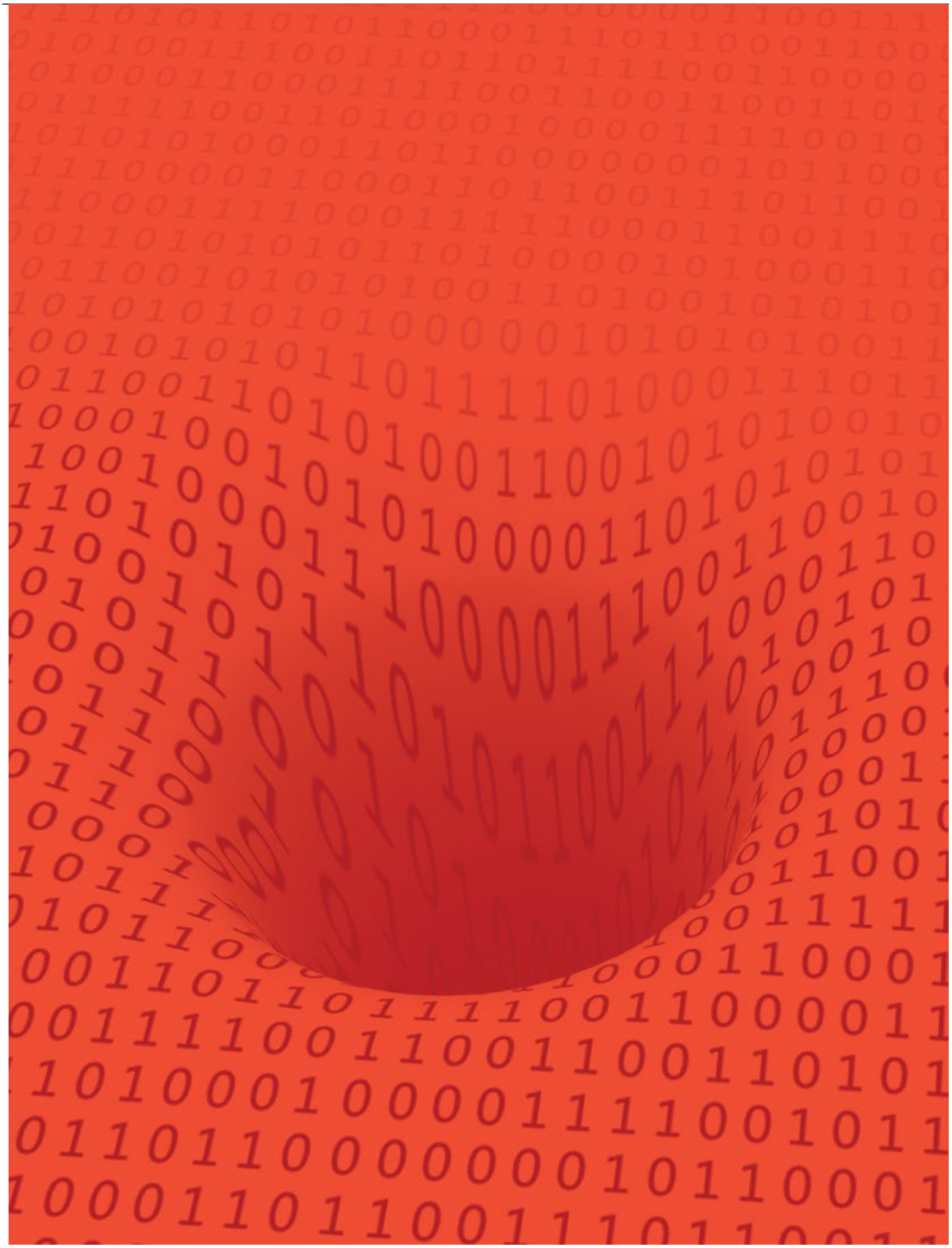
