoka, "Banking The Unbanked: A How-To," Forbes, June 14, 2013, *available at* http://www.forbes.com/sites/ashoka/2013/06/14/banking-the-unbanked-a-how-to.

177. For an in-depth discussion of these structures, *See generally* Ferretti.

178. Marc Rothmund and Maria Gerhardt, *The European Credit Information Landscape: An analysis of a survey of credit bureaus in Europe*, European Credit Research Institute, January 2011, 3, *available at* http://aei.pitt.edu/33375/1/ACCIS-Survey_FinalReport_withCover.pdf

179. Ferretti, "The Legal Framework of Consumer Credit Bureaus and Credit Scoring in the European Union," 796.

180. S*ee generally,* Upturn, *Knowing the Score: New Data, Underwriting, and Marketing in the Consumer Credit Marketplace,* October 2014, History Appendix, *available at* https://www.teamupturn.com/static/files/Knowing_the_Score_Oct_2014_v1_1.pdf; *see also* Sean Trainor, "The Long, Twisted History of Your Credit Score," Time, July 22, 2015, *available at* http://time.com/3961676/history-credit-scores/.

181. I*d.*

182. S*ee generally,* Rothmund and Maria Gerhardt, *The European Credit Information Landscape, available at* http://aei.pitt.edu/33375/1/ACCIS-Survey_FinalReport_withCover.pdf. Credit reporting coverages in Europe tends to be high, especially in Western European countries.

183. Paul Nadler, *Weekly Adviser: Horror at Credit Scoring Is Not Just Foot-Dragging*, 164 Am. Banker, no. 211, Nov. 2, 1999, at 9.

184. Robert B. Avery, Paul S. Calem, and Glenn B. Canner, *An Overview of Consumer Data and Credit Reporting*, Federal Reserve Bulletin, February 2003, *available at* http://www.federalreserve.gov/pubs/bulletin/2003/0203lead.pdf.

185. See *generally,* FICO, "About Us," *available at* http://www.fico.com/en/about-us. FICO works hand-in-hand with credit bureaus: by one estimate, more than 90 percent of the scores sold to firms for credit-related decisions in 2010 were scores created by FICO's algorithms. Consumer Financial Protection Bureau, *The Impact of Differences Between Consumer- and Creditor-purchased Credit Scores: Report to Congress*, July 19, 2011, 6, *available at* http://files.consumerfinance.gov/f/2011/07/Report_20110719_CreditScores.pdf

186. Ferretti, "The Legal Framework of Consumer Credit Bureaus and Credit Scoring in the European Union," 798.

187. Ibid., 798

188. Ibid., 798

189. Ibid., 798-99.

190. Rothmund and Maria Gerhardt, *The European Credit Information Landscape, available at* http://aei.pitt.edu/33375/1/ACCIS-Survey_FinalReport_withCover.pdf.

191. See *generally,* Upturn, *Knowing the Score, available at* https://www.teamupturn.com/static/files/Knowing_the_Score_Oct_2014_v1_1.pdf.

192. Sarah Davies, "Big Data and Credit Scoring, 2015 NCRC Annual Conference, March 2015 (on file with author).

193. Id.

194. See, *e.g.*, FICO, "What's in your score," *available at* http://www.myfico.com/crediteducation/whatsinyourscore.aspx.

195. Consumer Financial Protection Bureau, The Impact of Differences Between Consumer- and Creditor-Purchased Credit Scores, July 19, 2011, 1, *available at* http://www.consumerfinance.gov/reports/the-impact-of-differences-between-consumer-and-creditor-purchased-credit-scores.

196. Report of the Expert Group on Credit Histories, May 2009, 2 *available at* http://ec.europa.eu/internal_market/consultations/docs/2009/credit_histories/egch_report_en.pdf.

197. General Principles for Credit Reporting, 1.

198. General Principles of Credit Reporting, 1.

199. Consumer Financial Protection Bureau, Office of Resarch, *Data Point: Credit Invisibles,* May 2015, *available at* http://files.consumerfinance.gov/f/201505_cfpb_data-point-credit-invisibles.pdf.

200. Board of Governors of the Federal Reserve System, *Report to the Congress on Credit Scoring and Its Effects on* the *Availability and Affordability of Credit*, August 2007, S-3, *available at* http://www.federalreserve.gov/boarddocs/rptcongress/creditscore/creditscore.pdf.

201. Ibid.

202. Ibid. In fact, the use of credit scores might likely increases the consistency and objectivity of credit evaluation and thus may help diminish the possibility that credit decisions will be influenced by personal characteristics . . . including race or ethnicity."

203. Chi Chi Wu, National Consumer Law Center, *Solving the Credit Connundrum: Helping Consumers' Credit Records Impaired by the Foreclosure Crisis and Great Recession*, December 2013, 9, *available at* https://www.nclc.org/images/pdf/credit_reports/report-credit-conundrum-2013.pdf.

