Big Data Equals Big Need for Robust Data Governance

PRESENTED BY

SourceMedia™
MARKETING SOLUTIONS GROUP

SPONSORED BY

Io-Tahoe®

More than 80 countries and independent territories, including nearly every country in Europe and many in Latin America and the Caribbean, Asia, and Africa, have now adopted comprehensive data protection laws

I Introduction

Businesses are looking to strategically leverage their data assets for competitive advantage, operational agility and revenue generation. To do so, however, organizations need to first put effective data governance programs in place. More specifically, businesses need to comply with regulations such as the much-anticipated European Union's General Data Protection Regulation (GDPR). A look at the need for data governance, what's involved with GDPR compliance and its likely challenges illustrate the importance of data flow and point to the need for advanced data management solutions.

II Big Data: With Opportunity Comes Responsibility

The era of big data has arrived, and businesses can take advantage of all it has to offer. Organizations can benefit from big data by collecting market and customer intelligence, by improving internal efficiency and operations, and by extracting data-driven insights to improve their product offerings to enhance the customer experience, according to Bernard Marr, an internationally renowned business author.[1]

But there's also a flip side to working with data. More specifically, organizations need to manage the tsunami of data in a responsible manner. Indeed, various government regulations demand that businesses put data governance programs in place to protect data. The problem is most businesses don't really have a handle on their data coffers. In fact, only about 0.5 percent of all data is currently analyzed, and that percentage is shrinking as more data is collected.[2]

"The rate at which we're generating data is rapidly outpacing our ability to analyze it," said Professor Patrick Wolfe, Executive Director of the University College of London's Big Data Institute in an article that was published in Business Insider. "The trick here is to turn these massive data streams from a liability into a strength."[2]

II The Data Regulatory Landscape

The first step in ensuring that data does not turn into a liability comes with governing it correctly. More specifically, organizations need to understand and comply with a variety of data governance regulations, many of which assess financial penalties for non-compliance. More than 80 countries and independent territories, including nearly every country in Europe and many in Latin America and the Caribbean, Asia, and Africa, have now adopted comprehensive data protection laws.[3]

While data regulations exist across the globe, the GDPR, which comes into effect on May 25, 2018, is of particular interest as it represents perhaps the most comprehensive regulatory initiative designed in the EU to protect personally identifiable information to date. The GDPR requires compliance with a variety of stipulations for not only companies in the EU but carries a global footprint as well, as it affects any companies that have customers in the EU. Failure to fall into step with these regulations could result in serious financial repercussions. Organizations can be fined up to €20 million/ $23.5 million or 4 percent of the annual worldwide turnover of the preceding financial year for infringement of the regulation's provisions.

The GDPR regulations protect individuals' right to privacy by restricting organizations from processing personal information of individuals unless they have been freely given a specific, informed and unambiguous indication of consent.

IV An Overview of GDPR Compliance

The GDPR regulations protect individuals' right to privacy by restricting organizations from processing personal information of individuals unless they have been freely given a specific, informed and unambiguous indication of consent.[4] More specifically, the regulations are based on the following principles of personal data protection:

- Lawfulness, fairness and transparency: Personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject.
- Purpose limitation: Personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes.
- Data minimization: Personal data shall be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed.
- Accuracy: Personal data shall be accurate and, where necessary, kept up to date.
- Storage limitation: Personal data shall be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed.
- Integrity and confidentiality: Personal data shall be processed in a manner that ensures appropriate security of the personal data, including protection against unauthorized or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organizational measures.
- Accountability: The controller shall be responsible for and be able to demonstrate compliance with the GDPR.[5]

To comply with GDPR, organizations must make it simple and easy for customers and others to:

- Provide consent to sell their personal data to third parties.
- Notify their customers of any data breach or risk of breach within 72 hours.
- Provide customers with confirmation of whether their personal data is being processed and how.
- Support customers' "right to be forgotten" by enabling customers to request their personal data be erased and stop it from being distributed.
- Allow individuals to obtain and reuse their personal data for their own purposes by transferring it across different IT environments.
- Put processes in place that are designed to protect data from the get-go, when designing software, systems, websites etc.[6]