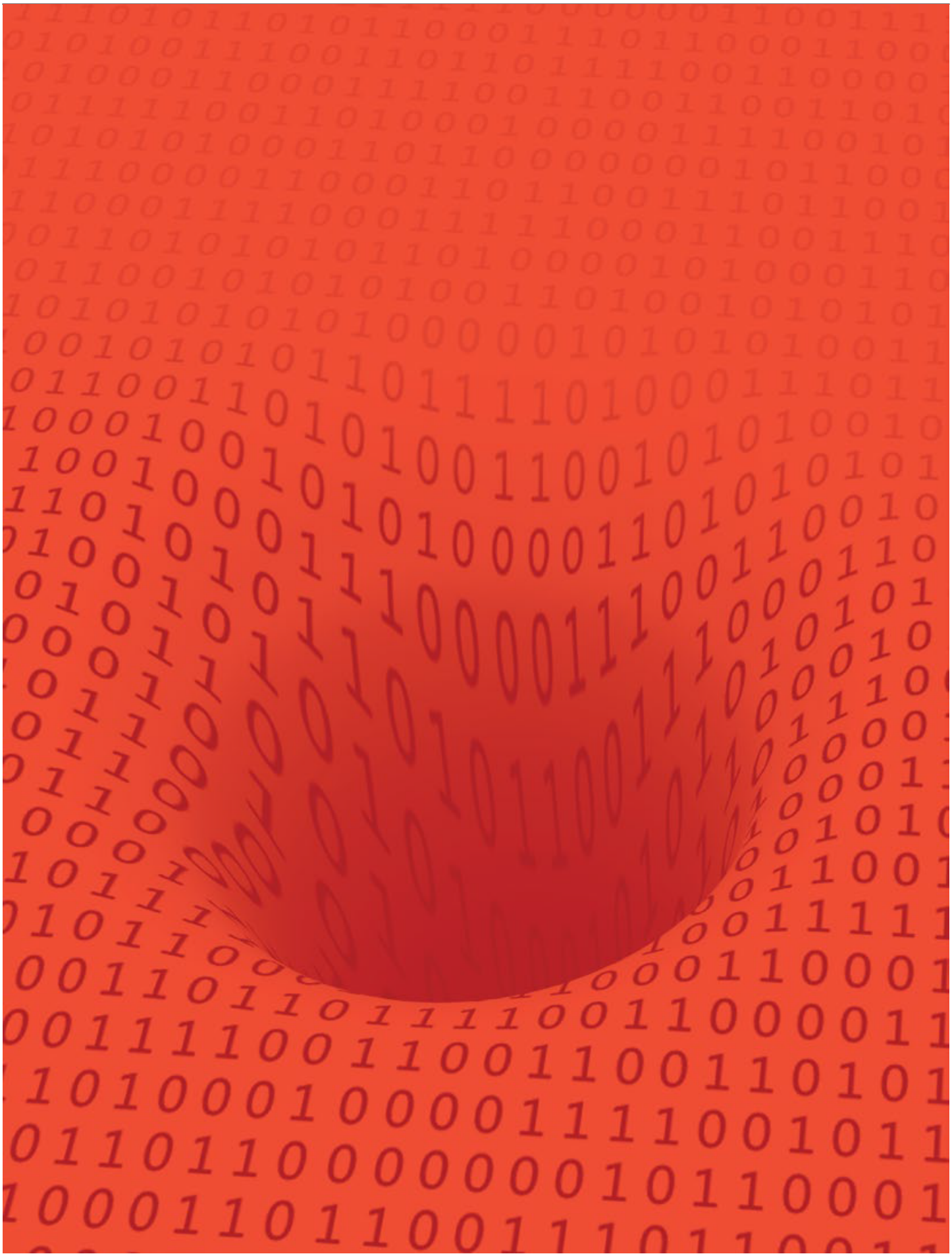


