From a Sea of Data to Actionable Insights:

Big Data and What It Means for Lawyers

This article was originally published in 26 *Intellectual Property & Technology Law Journal* No. 3, March 2014, at 8.

by Michael Murphy and John Barton

Michael Murphy
Global Sourcing
+1.415.983.1303
michael.murphy@pillsburylaw.com

John Barton
Global Sourcing
+1.202.663.8703
john.barton@pillsburylaw.com

Michael Murphy and John Barton are partners in the Global Sourcing Practice Group at Pillsbury Winthrop Shaw Pittman LLP. They can be reached at michael.murphy@pillsburylaw.com and john.barton@pillsburylaw.com, respectively. The authors wish to thank their colleagues who contributed to this article in the following subject areas: Paula Weber and Keith Hudolin (Employment Law); Joseph Lynyak (Lending Law); and Catherine Meyer (Privacy Law).

Big Data is one of the most hyped, and most confusing, terms in technology jargon. We know this to be true because Big Data tells us so. According to The Global Language Monitor,[1] "Big Data" has edged out such classics as "The Cloud" and "The Next Big Thing" as the most confusing tech buzzword of the decade. But even after we peel away the hype, it is clear that Big Data is changing our society—how we research, analyze, plan, think, make policy, form relationships, and shop. This article offers some observations about Big Data to help lawyers understand its scope, its implications for business, and the legal issues that come with it.

Big Data involves drawing data from a potentially wide variety of data sets that, historically, were never intended to be combined. Big Data applies analytical tools and processes to those data sets to see if meaningful correlations and relationships exist. Value is created when actionable insights are drawn from the analysis. The data sets may be drawn from separate systems within a single enterprise, for example, customer relationship management (CRM) and enterprise resource planning (ERP) systems or from external systems and data sources, for example, market analytics firms, geospatial records, government records, and weather systems.

Without a doubt, many enterprises were "doing" Big Data long before the term became worthy of capitalization, but several factors have brought the practice into the mainstream, including key enablers of cheap and massively scalable computing power and an exponentially growing ecosystem of networked (and therefore potentially accessible) data. Most importantly:

- Technologies have evolved to allow massively scalable data storage and data mining at an affordable cost. Until recently large scale data mining using older generations of technology simply was unaffordable for many users.
- The emergence of the "Internet of things"—networked devices that capture, store and transmit data in real time, such as mobile phones, networked computing equipment, industrial sensors, and a myriad of other sensors that are being deployed in our communities to monitor everything from traffic and weather to energy consumption. The Internet of things has created a vast new source of data that, individually,

may have little worth. But in very large quantities that data may reveal patterns of behavior or other characteristics of significant commercial or societal value.

- The emergence of online social media and new sources of information about individual preferences, "likes," habits, and social networks.

- The rapid emergence of Internet-based applications and software-as-a-service solutions to perform data analytics without crippling up-front investments.

The legal issues associated with such a broad set of technologies and potential applications are just beginning to be explored. Following are eight observations to help lawyers frame the issues and manage the risks.

Regulations Are Continually Evolving and Vary by Sector and Geography

Data privacy and consumer protection laws may apply to Big Data projects depending on the type of data involved, where it is collected, and how it is used. Because Big Data draws on a potentially wide range of data sources, more than one set of regulations may well apply to the data that is collected. This article focuses only on the European Union and the United States, but companies collecting or storing Big Data globally must consider that Canada, Australia, and most other countries now have their own data protection and/or privacy regulations.

Current Regulations

In the European Union, there is a well-established, broad-based EU Data Protection Directive to govern processing of personal data.[2] "Personal data" in this context means all information about an identified or identifiable natural person (a data subject).[3] "Processing" means any operation or set of operations that is performed on personal data.[4] Among other things, the EU Directive requires entities that process personal information (data controllers) to comply with principles[5] that restrict how data is used and protect certain rights of the data subjects. Data controllers are required to have a specific purpose for the data and to comply with the scope of that purpose. They must maintain the accuracy of the data collected, must destroy data when its purpose is over, must give data subjects access to the data collected and disclose who it is shared with, and must keep data secure from unlawful processing.

