insideBIGDATA

The GDPR Puzzle

insidebigdata.com/2017/09/07/the-gdpr-puzzle/

Editorial Team

9/7/2017

With the General Data Protection Regulation (GDPR) on the horizon, organizations have finally shifted their focus towards the issue of data privacy. At the risk of incurring unprecedented sanctions—up to 20M EUR or 4% of global revenue—all businesses with an EU footprint must enact a comprehensive strategy for protecting, managing, and remediating personal data by May 2018.

Although many businesses are coming to terms with this reality and putting in place technologies to better manage structured personal data, they're forgetting to ask one very important question: How will GDPR impact the management of *unstructured* data?

For years, file shares, SharePoint, and other unstructured repositories containing employee-created content have flown under the regulatory radar and thus remain largely unmanaged. These repositories have accumulated volumes and volumes of unmanaged ROT (redundant, obsolete, and trivial) data, business records, and, of course, personal data. Few organizations truly understand what data they're holding onto. Still fewer use appropriate processes and technologies to proactively govern it. With the GDPR's imminent arrival, if organizations don't act swiftly, personal data lurking within these unstructured repositories may create massive liability.

New Challenges

Of the regulation's many requirements, a primary area of concern for many organizations is a group of articles that mandate the ability to identify and act upon all data pertaining to a particular data subject. Under the new legislation, EU residents gain several important rights including: the right to be forgotten, the right to restriction of processing, the right to access, and the right to rectification.

The capabilities needed to comply with these requirements are fundamentally similar to those involved in early case assessment for eDiscovery, but research shows organizations are largely ill-equipped to handle such requests when it comes to employee-created data. While organizations may have gotten by without these capabilities under previous regulations, GDPR poses new risks and challenges impossible to ignore. Impacted organizations would be wise to implement technologies enabling the following key processes:

1. **Content Indexing** Most organizations cannot index the actual content held in employee-generated documents and repositories. Full indexing capabilities are necessary to comprehensively search for personal information and begin to act upon it.
2. **Minimizing Personal Data** Unstructured repositories often contain personal data, much of which is held unnecessarily. Minimizing it early will reduce the pressure generated pursuant to GDPR subject requests and therefore mitigate exposure to GDPR fines. To intelligently minimize data ROT, companies need the ability analyze data and holistically understand its value, risk, purpose, and location.
3. **Collection for Subject Access Requests** Organizations need the ability to quickly gather and handle personal data at will. This ultimately requires substantial search, collection and processing capabilities.

Unified Governance

GDPR will fundamentally change the way organizations handle their data. While indexing and managing unstructured repositories provides a quick fix to an urgent problem, for a truly defensible strategy we must add one last piece to our data-driven puzzle: holistic information governance across an enterprise's data silos and functions.

Imagine the following scenario: an employee requests their data be deleted per the GDPR's right to be forgotten. They're not requesting to be deleted from the company's personnel tax files or the holiday party listserv alone; they want their data deleted *everywhere*. So how will your organization respond to such a request? Will you search silo by silo, praying you gather every mention of this employee? Or will you bear the risk of massive fines?

Now imagine requests of this sort pouring in every day. Without proper data management policies in place, you'll be caught in the data deluge. As the number of repositories and functions for which data is used begin to multiply, so does the complexity of managing them all in synergy. Without a holistic system, it's like trying to force together pieces of a puzzle that just don't fit.

A unified approach to information governance is paramount for efficiently searching across disparate data stores for all of an individual's personal data and ensuring each action is reflected across the entire enterprise. With the advent of GDPR, it's time organizations rethink data privacy and take action.

About the Author

Kon Leong is CEO at ZL Technologies. Kon is responsible for managing all aspects of the business, including strategy, finance, sales and marketing. Earlier, Kon was co-founder and president of GigaLabs, a vendor of high speed networking switches. Prior to that, Kon was First Vice President of Mergers and Acquisitions at Deutsche Bank. Kon earned an MBA with Distinction from the Wharton School and received an undergraduate degree in Computer Science from Concordia (Loyola) University, after completing a year at the Indian Institute of Technology.