Understanding the General Data Protection Regulation (GDPR)

Joseph Bartolo, Jr.
jbartolo@valoratech.com

Sandra Serkes
sserkes@valoratech.com

As of May 25th, 2018 the European Union (EU) will alter business requirements for companies that possess personal information pertaining to EU residents. The General Data Protection Regulation (GDPR) applies to any company doing business with customers in the EU, and will have a far reaching impact, beyond just EU borders. While a primary purpose of the GDPR is to harmonize data privacy protection regulations across the various EU member nations, the potential business interruption for organizations around the world is a serious concern.

The applicable fines set forth in the GDPR for failing to comply with regulations are significant. Corporations that handle EU customer data, *regardless of where the company is based*, can face up to EUR20 Million (approximately $22.3 Million U.S. Dollars) in fines, or Four (4%) percent of their total global revenue for the preceding fiscal year, whichever is higher, for GDPR noncompliance[1].
The GDPR data protection protocols must be in place for "Personally Identifiable Information[2]" (PII), of all living EU citizens, regardless of where that information is sent, processed or stored. In addition, the company possessing such PII data must have a process in place to verify and prove that valid protections exist. Corporations are not exempt from the GDPR simply because they don't have offices or process data in the EU. The EU's concept of data privacy differs greatly from the United States', but US corporations doing business with EU citizens will still have to adhere to the strict requirements of the GDPR.

In certain specific circumstances, companies must create a position of "Data Protection Officer" (DPO), whom will address GDPR compliance. Hence, the costs to prepare for compliance will include requirements for trained personnel, and financial investment in technology. Having the means to comply with the stringent requirements of the GDPR is no simple task. Creating an effective compliance strategy will be costly and many companies have not set aside money in their projected annual budget for the funds required to address these concerns, which means they will come from emergency or other contingency planning budgets.

Those corporations who have already begun to address their information management capabilities have an advantage in complying with the GDPR requirements. Many of the key elements of a corporate "Information Governance" (IG) plan are similar to the issues of concern for GDPR compliance. The ability to manage information, and address data governance, corporate risk, and regulatory compliance, are existing concerns for corporations, notwithstanding the GDPR. Existing technology for cybersecurity and "Data Loss Prevention[3]" (DLP) can be utilized to help prepare for the GDPR. Moreover, search and retrieval technology and techniques used for eDiscovery purposes serve

[1] Article 83, 5. "*Regulation (EU) 2016/679 of the European Parliament and of the Council*," April 27, 2016. EU Legal Content
[2] Personally identifiable information (PII), or sensitive personal information (SPI), as used in information security and privacy laws, is information that can be used on its own or with other information to identify, contact, or locate a single person, or to identify an individual in context. Wikipedia
[3] Data loss prevention (DLP) is a strategy for making sure that end users do not send sensitive or critical information outside the corporate network. - TechTarget

as a means to assist in managing information. In addition "Knowledge Management" (KM) practices enhance a corporation's ability to comply with the GDPR. The illustration below from Susan Bennett, of Sibenco Legal & Advisory provides useful insight into aspects of information governance, many of which help address the specific needs of the GDPR.

(Source: "What is Information Governance and How Does it Differ from Data Governance", by Susan Bennett, of Sibenco Legal Advisory, April 2017, Information Governance vs. Data Governance)

Sensitive information handling is a challenge that pertains to both IG and GDPR compliance. Restrictions imposed on the transfer of PII by the GDPR can be addressed by the use of technology. Identification of sensitive content within a business record, and the ability to redact portions of content, can impact whether that specific file is transferrable under the rules set forth by the GDPR.

Addressing GDPR Compliance

Since the GDPR specifically requires the ability to prove data protections are in place, documentation of existing privacy safeguards is essential. All documentation and processes must clearly address issues such as: where is the data; what type of data exists; who has access to the data; what is in the data content; how is data stored; how is data transferred; and how is newly created data incorporated? Without answers to these questions, GDPR compliance is impossible.

Below are suggestions for IG best practices which can be specifically implemented to address the requirements of the GDPR:

Data Mapping
Data mapping is the process of identifying what type of information lives in which systems throughout the organization. When the data map for an organization is incomplete or inadequate, a discussion with I.T. stakeholders should take place to update it. Collaboration between I.T., management, and the corporate legal department in order to create a comprehensive data management plan is a vital step toward GDPR compliance. Corporate data stored by third-party providers, including cloud services vendors, or data archival companies, also requires attention. Any data in possession of third-party providers is still subject to GDPR regulations applicable to the corporation, including information retained by outside counsel.

Understanding File Contents.
Knowing *where* data resides is only part of the equation. A corporation must also know *what* the data is and contains. For example, are the files legally binding in nature, such as contracts and agreements? Do the files contain any sensitive data, such as PII or PHI?

Consent.
A key requirement of the GDPR is the need to obtain *specific consent* from an individual before obtaining, storing or utilizing their personal data. The corporation must provide a clear affirmative action or statement providing permission to process the individual's data. In addition, the GDPR establishes that the individual has a "Right to be Forgotten," and can request their personal information be explicitly removed from use. Without another legal reason to process an individual's information, the corporation must respect a request to delete data without undue delay.