In addition to the EU Data Protection Directive, there is the E-Privacy Directive enacted in the European Union in 2002 to further protect data processing across public communications networks[6] and the "Cookie Directive" enacted in 2009 that requires service providers to meet higher security standards when processing data and to notify both data protection authorities and individuals when the security of the data they are processing is breached.[7]

In contrast to the broad-based approach taken in Europe, regulation in the United States tends to focus on specific industry sectors and geographies:

- 46 States in the United States (plus Puerto Rico and Guam) currently have laws requiring companies to notify consumers if their personal data is improperly accessed or disclosed.

- Massachusetts has taken state regulation one step further by requiring companies to require third-party partners to contractually commit to implement and maintain security measures when their services access personal data.[8] Similar regulations are under consideration in California and other states likely will follow suit in the near future.

- At the federal level there are numerous statues that regulate specific categories and uses of data. For example, the Health Insurance Portability and Accountability Act (HIPAA)[9] and the Health Information Technology for Economic and Clinical Health Act (HITECH)[10] regulate protected health information of individuals; Fair Credit Reporting Act (FCRA)[11] and Fair and Accurate Credit Transactions Act (FACTA) are designed to promote the accuracy, fairness, and privacy of information in the files of consumer reporting agencies, and to regulate the use and dissemination of consumer reports; COPPA[12] is designed to protect the privacy of children under 13 on the Internet; and the Gramm-Leach-Bliley Act[13] requires financial institutions—companies that offer consumers financial products or services such as loans, financial or investment advice, or insurance—to explain their information-sharing practices to their customers and to safeguard sensitive data.

In addition to laws and governmental regulations, companies in some industries must follow standards adopted in those industries in order

to conduct business. For example, companies that want to accept credit cards will be required to adopt certain data protection measures to comply with the Payment Card Industry Data Security Standard imposed by the card networks.

Regulatory Trends

The data privacy and consumer protection regulations affecting Big Data will evolve as lawmakers struggle to strike the right balance between individual privacy rights and the commercial and societal benefits that can be derived from Big Data techniques.

The US sectorial approach to regulation may trigger a backlash from companies and consumers as established data categories break down. For example, HIPAA/HITECH focuses on the health sector because, traditionally, that is where the personal health information has been created and stored among doctors, hospitals, insurers, and their service providers. Consequently, those laws regulate the use of personal health information obtained by healthcare providers in the course of providing healthcare services. Big Data has created an entirely new category of "medically inflected data"—information (such as online searches, GPS, and shopping data) about a person that is derived outside the health sector but which can be used to make health-related predictions and potentially profile individuals. In one famous example, a major US retailer identified potentially pregnant women from their online browsing habits and sent them targeted advertising. HIPAA/HITECH does not regulate this activity.[14] It is possible that regulators may seek to expand the scope of sectorial regulations in an effort to plug these gaps.

Examples of regulators and policy makers seeking to update consumer protections can be seen at all levels of government.

The European Commission is currently working on a General Data Protection Regulation to address challenges arising from new technology, the dramatic increase in the quantity and availability of Big Data and inconsistencies in the way in which the EU Data Protection Directive has been implemented and enforced across the EU Member States.[15] Among other things, the regulation would

- Extend to new types of data, for example, genetic data and online identifiers such as email addresses, IP addresses, and cookie identifiers;
- Require data controllers to obtain explicit consent for a specific purpose (implied consent would no longer be valid);
- Establish a "right to be forgotten" by enhancing an individual's right to have his personal data erased; and
- Require data controllers to implement transparent and easily accessible data processing policies.[16]

The White House recently published a Consumer Privacy Bill of Rights that incorporates many of the principles already adopted in the European Union including rights for consumers to control how their data is used, rights to access and correct data, and rights to limit the data that companies retain.[17] Although not binding on companies today, the Consumer Privacy Bill of Rights is intended to provide a framework for lawmakers, industry groups, companies, and others to consider as they make policy relating to Big Data in the future.