204. CFPB, *Credit Invisibles*, 4.

205. Ibid., 6.

206. Almost 30 percent of low-income neighborhoods are credit invisible, and an additional 15 percent have unscored records, and Black and Hispanics are more likely than Whites or Asians to be credit invisible or to have unscored credit records.

207. Ferretti, "The Legal Framework of Consumer Credit Bureaus and Credit Scoring in the European Union," 799 (see also sources in n. 24).

208. Federal Trade Commission, "FTC Issues Follow-Up Study on Credit Report Accuracy," January 21, 2015, *available at* https://www.ftc.gov/news-events/press-releases/2015/01/ftc-issues-follow-study-credit-report-accuracy.

209. Id.

210. The Economist, "Stat oil: Lenders are turning to social media to assess borrowers," February 9, 2013, *available at* http://www.economist.com/news/finance-and-economics/21571468-lenders-are-turning-social-media-assess-borrowers-stat-oil.

211. Robinson Meyer, "Could a Bank Deny Your Loan Based on Your Facebook Friends?" The Atlantic, September 25, 2015, *available at* http://www.theatlantic.com/technology/archive/2015/09/facebooks-new-patent-and-digital-redlining/407287/.

212. National Public Radio, All Tech Considered, "Could Your Social Media Footprint Step On Your Credit History?" November 13, 2015, *available at* http://www.npr.org/sections/thetwo-way/2015/11/04/454237651/could-your-social-media-footprint-step-on-your-credit-history.

213. Ferretti,"The Legal Framework of Consumer Credit Bureaus and Credit Scoring in the European Union," 802.

214. TK EU report.

215. Ferretti, "The Legal Framework of Consumer Credit Bureaus and Credit Scoring in the European Union," 814.

216. Sharon Chamard, *The History of Crime Mapping and Its Use by American Police Departments*, University of Alaska Anchorage, Fall 2006, *available at*http://justice.uaa.alaska.edu/forum/23/3fall2006/a_crimemapping.html.

217. Thales Group, "Smarter Data for Public Safety," *available at* https://www.thalesgroup.com/en/worldwide/big-data/solutions/public-safety.

218. LexisNexis, "Accurint LE Plus," *available at* http://www.lexisnexis.com/risk/products/government/accurint-le-plus.aspx.

219. LexisNexis, "Case Studies: LexisNexis Accurint for Law Enforcement," *available at* http://www.lexisnexis.com/government/solutions/casestudy/accurintle.pdf.

220. TransUnion, "TLOxp," *available at* https://www.tlo.com/law-enforcement/.

221. Ann Woolner, "Ex-Drug Smuggler Turned Data Miner Reclaims Field He Created," September 15, 2011, *available at* http://www.bloomberg.com/news/articles/2011-09-15/ex-cocaine-smuggler-turned-data-miner-

seeks-to-conquer-a-field-he-created. ("Anything you've posted without privacy limits on a social media site like Facebook might wind up there, said TLO's chief privacy officer, Martha Barnett.").

222. LexisNexis, "LexisNexis Risk Solutions Investor Seminar," May 10, 2011, 3, 75, *available at* http://www.relx.com/investorcentre/Documents/presentations/lexis-nexis-risk-solutions-investor-seminar-100511.pdf.

223. LexisNexis, "LexisNexis Accurint for Law Enforcement," *available at* http://www.lexisnexis.com/government/solutions/literature/accurintle_ss.pdf.

224. https://www.ftc.gov/system/files/documents/reports/section-319-fair-accurate-credit-transactions-act-2003-sixth-interim-final-report-federal-trade/150121factareport.pdf.

225. Bluejay, *available at* http://brightplanet.com/bluejay/.

226. Logan Koepke, "The feds are tracking #BlackLivesMatter. But does that violate anyone's rights?" Equal Future, July 29, 2015, *available at* https://www.equalfuture.us/2015/07/29/feds-tracking-blacklivesmatter/.

227. S*ee, e.g.*, Kevin Bankston and Ashkan Soltani, *Tiny Constables and the Cost of Surveillance: Making* Cents Out of United States v. Jones, The Yale Law Journal Online, 2014, *available at* http://www.yalelawjournal.org/forum/tiny-constables-and-the-cost-of-surveillance-making-cents-out-of-united-states-v-jones.

228. *See* Ben Popper, "How the NYPD is using social media to put Harlem teens behind bars," The Verge, December 10, 2015, *available at* http://www.theverge.com/2014/12/10/7341077/nypd-harlem-crews-social-media-rikers-prison.

229. TransUnion, "TLOxp," *available at* https://www.tlo.com/vehicle-sightings.php.

230. Ibid.

231. Adam Tanner, "Data Brokers Are Now Selling Your Car's Location For $10 Online," Forbes, July 10, 2013, *available at* http://www.forbes.com/sites/adamtanner/2013/07/10/data-broker-offers-new-service-showing-where-they-have-spotted-your-car/.