While GDPR protects all personal information, additional protections are applied to "sensitive personal data." Such data includes information concerning an individual's racial or ethnic origin, political opinions, religious beliefs, trade union activities, physical or mental health, sexual life, or details of criminal offences. Even though there are circumstances that enable the processing of sensitive personal data without consent, if consent is used as a way to process such data, the Act requires explicit consent.[7]

V Coping with Compliance Challenges

Complying with these stipulations is anything but easy – and requires sophisticated data governance programs that specifically help organizations understand the flow of their data. Data flow is defined as the movement of data from one point to another at a specific level of granularity such as application or organization to transport data.[8]

An understanding of data flow is a crucial component of success as organizations strive to comply with regulations such as GDPR. For example, if a person asserts their "right to be forgotten," an organization would have to deploy staff members to thoroughly comb through reams of data in undocumented systems to make sure that all identifying data attributes have been eliminated – a needle in a haystack proposition if ever there was one.
Consider the following scenario: If a business operates an e-commerce website, a customer's personal information could be captured and stored in the ordering system. That information, however, might traverse to several other systems. For example, the information might be housed in the analytics system, as marketers develop targeting strategies. The information could, in fact, find its way to thousands of databases within the organization. An understanding of data flow could help to identify the 10 or so databases where the personal or sensitive information traveled to – and was eventually stored in.
The challenge is exacerbated by the fact that current undocumented information systems are not set up in a way that makes it easy to locate data. Sure, some information systems and databases will have columns labeled as "social security number" or by some other moniker that will immediately call out the fact that sensitive, personal information lies within that data field. But many systems do not contain such helpful labeling. Even in organizations that have built systems that catalog data in a manner that makes it easy to identify and retrieve personal information, most – if not all – organizations also use older legacy systems or systems that have been inherited through mergers and acquisitions. And, frequently, these systems do not have data properly catalogued or labeled.

VI Conclusion

Conducting manual data flow analysis initiatives in the undocumented systems throughout a company is labor intensive and often not within the grasp of an organization's human resources. To understand how data moves throughout an organization's information systems at any point in time, organizations need to implement an automated, intelligent system that can provide insight into data flow. Such efforts can result in a more sophisticated, in-depth understanding. This, in turn, can make it easier to construct a full inventory of data and better understand how data moves throughout an organization. As such, businesses can confidently move forward knowing that they have an understanding of data flow, which can make it easier to comply with data regulations and to strategically leverage data to improve performance. For more information on how your company can leverage Io-Tahoe's sensitive data discovery product, go to: www.io-tahoe.com.

References

1. Marr, B. 4 Ways Big Data Will Change Every Business. https://www.forbes.com/sites/bernardmarr/2015/09/08/4-ways-big-data-will-change-every-business/2/#12fa1f625c77
2. Browning, L. We sent men to the moon in 1969 on a tiny fraction of the data that's in the average laptop. Business Insider. http://www.businessinsider.com/mind-blowing-growth-and-power-of-big-data-2015-6
3. Greenleaf, Graham. Global Data Privacy Laws: 89 Countries, and Accelerating. Social Science Electronic Publishing, Inc. *SSRN 2000034*
4. EU GDPR. Definitions. http://www.privacy-regulation.eu/en/article-4-definitions-GDPR.htm
5. TaylorWessing. The data protection principles under the General Data Protection Regulation. https://united-kingdom.taylorwessing.com/globaldatahub/article-the-data-protection-principles-under-the-gdpr.html
6. Burton, C. et. al. The Final European Union General Data Protection Regulation. Bloomberg Law. http://www.bna.com/final-european-union-n57982067329/#!
7. The University of Manchester. Processing Sensitive Personal Data. http://www.dataprotection.manchester.ac.uk/whatisdataprotection/sensitivepersonaldata/
8. Ortecha, Ltd. Data Jigsaw. Differences between Data Flows, Lineage, Provenance and Traceability. https://www.datajigsaw.com/articles/dm/DataFlowsLineageAndMore