The Federal Trade Commission (FTC) also has published a framework to inform policymakers and encourage industry to self-regulate in accordance with three guidelines:

1. Companies should build privacy and security protection into new products (often referred to as Privacy by Design);
2. Privacy policies should be written in plain language that consumers can understand; and
3. Companies should provide greater transparency regarding data collection, use and retention.[18]

Specifically the framework recommends the development of "do not track" mechanisms that provide consumers with control over how their information is collected and used. Although there is no consensus on "do not track" at the national level among policymakers and industry, it continues to gain traction.[19]

Since 2003 California law has required operators of Web sites and online services to make clear disclosures about the personally identifiable information they collect. Effective January 1, 2014, the law will require additional disclosures about (1) how the operators respond to Web browser "do not track" and similar mechanisms designed to give

consumers control over how their data is collected and used; and (2) whether other parties may collect information about a consumer's online activities over time and across different Web sites when they use an operator's Web site or online service.[20] Although state laws do not apply nationally, when passed by a large state like California they often have the practical effect of ratcheting up the disclosure and consumer choice requirements for all large companies in the United States.

Not to be left out, the US Congress is getting in on the action as well. In October 2012, a bi-partisan Congressional Privacy Caucus sent inquiry letters to credit reporting agencies and data brokers, including Experian, Equifax, Acxiom, Epsilon, and Intelius, requesting information about how they collect, analyze, and then sell consumer information. The letters reflect a concern by legislators about the lack of transparency and potential harm that may result from these companies collecting, mining, and selling vast amounts of personal, medical, and financial data about consumers.

Perhaps in response to the Congressional inquiries mentioned above, some companies in the United States appear to be trying to stay ahead of the curve by providing consumers with additional transparency and choice about how their data is collected. For example, data broker Acxiom Corporation recently launched aboutthedata.com, a Web site that gives consumers the opportunity to see some of the data collected and profiles created about them.

Existing Regulations Will Continue to Be Applied in New Ways

Although none of the existing laws summarized above were designed specifically with Big Data in mind, many are written broadly enough to apply in new ways to companies that use Big Data techniques, both as service providers and as users. For example, in recent remarks, FTC Commissioner Julie Brill expressed concern about how companies may now combine data from multiple sources and replace traditional credit reports as the basis for making determinations about a consumer's credit or suitability for an insurance policy or mortgage.[21] For example, a life insurer might use data about a consumer's consumption patterns to evaluate the consumer's health or predict life expectancy; a potential employer might purchase geolocation data to vet potential employees; or a bank might use credit card spending data to create consumer profiles that determine the terms of a mortgage.[22] The Commissioner highlighted the need to ensure that the strict rules under the FCRA governing the use of traditional credit reports for these purposes are applied to Big Data techniques that are used for similar purposes.[23]

The FTC's willingness to act in this context is apparent from an enforcement action it brought against Spokeo, Inc.[24] In 2012, the FTC ordered Spokeo (a data broker that compiles and sells information about consumers) to pay $800,000 to settle FTC charges that it marketed information to companies in the human resources, background screening, and recruiting industries without taking steps to protect consumers as required under the FCRA. According to the FTC, Spokeo collects personal information about consumers from hundreds of online and offline data sources, including social networks. It merges the data to create detailed personal profiles[25] of consumers. The FTC alleged that in doing so Spokeo operated as a consumer reporting agency and violated the FCRA by:

1. Failing to make sure that the information it sold would be used only for legally permissible purposes;

2. Failing to ensure the information was accurate; and

3. Failing to tell users of its consumer reports about Spokeo's obligations under the FCRA, including the obligation to notify consumers if the data user takes an adverse action against the consumer based on information contained in the consumer report.

The FTC noted that this was the first FTC case to address the sale of Internet and social media data in the employment screening context.

New Uses for Data Will Test the Privacy Promises Made When the Data Was Captured

Data scientists search continuously for correlations, relationships, and actionable insights. This may involve drawing data from unexpected sources and combining it in ways that were never contemplated when the data was collected. As a result, it is increasingly common for companies to collect data that far exceeds their immediate needs with the hope that they can put it to productive use in the future.

This raises obvious issues in European countries where data controllers must

collect personal data for specified, explicit, and legitimate purposes; not use it in a way that is incompatible with those purposes; and then destroy it when it is no longer necessary for the purposes for which it was collected.