Information Request.
Similarly, an individual has a right to request access to the personal information being gathered and stored about them. The individual may request information from a company about any of their personal data, including: who has access to their information, how the data is accessed; where it is being accessed; and the purpose for which it is being accessed. Furthermore, an individual can also seek corrections about their personal data, if the EU resident feels the information is inaccurate. The individual may object to the use of their data for profiling by the corporation.

Retention Schedules.
Enforcing corporate document retention schedules, while also maintaining proper litigation hold protocols can be quite challenging. There are inherent risks associated with maintaining information when there is no legal obligation to retain possession of that data. An effective means of dispensing with specific information that is outside of an applicable document retention schedule is an important component for both IG and GDPR compliance.

Security Breaches.
An overarching component of the GDPR is the need to provide cybersecurity protections to prevent data breaches, as well as express provisions regarding notifications of data breaches to both the supervisory authority and to individuals whose information has been exposed. Hence, corporations must not only be aware when a breach has *occurred* but also must have a means to notify those impacted by the breach of *what specifically* was exposed.

Data Transfer.
The GDPR places explicit restrictions on transfers of personal information and organizations must have an enforceable plan to prevent such unauthorized data transfer, particularly regarding transfers to locations outside of the EU. Whether a data transfer is permissible under the rules of the GDPR requires answers to a series of queries about the content of the information. When PII, or otherwise sensitive information, exists in the data, additional restrictions may be applied, possibly revoking permission for the transfer of that information. An entire file might be improper to transfer under certain circumstances, thereby prohibiting access for persons outside of the EU to view such information. In other instances, a portion of the content of a file might block the permissible transfer, however if actions are taken to *redact* the specific content in question, the remainder of the file might be permissible for a data transfer.

How Auto-Classification Assists with GDPR Compliance

It is clear that properly managing all data in a corporation's possession to comply with GDPR regulations is an onerous task. The GDPR requirements thus create an increased reliance upon automation in order to properly manage the lifecycle of corporate information.

Corporate best practices for IG, KM, E-Discovery, compliance and cybersecurity, all provide guidance for the use of technology to help address GDPR regulations. One particularly promising technology for corporations struggling with GDPR mandates is referred to as "Auto-Classification."

Auto-Classification Software data mines information at the content level, and then categorizes files based on the information's substance. This technology is already being utilized by many corporations as part of their IG, KM or eDiscovery strategies. Auto-Classification's ability to group information by category or by specific characteristics proves useful for GDPR compliance. Similarly, Auto-Classification's ability to detect the presence of PII and other sensitive content will likely become a best practice when it comes establishing GDPR protections.

Auto-Classification software uses both pattern-matching algorithms as well as artificial intelligence to detect file contents and attributes such as: personal information; authorship and origin; type or format of document; and expected retention period for different purposes. In addition, Auto-Classification technologies follow a set of customized rules regarding file disposition. For example, a rules-based Auto-Classification system will enforce a specific document's retention schedule, and then place the file into the proper folder taxonomy structure. Auto-Classification technology specifically meets the GDPR requirements to have a system in place that can detect what information it has, where it lives and how it will be handled under differing circumstances.

With a proper Rules Engine, sensitive information is protected via individual security level restrictions, including limitations based on the geographical location of the user attempting access. Rules are further used to block improper information transfer across country or network system borders. Finally, rules are used to trigger certain events, such as an expiration date associated with certain data which would make such information eligible for deletion.

Conclusion – Advantages of Using Automation for GDPR Compliance

While compliance with GDPR regulations will be no small task for most enterprises, the use of automation makes the task more manageable. There is still time to prepare for the GDPR regulations, and avoid the imposition of fines by utilizing technologies specifically designed for content-level file analysis. Enterprises which are proactive in automating their IG strategies are in a better position to comply with the GDPR than others.

Return on investment is often a key metric required to approve corporate expenditure of funds. While companies may have been previously reticent about investing in IG technology, the GDPR requirements serve as a stark turning point to that strategy. The potential for business interruption caused by the GDPR, not to mention its stringent fines for non-compliance, justify return on investment calculations several times over. Furthermore, the benefits derived from improved information management techniques assist not only GDPR compliance, but also related areas of corporate efficiency, eDiscovery, Records and Information Management, Information Governance and Knowledge Management.

Auto-Classification technology is proving itself to be a vital tool to meet GDPR compliance. With its ability to determine what specific information lives where, along with indicators for how to classify and redact sensitive content, Auto-Classification is likely to benefit many organizations struggling with large volumes of data and short timeline for solution implementation.
Lack of preparedness for GDPR is an alarming concern. According to a Symantec survey in 2016, "91% (Ninety-One Percent) of 900 business IT decision makers polled in the U.K., France and Germany have serious concerns about their ability to be compliant by May of 2018[4]. The attention paid to the looming threat from the GDPR's effective date of May 25, 2018, will only grow as the date approaches.

[4] "Business Unprepared for GDPR, Survey Show", Computer Weekly, October 18, 2016, Business Unprepared for GDPR)