Avoiding the Ethical Perils and Pitfalls of Big Data

Compliance and Ethics

Technology, Privacy, and eCommerce





CHEAT SHEET

- ***Violations ahead.*** By 2018, half of business ethics violations will occur through the improper use of big data analytics, according to the research firm Gartner.
- ***If the shoe fits.*** While the use of data analytics allows an organization to tailor its user experience, it can also blur ethical boundaries by prioritizing revenue over privacy.
- ***Looking abroad.*** In Europe, unless express consent of the target is provided, legislation specifically prohibits the use of data analytics to conduct "automated individual decision making."
- ***Maintaining your integrity.*** To mitigate risk, it's important to ensure that data is used in compliance with international privacy standards and is collected with respect for an individual's rights.

The increased connectivity of people and things is creating previously unimaginable amounts of data. This volume, coupled with the rapid pace of data generation, provides unprecedented real-time insights into the habits, statistics, and patterns of people and processes. The ability to leverage data into actionable intelligence is now critical when making strategic business decisions.

Through the use of data analytics, businesses can understand consumer behavior to refine marketing efforts, predict future personnel behavior, as well as profile individuals and forecast behavior to drive decisions based on data voluntarily provided or gleaned from behavior.

As new algorithms pioneer how we gain actionable intelligence, the nature of the data — and the resulting outcomes — can expose a business to ethical challenges. The line between appropriate action and misuse is growing increasingly unclear. With the lack of clarity, it is no surprise that by 2018 half of business ethics violations will occur through the improper use of big data analytics, according to the research firm Gartner. As your company adopts or continues to use big data analytics, how will you properly guide through this ethical quagmire?

This article looks at various stages of the data analytics process, where decisions to move forward or exercise restraint are key factors in the ethical equation.

Data information collection

Before you can take action on data, you first have to collect it. This is where opportunities for ethical and legal missteps begin to arise, particularly in terms of data veracity and maintenance.

In the United States, there is no comprehensive federal law that regulates privacy. Instead, privacy is generally governed by state and by industry sector. For example, both the Fair Credit Reporting Act (FCRA) and the Health Insurance Portability and Accountability Act (HIPAA) contain mechanisms for individuals to access their personal information and rectify inaccuracies. Recently, the collection and use of inaccurate personal information under the FRCA came under scrutiny in both a class action lawsuit and a regulatory investigation.

Spokeo, Inc. operates a website where users can access and run queries about individuals by name, email, or phone number. When a query is submitted, Spokeo searches a wide range of databases and provides a profile containing information such as the individual's address, phone number, marital status, approximate age, occupation, hobbies, finances, shopping habits, and musical preferences. Thomas Robins alleged that, at some point and time, he became aware that someone submitted a Spokeo search on him and that his profile contained inaccurate information, stating that he was married, had children, was in his 50s, had a job, was relatively affluent, and held a graduate degree. Robins filed a class-action complaint claiming, among other things, that Spokeo willfully failed to comply with the FCRA's requirement to "follow reasonable procedures to assure maximum possible accuracy of " consumer reports. As it currently stands, this case has been remanded back to the Ninth Circuit Court of Appeals by the US Supreme Court for further deliberation on standing.

Spokeo, Inc. v. Robins, No. 13-1339, slip op. at 3-4 (S. Ct. May 16, 2016).

Spokeo was also the subject of a Federal Trade Commission investigation for allegedly violating the FRCA by, among other items, failing to ensure the accuracy of information collected and conveyed. The FTC and Spokeo settled the matter agreeing to the imposition of an US$800,000 civil penalty, as well as barring Spokeo from future violations of the FRCA. To ensure compliance with the order, Spokeo is obligated to submit compliance notices with the FTC for 20 years, and is subject to other recordkeeping and monitoring requirements.

Spokeo highlights the ramifications that can result from failing to properly assess data collection and use for an ethical perspective. If the information collected from sources cannot be verified or processs are not in place to account for accuracy: Is it ethical to rely on that data? Failure to appropriately take these variables into consideration may result in reputational and monetary consequences.

In Europe, data privacy is a fundamental right. The European legislative framework includes a variety of obligations for businesses and organizations to ensure data accuracy and the safeguarding of individuals' rights. These measures include embedding privacy obligations within the design of business procedures (known as "privacy by design"), requiring a response to any petition from any individual to correct wrong data or to erase data, discontinuing the use of personal data when not explicitly required for business or organizational purposes. Lack of compliance with such legal obligations may be subject to heavy fines — up to €20 million or four percent of annual revenue.