It also can create challenges for companies in the United States. Although US regulations generally are less restrictive than European regulations, most laws require companies to notify individuals in privacy policies about the data they collect and to then comply with those policies. Companies that violate this obligation may face enforcement action from the FTC.[26] A recent high profile example of this is Google, Inc.'s agreement to pay a $22.5 million civil penalty to settle FTC charges that it misrepresented to users of Apple Inc.'s Safari Internet browser that it would not place tracking "cookies" or serve targeted ads to those users, violating an earlier privacy settlement between the company and the FTC. The FTC notes that the settlement is part of the FTC's ongoing efforts to make sure companies live up to the privacy promises they make to consumers, and is the largest penalty the agency has imposed for a violation of an FTC order. In addition to the civil penalty, the order also requires Google to disable all the tracking cookies it had said it would not place on consumers' computers.[27]

As companies discover new ways to use and commercialize the data they collect, they must have the technical capabilities and policies to track the source of the data they collect and ensure that their new uses of that data are within the scope of the disclosures made when the data was collected.

Collection and Storage of Big Data Heighten the Risk and Magnitude of Security Breaches

The practice described in the section above of collecting and retaining Big Data for uses that may be determined in the future also conflicts with one of the fundamental principles of the "privacy by design" guidelines published by the FTC and others—data minimization. The idea of data minimization is that companies should use personal data only for the purposes for which it was collected, and then promptly destroy it in order to minimize the impact of a potential data breach. In addition to FTC guidelines, this concept is reflected in many sectoral and private regulations. For example, merchants are prohibited by payment network rules from storing card identification numbers and similar data obtained from credit cardholders in order to reduce credit and fraud losses following a data breach; the EU Data Protection Directive requires destruction of data after the specific purpose for which it was collected is achieved.

Collecting more data than is needed may result in increased costs to notify individuals of data breaches, exposure to lawsuits from individuals affected by the breach, loss of reputation or goodwill, credit and fraud losses if credit card data was implicated, and possibly enforcement actions from the FTC or other government authorities with jurisdiction over the matter. For example, the FTC recently brought a claim against Wyndham Hotels alleging that Wyndham failed to use adequate security measures to protect the customer data it collected.[28] Wyndham is defending itself by arguing (among other things) that the FTC does not have the authority to bring this type of action. The ability of the FTC to bring these types of claims in the future may be solidified or significantly reduced depending on how this case is decided.

In some cases, the potential benefits of collecting and storing Big Data indefinitely will outweigh the associated risks. Companies that make this determination, however, should consider whether there are additional things they can do to reduce the risk. For example:

- Ensure that the company understands the type of data being collected. As companies widen their data collection nets, they increase the likelihood of inadvertently collecting sensitive financial, health, and other data that may subject them to new regulations and increased liability. Many companies discover that they are storing personally identifiable information only after a data breach occurs and it's too late to limit their exposure.

- Design or modify systems to include reasonable safeguards and controls to protect the data collected. For example, "toxic data" (*e.g.*, personal information, credit card data) should be segregated from other less sensitive data and subject to more stringent access and retention policies. According to the FTC, the safeguards and controls should correlate to the sensitivity of the information collected, the amount of information collected, threats attendant to a particular network structure, the evolving field of commonly targeted vulnerabilities, and many other factors.[29] In the case of Wyndham the FTC

alleged that Wyndham failed to meet this standard in numerous ways: failing to limit access among different computer networks through the use of readily available measures, such as firewalls; failing to configure software properly to prevent the storage of payment card information in clear text; failing to ensure the Wyndham-branded hotels had adequate information security policies in place prior to allowing them to access Wyndham's computer network; failing to require servers attached to its networks to have the latest security patches from manufacturers; failing to change commonly known default passwords within its network; failing to follow best practices for password complexity; failing to inventory the computers on its network in order to permit Wyndham to identify the origin of intrusion efforts; failing to employ reasonable measures to detect and prevent unauthorized access; failing to follow proper procedures to prevent repeated intrusions; and failing to restrict third-party access to its network.[30]

- Understand the extent to which existing insurance policies cover data security breaches and consider whether it makes sense to supplement them with additional cyber-insurance.

De-Identifying Data May Not Be Sufficient

Advertisers, researchers, and users of data in many other industries have long argued that aggregating or de-identifying personal data can render it anonymous and thus allow unrestricted use without compromising individual data subject privacy. Until very recently, most regulators have accepted this argument as well in granting safe harbors or similar exceptions to data privacy regulations for data that has been anonymized. In the outsourcing and cloud-computing industry, customers have followed suit in routinely granting their service providers the right to use customer data so long as the service providers aggregate it with other data and remove personally identifiable data prior to disclosing it.

