Portfolio Media. Inc. | 111 West 19th Street, 5th Floor | New York, NY 10011 | www.law360.com
Phone: +1 646 783 7100 | Fax: +1 646 783 7161 | customerservice@law360.com

Data Inventory: The Critical 1st Step In Data Security

By **David Manek, Navigant Consulting Inc., and Bruce Radke and Michael Waters, Vedder Price PC**

Law360, New York (April 28, 2017, 11:16 AM EDT) --

David Manek

As high-profile data system breaches continue to make headlines, businesses and public agencies are focusing with increasing intensity on data privacy and security concerns. Upgrading information security systems, developing incident response plans, conducting penetration testing and carrying out other information governance, privacy and data management activities are time-consuming, expensive and strain limited IT resources. Yet they are essential in today's information-driven organizations, where the theft, loss or unauthorized exposure of sensitive information can inflict devastating consequences in terms of reputational risk, increased compliance and regulatory costs, and direct financial losses.

Bruce Radke

Unfortunately, the effectiveness of data privacy and security initiatives is compromised if the organization has not taken one critical first step: developing and maintaining an accurate and comprehensive data map. In addition, as regulatory agencies increase their scrutiny of data security issues, the ability to identify and locate sensitive data quickly becomes even more critical. Effective data mapping has become an essential compliance and regulatory response tool.

Michael Waters

Data Mapping Basics

Fundamentally a data map shows where certain types of data are located within a system and how that data can be accessed. A data map typically provides both a data inventory and flow diagrams that depict how data moves through the organization. These data flow charts can be diagrams, other graphic representations or spreadsheets. Graphic representations are useful for visualizing data flows and connections, but spreadsheet-based or matrix-based data inventories are often more practical for end users. The terms "data map" and "data inventory" are used interchangeably throughout this article; however, one should envision a matrix-based data inventory as the ultimate goal.

Data Inventory Drivers — Why It Matters

The specific features and characteristics of the data inventory and types of data to be mapped must be aligned to each organization's business needs and customized to its specific legal and regulatory

environment. There are typically four drivers that prompt the creation of a data map:

1. Legal, Regulatory and Privacy Compliance

A data map identifies specific types of information that must be tracked and reported under the requirements of various laws, regulations or professional standards. (For example, the Gramm-Leach-Bliley Act, Health Insurance Portability and Accountability Act and data requirements imposed by organizations such as the Financial Industry Regulatory Authority or the International Organization for Standardization.) State governments also can require businesses that retain personal information of state residents to maintain a written information security program which includes certain minimum administrative technical and physical safeguards.

The General Data Protection Regulation, whereby the European Union intends to strengthen protection of data collected from EU data subjects, will be enforced in May of 2018. It extends to organizations collecting personal information from EU data subjects regardless of where the organization is headquartered or hosting the data. U.S.-based companies collecting personal information from EU citizens may be forced to demonstrate compliance with GDPR. Penalties for noncompliance can reach €20 million or four percent of total revenue, whichever is higher. Maintaining a data inventory is a fundamental element of demonstrating compliance with GDPR.

2. Intellectual Property Risk

A data inventory is essential for identifying and tracking financial data, intellectual property and trade secrets as well as other sensitive data such as personally identifiable information, protected health information or nonpublic personal information. In addition to helping organizations apply required access controls and security safeguards, a data map also demonstrates compliance with those controls, and shows that adequate safeguards are in place.

3. E-Discovery

A data map simplifies responses to e-discovery requests. Legal departments can use the data inventory as a menu to select which systems they want collected and preserved in response to pending or potential litigation. In public sector organizations, a data map is helpful in replying to Freedom of Information Act requests.

4. Data Management, Retention and Disposition Policies

A data map is the first step in establishing effective data management policies that prioritize what data should be retained, the resources needed to safeguard that information and the location of information to be disposed of when their retention periods expire.

Breach Response — When a Data Map Is Vital

While the compliance and business drivers discussed thus far demonstrate the benefits of maintaining data maps, the most crucial need relates to breaches that lead to the exposure of sensitive information, outright theft of proprietary data, or a ransomware attack in which data is held hostage. Was the data accessed by an external attacker or leaked from an internal source? Or, did the leak occur when a third party with authorized access failed to protect the data? What steps must be taken immediately to identify the source of the breach, contain the damage and prevent exposure of additional information?