Additionally, members of international organizations such as the Organisation for Economic Co-operation and Development (OECD) and Asia-Pacific Economic Cooperation (APEC), along with jurisdiction-specific data protection schemes (i.e., the United Kingdom's Information Commissioner's Office's Data Protection Principles), charge businesses with the responsibility to abide by certain data protection principles, which include the obligation to maintain accurate personal data.

> As data collection increasingly comes under scrutiny, companies are taking a proactive approach by being more transparent in their public-facing privacy policies, adopting verbiage that reflects aspects of applicable privacy principles.

As data collection increasingly comes under scrutiny, companies are taking a proactive approach by being more transparent in their public-facing privacy policies, adopting verbiage that reflects aspects of applicable privacy principles. But beyond this, data collection points and work process flows will also need to be assessed internally to account for these ethical pitfalls. Questions to initiate

conversation regarding data accuracy include:

- How often is data verified? (to the extent possible)
- How “fresh” is the data, and what parameters reinforce this?
- What steps are taken to vet mined and third party data?

Abiding by ethical principles, such as ensuring the accuracy of data through the entire collection lifecycle, will help protect the company from exposure to ethical and legal claims. It will also support effective decision making that will achieve organizational objectives.

Analysis and use

The next stage in this process is the analysis and use of data. It presents ethical challenges ranging from the development of the algorithm or query and implementation of automated analysis processes to the interpretation and use of the results.

Precautions should be taken when developing search queries. Avoid queries where the responses, despite the anonymization, will lead to the exclusion of certain races or classes. Remember, if it is illegal to do it in person, it is illegal to do it electronically.

In the 2015 case of *State of Wisconsin v. Eric L. Loomis*, data analytics came to the forefront with the use of the COMPAS assessment, an algorithm used to calculate the likelihood that someone will commit another crime. The results come from an analysis of past conduct that included data such as criminal and parole history, age, employment status, social life, education level, community ties, drug use, and religious beliefs.

The pitfall of using such an algorithm to influence sentencing determinations is that the individual will be lost in group characteristics. Eric L. Loomis appealed his sentence when he was subject to a higher prison term due to the COMPAS score, which indicated that he was at “high risk” of committing another crime. Loomis argued that his due process rights were violated because the company that makes the test does not reveal how COMPAS weighs answers to arrive at the risk score. He also argued that the evaluation treat men as higher risk than women.

The Wisconsin Supreme Court ruled that the decision regarding a longer or shorter sentence cannot solely be based on scoring systems (or the scoring system cannot be the only reasoning or factor to establish the length of the sentence). Sentencing courts, however, are entitled to include scoring systems among the many factors used to determine lengths of sentences.

While results leading to disparate impact should be avoided, the lines are not always so clear-cut. As more companies increase the use of data analytics to tailor experiences for consumers, questions of about big data ethics will continue to arise. Take Orbitz Worldwide, Inc., which discovered in 2012 that people who use Mac computers were more willing to pay higher nightly rates. Based on this, Orbitz chose to alter the view of search results dependent on whether the consumer was using a Mac or a PC, listing the more expensive options first for Mac users.

The retailer Target also experienced issues when applying big data and data mining analysis results in order to predict which of their customers were pregnant — a time when purchase behavior is most in flux. The predictive effort was designed to identify these customers and deliver targeted and timely promotional offers to reach consumers when they are more likely to change spending habits and embrace new ones.

When data analytics results in categorizing customers and offering different products or prices based on those categories, problems can arise. Pam Dixon, founder of World Privacy Forum, notes "[D]etermining whether someone is going to be a loyal customer is fine. But then if you're changing the way you treat your customer based on that, that's where the questions come in."

From a European perspective, legislation specifically prohibits this type of analysis as well as the use of *"automated individual decision making, including profiling."* Data controllers are forced to introduce suitable measures to safeguard data subjects' rights and freedoms, including the right to obtain human intervention on the part of the organization that has implemented such automated procedures — allowing individuals to express their point of view and to contest the decision.