In recent years, computer scientists have demonstrated that anonymized data can be "re-identified" by linking anonymized records to outside information.[31] For example, with an individual's zip code, birth date, and gender, researchers can identify the person with certainty 87.1 percent of the time (based on 1990 census data). More recently, researchers have been able to identify individuals using only the reviews they posted on Netflix combined with publicly available information.[32] In each case, researchers found that seemingly anonymous data contained unique attributes and other clues that enabled them to re-identify it with individuals. Once a person has been identified, the effect is compounded as it becomes easier to associate more and more information with that person.

The ease with which researchers can re-identify anonymized data has several implications in the outsourcing and cloud-based service industry. Among them:

- Regulations generally define the "personal data" that they cover broadly as information that can be used to identify a person. With re-identification, seemingly innocuous information such as search queries and Netflix reviews could arguably fall within the definition of personal data and be subject to additional regulation.

- Regulators are beginning to explicitly address new types of data (*e.g.*, IP addresses, cookie identifiers).[33]

- Re-identification also may lead to increased liability. For example, if personal information collected by a company is disclosed by the company's service provider and later re-identified, the company may face claims from its end users and possibly fines from regulators; the service provider may face claims from the company for failing to adequately anonymize the data.

- If the "right to be forgotten" is implemented under the proposed General Data Protection Directive in Europe, customers that make personal data about individuals in Europe available to their service providers will need to ensure that they can direct their outsourcing and cloud-service service providers to "erase any links to, or copies or replications of that personal data."[34] Many service providers likely will lack the capability to do this with data that has been aggregated with other personal data.

Actions Based on Some Data Correlations Could Be Unlawfully Discriminatory

It is easy to imagine scenarios in which practices taken based on Big Data analysis could fall afoul of anti-discrimination laws. Most, if not all, Big Data analysis involving consumers also involve "profiling" in

some form or another. Profiles based on age, race, sexual orientation, or other characteristics protected by anti-discrimination laws obviously are a problem, but what about profiles based on characteristics that are not directly subject to anti-discrimination laws, but which are disproportionately associated with protected classes? Examples could include profiles built on job applicants, existing employees, and loan applicants.

Anti-discrimination law supports claims based not only on intentional discrimination but also on discrimination arising from "disparate impact" which can be proved by statistical evidence that a protected class has been adversely impacted by a practice as compared to similarly situated persons or groups. Federal government agencies have been more aggressive in recent years when pursuing enforcement actions based on disparate impact. Big Data techniques could expose those who deal with consumers based on their Big Data analytics to claims that the decisions (*e.g.*, differential pricing of financial services) have an unlawful discriminatory impact on a protected class. In addition, if a claim is litigated the effort associated with retrieving and analyzing the data in issue could add significant cost to the litigation.

Actions Based on Some Data Correlations Could Run Afoul of Employment Law

In the employment context, job candidates may be subject to a range of pre-employment screens including prior employment, personal references, criminal records, driving history, credit history, and other checks such as Facebook pages. Some employers also monitor existing employees including their online activities. Federal and State laws regulate the collection and use of such information.

Automated screening systems that use Big Data techniques have emerged to help employers deal with very high numbers of job applications. HR service companies also run screens on behalf of employers, including consumer reports. If not carefully managed, employment related decisions based on information obtained from these systems could fall afoul of hiring and employment laws. For example:

- FCRA applies when an employer obtains a background check from a "consumer reporting agency"[35] regarding an applicant or employee when the information sought by the employer pertains to credit information or the individual character, general reputation, personal characteristics, or mode of living. Certain specific disclosures must be made and consent given by the applicant before an employer can undertake such a background check. Some states have similar laws with their own specific disclosure and consent requirements.[36] Data analytics companies that collect this sort of information could fall within the definition of a "consumer reporting agency" although they may not think of themselves as such.[37]

- Federal law and certain state laws restrict information that can be included in a consumer report. For example, in California the report cannot include bankruptcies more than 10 years old[38] or lawsuits, or judgments more than seven years old.[39]

- The Federal Genetic Information Nondiscrimination Act of 2008[40] (GINA) prohibits employers from using an individual's genetic information when making employment decisions.