All of these questions must be answered quickly, not only to minimize the damage but to comply with various notification requirements. Many state and federal data breach notification laws require notice to affected individuals and regulators within a short amount of time. Without an accurate and up-to-date data map, your organization has three costly and complex options to comply with these requirements:

1. Install an enterprise software solution that searches your entire infrastructure by document name or other characteristics. Generally, they are unable to index “loose devices” such as mobile devices, flash drives or data stored by third parties. The time required to identify, select and install such software is another significant issue when timely response is critical.

2. If you have some idea what servers or devices were compromised, you can hire a consulting firm to create an image of the device, process and scan the information to look for the specific document or other information. While this approach is faster than the first option, it is still time-consuming and, like enterprise software solutions, it can still be unable to address all sources or devices.

3. The third option is to conduct a manual review or “treasure hunt.” This involves combing through the data of the entire organization to identify who had access to the exposed documents and where the information was stored. When deadlines are tight and potential penalties are mounting, such a process is not only laborious and costly, it is prone to potential errors and consumes an inordinate amount of resources.

Obviously, knowing what systems contain PII, PHI, nonpublic personal information or other sensitive data can dramatically accelerate incident investigation and containment. A comprehensive and accurate data map makes it possible to respond to a breach quickly, efficiently and systematically.

In addition, the process of developing the data map minimizes the amount of sensitive information that is stored, reducing both the risk of a breach occurring and the extent of the damage if one does occur.

Building a Data Inventory in Four Steps

1. Define the scope and other business requirements. Is the map enterprise-wide or focused only on certain departments or functions? What data repositories are included and what data elements within those repositories will be collected and maintained?

2. Develop the design and framework. Identify the users of the data map and get their input and commitment. Key stakeholders will likely include legal, compliance and risk management functions, as well as IT.

3. Populate the data map. Identify what information is stored and where that information resides. Generally this is accomplished using in-person interviews, online surveys or both.

4. Develop a maintenance process. Most organizations are adding new systems and new business processes routinely as they acquire new business or development information. This makes it essential that the data map be regularly updated and improved as well. Quarterly maintenance of the data map is a good minimum, along with an annual full-scale audit and update.

Getting Started

The importance of achieving clarity on scope and other requirements at the outset cannot be overstated.

For organizations new to developing a data inventory, begin with a pilot program focused on high-risk records or specific types of information such as PII or PHI. An alternative is to conduct the pilot program within a single business unit only, and expand to other segments in later phases.

Once the overall scope of the data map is defined, consider which data repositories will be covered. Servers, workstations and databases only? Or, also online sharing sites, laptops, smartphones and other mobile devices? Will the focus be limited to electronic records and data only, or will it cover both electronic and hard copy records?

Best Practices and Next Steps

In terms of overall best practices, a National Institute of Standards and Technology publication that is particularly on point regarding data mapping is, "Guide to Protecting the Confidentiality of Personally Identifiable Information (PII): Recommendations of the National Institute of Standards and Technology,"[1] published in 2010.

Regarding the actual content of the data map, the amount of information contained can range from the bare minimum — such as designating a subject matter expert for each affected system and a description of the nature of the data it contains — to very mature data maps that address a wide range of parameters, including:

- Nature of the data
- System format
- Retention policy
- Backup procedures and frequency
- Identification of all PII in the system
- Confirmation that the use, collection and retention of PII is limited to that which is strictly necessary
- Categorization of all PII by its confidentiality impact and value to the organization
- Application of appropriate safeguards, and
- Annual review policy and personnel.

Regardless of the format and structure that is eventually chosen, the most important point is to start the process of scoping and developing the data map. Have a conversation with your IT team to identify what types of data inventory might already be in place.

If your organization has not yet begun this process, the time to begin is now. If you already have data mapping in place or under development, it is important to regularly review and update your data map to ensure it remains current with your entity's ever-changing data environment. High-profile data system breaches continue to make headlines and regulatory agencies focus intently on privacy and data security issues. An accurate data map is an indispensable tool in identifying, locating and managing sensitive personal and organizational data.

David Manek is a director at Navigant Consulting Inc. in Chicago and national leader of the firm's technology solutions sructured data practice.

Bruce A. Radke and Michael J. Waters are shareholders in the Chicago office of Vedder Price PC and co-chair the firm's privacy, cybersecurity and media practice group.

[1] http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-122.pdf

---