Safeguards also include applying restrictions to automated decision-making processes on sensitive data, such as health and ethnicity. In addition to the above described right of not being submitted to *"automated decision making,"* the legislative framework provides additional safeguards to individuals submitted to big data procedures: If any data controller wishes to use an individual's data for purposes differing from those that have been consented to consented to, the data controller must consider several factors to ascertain whether the new purpose is compatible with the initial purpose.

According to Article 6.4 of the REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, of 27 April 2016, on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), such factors are (a) any link between the purposes for which the personal data have been collected and the purposes of the intended further processing; (b) the context in which the personal data have been collected, in particular regarding the relationship between data subjects and the controller; (c) the nature of the personal data, in particular whether special categories of personal data are processed, pursuant to Article 9, or whether personal data related to criminal convictions and offences are processed, pursuant to Article 10; (d) the possible consequences of the intended further processing for data subjects; (e) the existence of appropriate safeguards, which may include encryption or pseudonymisation. This compatibility test was firstly introduced in the European legislative framework by an Opinion issued in 2013, by the Data Privacy Working Party, on Purpose Limitation.

If this "test of compatibility" confirms the new purposes are incompatible with the ones consented to by the concerned individual, the organization is forced to choose between two options:

1. anonymize data; or,
2. contact the individuals to obtain consent for the new purposes.

This requirement is prompting many European organizations to refocus procedures related to personal data processing by obtaining blanket consent for targeting, profiling, improving the experience of clients, risk assessment, and other data analysis measures.

Lack of compliance with such obligations may not only lead to the aforementioned heavy fines, but also to the obligation to indemnify the concerned individuals for any damages caused. As counsel, you will need to be prepared to offer legal guidance when developing questions in addition to guidance on the evaluation and use of results.

Data protection impact assessment

One final obligation introduced by the new European General Data Protection Regulation (GDPR),

which applies to big data projects, is related to the implementation of Data Protection Impact Assessment (DPIA) and requires prior approval from data protection authorities for any data processing with a high-risk of negative impact on the data protection rights of individuals.

According to the GDPR, which will go into effect May 2018, the minimum contents of a DPIA must include the following:

- A systematic description of the envisaged processing operations and the purposes of the processing, including, where applicable, the legitimate interest pursued by the controller;
- An assessment of the necessity and proportionality of the processing operations in relation to the purposes;
- An assessment of the risks to the rights and freedoms of data subjects; and,
- The measures envisaged to address the risks — including safeguards, security measures and mechanisms to ensure the protection of personal data and to demonstrate compliance with the law, taking into account the rights and legitimate interests of data subjects and other persons concerned.

Under US legislation, such assessments are not compulsory when dealing with personal data collected for commercial purposes. However, implementation should be considered. Besides showing proof of due diligence, it will be beneficial when extending the activity to European jurisdictions, or when considering certification under Privacy Shield, the new framework that replaces the Safe Harbor agreements that previously addressed exchange of personal data between the European Union and the United States.

Conclusion

When undertaking big data projects, in-house counsel should consider the main challenges from an ethical point of view:

- Ethical lines are prone to be violated when objectifying or classifying the data of individuals.
- Likewise, data tied to personal aspects of an individual’s life versus aspects of customer behavior present greater risk.
- Even if managed ethically, using data to direct operations can backfire — tarnishing the organization’s reputation with customers and stakeholders.

To mitigate ethical risks, here are some final recommendations:

- Define and enforce rules for collection of data, to ensure veracity and accuracy;
- Ensure use of data is obtained according to corporate ethical standards and with respect for individuals’ rights;
- Do not misrepresent the quality or completeness of data;
- Be transparent with individuals regarding the collection and use of their data; and,
- Use data fairly to avoid objectification and manipulation.