- Various states prohibit employment decisions based on certain criminal histories.[41]

- An automatic decision to deny employment based on a criminal conviction could result in a claim of race discrimination on the theory that such a rule has a disproportionate impact on certain protected groups and has a tenuous or insubstantial relation to job qualifications.[42] In California credit checks may only be run on employees or applicants who hold or apply for certain job positions.[43] The job applicant must be notified of the nature and scope of the inquiries being made.[44]

The examples are illustrative only; other states have similar laws. Employers must be aware of the laws applying to job candidates and employees in each jurisdiction in which they work.

Clearly, merely collecting information creates a potential liability risk, regardless of whether the employer takes the information into account in making an employment decision. Employers, therefore, should tailor their information collection efforts to the needs of each job classification to avoid collecting data that is not necessary for the employment decision and to comply with applicable law. The technique of collecting large quantities of data, as is commonly the case in Big Data projects, may put an employer at increased risk of legal claims.

Analytics-as-a-Service, Cloud Technology and Outsourcing Complicate the Analysis of Big Data Options and Risks

Increasingly, Big Data projects draw on multiple internal and external sources of data, analysis, and value. They rely on a patchwork of technologies and services that are sourced internally and from outside the enterprise. Analytics-as-a-service solutions have emerged ranging from specific data or transaction types to generic capabilities in enterprise platforms including SAP, Oracle, and Microsoft analytics. Moreover, external service providers who may not think of themselves as part of a Big Data solution may still generate or capture useful data as a byproduct of their core activities. These service-based platforms complicate how a company mines its data and controls for the risks of its use.

For example, a retailer may wish to analyze browsing and purchase data drawn from its internal systems, customer interactions managed by an outsourced call center, and demographic or geographic data drawn from independent Web sites and data sources. The results of the analysis may be visible to the external service providers. The retailer will need to look at each component of the value chain and assess its internal policies and contracts to make sure they do the following:

- Correctly acknowledge the source of data that is critical to the analysis.

- Appropriately assign ownership of the resulting analytics and information. External service providers often will have a stake in the data analysis. If the service provider's proprietary processes or data are used in the analysis, the provider may expect to own or have certain usage rights in the resulting analysis. The company should be wary of making itself dependent on critical data or analytics that it cannot obtain from any other source.

- Appropriately protect the company's competitive advantage in that data. Is that advantage protectable under copyright law, by patenting, or must it be kept secret? The company should address rights to use and further analyze data for other purposes; and to use and exploit new data modeling inventions and insights.

- Do not infringe external data suppliers' limits on use. It is not unusual for data providers to specify limits, particularly on the repackaging or resale of their data.

- Do not use or disclose consumer data in a way that conflicts with the company's published data use policies and data subject consents, where applicable.

Niche analytics-as-a-service providers offer relatively fast and cheap implementations with minimal up-front investments of time or effort. These solutions can be ideal for companies that pursue an R&D strategy of "fail often, fail fast." However, niche providers might not be able to integrate with more complex solutions, and this could impede later efforts to share and exploit data in new ways. The company should evaluate the adaptability of these external analytics and data sources.

Another issue to be aware of is that company information held by an external service provider may be at higher risk of compelled disclosure in criminal investigations than if the information were stored in the company's internal systems.[45] If presented with a subpoena or court order to produce records stored electronically for a client, the service provider may lack the knowledge and motivation to push back against over-broad demands. Importantly, the data owner may not be informed of the subpoena or order until it is too late.[46]

Conclusion

Over the past decade there has been exponential growth in the quantity and types of data created. According to IBM, 90 percent of all the data in the world has been generated over the last two years alone, and we have every reason to expect this trend to continue. Individuals will generate more and more data through mobile devices, cloud-based services, social media platforms, and other new technologies, and companies will continue to collect and find new productive ways to use it. While this increased collection and use of data has the potential to bring new efficiencies and value to society, it also presents challenges for lawmakers and regulators as they seek to balance the rights of individuals, businesses, and researchers.