232. Shawn Musgrave, "A vast hidden surveillance network runs across America, powered by the repo industry," The Boston Globe, March 5, 2014, *available at* http://www.betaboston.com/news/2014/03/05/a-vast-hidden-surveillance-network-runs-across-america-powered-by-the-repo-industry/.

233. See, e.g., Smith v Maryland, *available at* http://caselaw.findlaw.com/us-supreme-court/442/735.html; US v. Miller, *available at* https://www.law.cornell.edu/supremecourt/text/307/174.

234. James X. Dempsey & Lara M. Flint, Commercial Data and National Security, 74 Geo. Wash. L. Rev. 1474 (Aug. 2004), *available at* https://www.cdt.org/files/publications/200408dempseyflint.pdf.

235. European Parliament, The US Legal System on Data Protection in the Field of Law Enforcement. Safeguard, Rights and Remedies for EU citizens (2015), at 10.

236. Dempsey & Flint, *supra* note 236.

237. *Id*.

238. EU Law Enforcement Directive, 2016/XX, consolidated text available at http://www.emeeting.europarl.europa.eu/committees/agenda/201512/LIBE/LIBE%282015%291217_1/sitt-1739884.

239. Koops, Police-Investigations in Internet Open Sources, 2013.

240. Maurits Martijn, "Politie en inlichtingendiensten kunnen via een achterdeur bij gegevens van de Belastingdienst," De Correspondent, September 30, 2014, *available at* https://decorrespondent.nl/2720/Baas-Belastingdienst-over-Big-Data-Mijn-missie-is-gedragsverandering/83656320-f6e78aaf.

241. *See* Ari Gesher, "Palantir: Like an Operating System for Data Analysis," The Palantir Blog (Nov. 6, 2009), *available at* https://www.palantir.com/2009/11/palantir-like-an-operating-system-for-data-analysis/.

242. Thales Group, "Smarter Data for Public Safety," *available at* https://www.thalesgroup.com/en/worldwide/big-data/solutions/public-safety.

243. Jonathan Zittrain, "Facebook Could Decide an Election Without Anyone Ever Finding Out," The New Republic, June 1, 2014, *available at* https://newrepublic.com/article/117878/information-fiduciary-solution-facebook-digital-gerrymandering.

244. *See generally* Fairness, Accountability, and Transparency in Machine Learning (FAT ML), *available at* http://www.fatml.org/.

245. *See generally* IEEE ICDM 201 International Workshop on Discrimination and Privacy-Aware Data Mining (DPADM), *available at* https://sites.google.com/site/dpadm2012/.

246. Amit Datta, Anupam Datta, Michael Carl Tschantz, and Jeannette M. Wing, "Information Flow Experiments: Determining Information Usage from the Outside," *available at* https://www.cs.cmu.edu/~mtschant/ife/.

247. Roxana Geambasu, "Research—Current Projects," *available at* https://roxanageambasu.github.io/01-research/.

248. National Science Foundation, "Standard Grant—Understanding and Illuminating Non-Public Data Flows," August 17, 2015, *available at* http://www.nsf.gov/awardsearch/showAward?AWD_ID=1514509.

249. Discrimination and Privacy in the Information Society, 3 (Bart Custers et al. eds., 2013), http://link.springer.com/10.1007/978-3-642-30487-3 (last visited Sep 30, 2014).

250. Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, Richard Zemel, *Fairness Through Awareness*, November 30, 2011, *available at* http://arxiv.org/pdf/1104.3913.pdf.

251. 15 USC § 1691 et seq.

252. *See generally*, The Health Insurance Portability and Accountability Act, 110 Stat. 1936.

253. Tara Parker-Pope, "Keeping Score on How Yout Take Your Medicine," The New York Times, June 20, 2011, *available at* http://well.blogs.nytimes.com/2011/06/20/keeping-score-on-how-you-take-your-medicine/.

254. *See generally*, Solon Barocas and Andrew Selbst, *Big Data's Disparate Impact*, California Law Review, Vol. 104, 2016, *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899.

255. International Association of Privacy Professionals, "IAPP Facts," *available at* https://iapp.org/about/iapp-facts.

256. International Association of Privacy Professionals, "IAPP Mission and Background," *available at* https://iapp.org/about/mission-and-background.

257. Andrew Clearwater and J. Trevor Hughes, *In the Beginning... An Early History of the Privacy Profession*, 74 Ohio St. L.J. Vo. 74:6, 916–917, *available at* http://moritzlaw.osu.edu/students/groups/oslj/files/2013/12/8-Clearwater-Hughes.pdf.

258. Solon Barocas and Helen Nissenbaum, "Big Data's End Run Around Procedural Privacy Protections," Communications of the ACM, Vol. 57 No. 11, pp. 31–33, *available at* http://cacm.acm.org/magazines/2014/11/179832-big-datas-end-run-around-procedural-privacy-protections/fulltext.


Upturn


OPEN SOCIETY
FOUNDATIONS