Soo Y. Kang

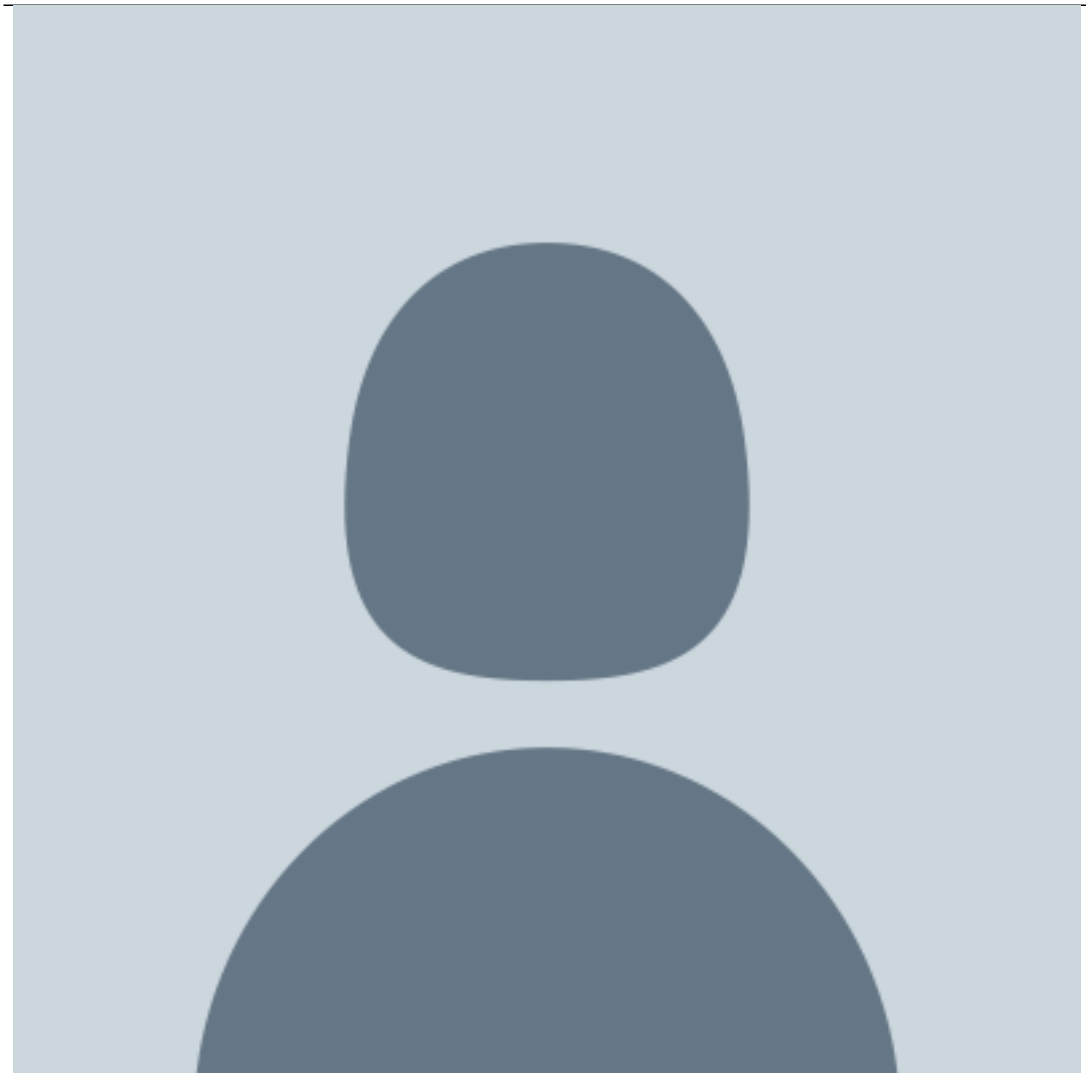


General Counsel and Director of the Consulting Division

Zasio Enterprises Inc.

Zasio Enterprises Inc. is a global leader in information governance offering technology and strategic consulting services.

Carlos Pérez



Partner and Head of Information Technology and Compliance

ECIJA

ECIJA has offices in Madrid and Barcelona, Spain, and is a Meritas member firm. With experience in information technology and telecommunication law, Pérez has advised leading Spanish and international companies on data protection, IP, telecommunications, IT outsourcing, and more.