With the General Data Protection Regulation, the European Union is attempting to address many of the data privacy and consumer protection concerns raised by Big Data through comprehensive legislation that will apply across all member states. In the United States, at least for the foreseeable future, Big Data will continue to be regulated through a patchwork of State and Federal regulations that lawmakers and regulators will continue to supplement and interpret to address concerns in specific sectors and geographies.

Lawyers who advise Big Data users will need to monitor these trends closely.

Endnotes

1 *www.languagemonitor.com.*

2 Directive 95/46/EC of the European Parliament and of the Council of October 24, 1995, on the protection of individuals with regard to the processing of personal data and on the free movement of such data, 1995 O.J. (L281) 31 (the EU Data Protection Directive).

3 EU Data Protection Directive, article 2(a).

4 EU Data Protection Directive, article 2(b).

5 EU Data Protection Directive, articles 6, 12 and 15.

6 Directive 2002/58/EC of the European Parliament and of the Council of July 12, 2002, concerning the processing of personal data and the protection of privacy in the electronic communications sector, 2002 O.J. (L 201) 37 (the E-Privacy Directive).

7 Directive 2009/136/EC of the European Parliament and of the Council of November 25, 2009, amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws, 2009, O.J. (L 337) 11.

8 *See* The Massachusetts General Law Chapter 93H and its new regulations 201 CMR 17.00.

9 Health Insurance Portability and Accountability Act of 1996 (HIPAA), Pub. L. No. 104-191, 110 Stat. 1936 (1996), codified at 42 U.S.C. § 300gg and 29 U.S.C § 1181 *et seq.* and 42 U.S.C. § 1320d *et seq.*

10 Health Information Technology for Economic and Clinical Health (HITECH) Act, Title XIII of Division A and Title IV of Division B of the American Recovery and Reinvestment Act of 2009 (ARRA), Pub. L. No. 111-5, 123 Stat. 226 (Feb. 17, 2009), *codified* at 42 U.S.C. §§300jj *et seq.*; §§17901 *et seq.*

11 Fair Credit Reporting Act (FCRA) and Fair and Accurate Credit Transactions Act (FACTA), 15 U.S.C. § 1681 *et seq.*

12 Children's Online Privacy Protection Act of 1998 (COPPA), 15 U.S.C. §§ 6501–6506 (Pub.L. 105–277, 112 Stat. 2581-728, enacted October 21, 1998).

13 Public Law 106-102, 15 U.S.C. § 6801, et seq. and 16 C.F.R. § 313, 65 Fed. Reg. 33646 (May 24, 2000)).

14 For a detailed discussion of the limitations of current HIPAA/HITECH laws, see Nicolas P. Terry, "Protecting Patient Privacy in the Age of Big Data," 81 *UMKC L. Rev.* 385 (2012).

15 Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), COM (2012) 11 final (Jan. 25, 2012).

16 Marc Rotenberg and David Jacobs, "Updating the Law of Information Privacy: The New Framework of the European Union," 36 *Harv. J. L & Pub. Pol'y* 605, 631–634 (2013).

17 The White House, "Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting innovation in the Global Digital Economy" (2012).

18 *See* FTC, "Protecting Consumer Privacy in an Era of Rapid Change: A Proposed Framework for Businesses and Policymakers," 39-42 (Dec. 1, 2010).

19 *See, e.g.*, (1) On February 28, 2013, Sen. Jay Rockefeller introduced the "Do-Not-Track Online Act of 2013" for consideration by Congress; (2) multiple references to "do not track" in White House framework referenced above; (3) CA law referenced below.

20 California Business & Professions Code § 22575, as amended on September 27, 2013 by Assembly Bill 370.

21 *See* "Introduction to the North Carolina Law Review Symposium, *Social Networks and the Law: Privacy & Consumer Protection in Social Media*," 90 N. C. L. Rev. 1295, 1304 (2011-2012) (Commissioner Brill Remarks).

22 *Id.*

23 *Id.*

24 *http://www.ftc.gov/opa/2012/06/spokeo.shtm.*

25 The profiles contain such information as name, address, age range, and email address. They also might include hobbies, ethnicity, religion, participation on social networking sites, and photos.

26 See *http://www.ftc.gov/opa/reporter/privacy/privacypromises.shtml* for links to 32 legal actions taken by the FTC as of May 2011 against organizations that have violated consumers' privacy rights, or misled them by failing to maintain security for sensitive consumer information.

27 *http://www.ftc.gov/opa/2012/08/google.shtm.*

28 *See* Federal Trade Comm'n v. Wyndham Worldwide Corp., NO CV 12-1365-PHX-PGR, filed in US District Court, D. Arizona, March 25, 2013 (FTC Complaint).

29 *See* page 20 of the FTC's response to Wyndham Worldwide Corporation's motion to dismiss in Civil Action No. 2:13-CV-01887-ES-SCM in the US District Court for the District of New Jersey.

30 *See* amendment to FTC complaint on August 8, 2012 (ECF No. 28).

31 Paul Ohm, "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization," 57 *UCLA L. Rev.* 1701 (2009-2010).

32 Arvind Narayanan & Vitaly Shmatikov, "Robus De-Anonymization of Large Sparse Datasets," 2008 *IEEE Symp. On Security & Privacy* 111.

33 *See* paragraph 24 of the General Data Protection Directive: "When using online services, individuals may be associated with online identifiers provided by their devices, applications, tools and protocols, such as Internet Protocol addresses or cookie identifiers. This may leave traces, which combined with unique identifiers and other information received by the servers, may be used to create profiles of the individuals and identify them. It follows that identification numbers, location data, online identifiers or other specific factors as such need not necessarily be considered as personal data in all circumstances."

34 Paragraph 54 of the General Data Protection Directive.

35 A list of self-reported consumer reporting agencies can be found at: *http://files.consumerfinance.gov/f/201207_cfpb_list_consumer-reporting-agencies.pdf.*

36 *See, e.g.*, California's Consumer Credit Reporting Act (Cal. Civil Code §1785.1, *et seq.*) and the California Investigative Consumer Reporting Agencies Act (Cal. Civil Code §1786, *et seq.*

37 Under the FCRA, the term "consumer reporting agency" means "any person which, for monetary fees, dues, or on a cooperative nonprofit basis, regularly engages in whole or in part in the practice of assembling or evaluating consumer credit information or other information on consumers for the purpose of furnishing consumer reports to third parties, and which uses any means or facility of interstate commerce for the purpose of preparing or furnishing consumer reports." 15 U.S.C. § 1681a(f). Under California's Investigative Consumer Reporting Agencies Act the term "investigative consumer reporting agency" means any person who, for monetary fees or dues, engages in whole or in part in the practice of collecting, assembling, evaluating, compiling, reporting, transmitting, transferring, or communicating information concerning consumers for the purposes of furnishing investigative consumer reports to third parties, but does not include any governmental agency whose records are maintained primarily for traffic safety, law enforcement, or licensing purposes, or any licensed insurance agent, insurance broker, or solicitor, insurer, or life insurance agent." CC §1786.2(d).

38 15 U.S.C. § 1681c(a)(1).

39 15 U.S.C. § 1681c(a)(2).

40 Public Law 110-233.

41 *See, e.g.*, Cal Lab. Code § 432.7; NY Exec. Law. § 296.16; N.Y. Correct. Law § 752; MGL c. 151B § 4(9) (Massachussetts).

42 *See* Griggs v. Duke Power Co., 401 US 424, 431 (disparate impact theory of liability); Gregory v. Litton Sys., Inc., 472 F.2d 631, 632 (9th Cir. 1972) *aff'd as mod.* (9th Cir. 1972) (facially neutral employment questionnaire that requires applicant to reveal arrest record may violate Title VII if it operates to bar employment of black applicants in greater proportion than white applicants.)

43 Cal. Labor Code § 1024.5.

44 California Civil Code § 1786.16.

45 Under the Federal Stored Communications Act, 18 U.S.C. § 2701 *et seq.*, government entities may require service providers holding electronically stored data to disclose the contents of the data by means of a warrant or, in certain circumstances, a court order or subpoena. 18 U.S.C. § 2703(b).

46 Notice of the subpoena or order must be given to the data owner but that notice may be delayed for up to 90 days if a court determines that the notice may have an adverse result including danger to safety, flight from prosecution, or destruction or tampering with evidence. 18 U.S.C. § 2705(a).

Pillsbury Winthrop Shaw Pittman LLP | 1540 Broadway | New York, NY 10036 | +1.877.323.4171