COSMOS
Navigate the Compliance Universe

CEP Magazine - August 2021
Implement information governance to successfully manage ungoverned data

By Dean Gonsowski, JD

Dean Gonsowski *(dean.gonsowski@activenav.com) is the Chief Revenue Officer of ActiveNav in Reston, Virginia, USA.*

- twitter.com/dean_gonsowski
- linkedin.com/in/dean-gonsowski

Are we currently experiencing a data *explosion,* as many pundits claim? If the general volume of data is truly exploding, it is the longest and slowest explosion that we have ever seen. Despite the poor analogy, one thing is very clear: Organizations are struggling to control this growth of data, leaving them subject to significant compliance, privacy, and security risks.

As much as 90% of a company's data environment can be made up of unstructured data.[1] Unstructured data doesn't fit into predefined data models and is therefore inherently full of risk because it's hard to search, analyze, and control. No one knows what's hiding in this mess of data, which is often created by users and usually includes sensitive information that hackers look for, such as personally identifiable information (PII), Social Security numbers, and credit card information. Unstructured data is typically ungoverned, meaning that companies are sitting on huge amounts of risk, often unaware. It is a target for hackers because many companies do not know how much of it they have, where it's stored, or who owns it.

Meanwhile, the looming threat of data breaches continues to grow. Alarmingly, in 2020, despite huge investments in cybersecurity technologies, the number of reported data breaches was the highest on record. On average, the cost per record breached was $146. That cost went up to $150 if the record contained PII, which is still the most compromised type of record with PII being breached in 80% of cases.[2] Other significant costs stemming from security incidents (that should not be overlooked) include reputational damage and loss of productivity.

Companies, as they should, have been investing heavily in cybersecurity technologies like endpoint protection, identity access management, and third-party risk management to improve their security posture. Despite these prophylactic measures, data breaches continue to occur at an alarming rate, even at some of the most sophisticated companies in the world. For example, in January 2020, a Microsoft customer support database containing 280 million records was left unprotected, exposing email addresses and support case details.[3] If Microsoft, with all its resources, gets breached, it's clear that it's now a matter of *when* you will get breached, not *if*, which is a sobering thought, to say the least.

Information governance can complement cybersecurity efforts to help better manage information prior to and when the inevitable occurs. Information governance is the suite of activities and technologies that organizations can employ to maximize the value of their information while minimizing associated risks and costs. When done effectively, information governance reduces the amount of data under management and provides visibility into what data was breached if that occurs—or rather, *when* that occurs.

Information governance vs. information management

Traditional approaches to governing information are no longer working. Different functional areas within companies have unique objectives and value information differently. These functional areas tend to establish policies and practices that satisfy their lines of sight but often lead to conflicting silos between sales, marketing, finance, operations, etc.

Where information management focuses on *how* information flows through an enterprise, information governance asks *why* we have the information in the first place. We need to govern information in a way that works for everyone, no matter their role or department.

In our experience, the best and most efficient way to do this is to link information creation, use, and disposition to business objectives. Organizations must ask these questions:

- **What** are our organizational objectives (business, legal, and regulatory)?
- **What** information is needed to achieve our objectives?
- **How long** is that information useful?
- While it is useful, **how** does it need to be organized (access, security, privacy)?
- **What** do we do when information is no longer useful?

You can't get value from data you can't access

Too much data makes everything less efficient, yet organizations still tend to hoard data "just in case." Other problems created by collecting and storing too much data include increased regulatory burdens, storage costs, and security risks.

Data remediation addresses these problems. Data remediation is a process directed at bringing order to information, and that doesn't just mean deletion. The goal of remediation is to help ensure that companies retain only valuable information—i.e., information that is necessary to meet the organization's business, legal, or regulatory objectives and obligations. Remediation helps to ensure that information no longer useful to the organization is deleted in a defensible manner.

A framework for remediation

Data discovery: The first step is to gain visibility into your data estate. You can't protect and manage what you don't know you have. Enterprise software solutions such as "file analysis" can help you gain insight into all your sources of data, what's being collected, and where it's being stored. By discovering and mapping your data, you'll be able to find out what is actually there. After all, most business processes and collaboration activities are not captured with a top-down approach.

Data classification: The goal of data classification is to make content easier to govern and access. Data classification automates the ability to audit policies, request responses, and protect data. Leverage taxonomies, similarity clusters, and predictive learning classifiers to make it easier to understand the sensitivity and value of the data you have. Data can be classified in several different ways depending on business needs, but common categories include:

- Content

- Security
- Confidentiality
- Privacy
- Regulatory
- Legal hold

Metadata management: Metadata is simply data about data. Creating key metadata helps with the ongoing governance of content. Metadata helps determine how content will be discovered. So make use of existing metadata as much as possible. From there, develop rules for identifying and extracting additional metadata that can help with further data classification.

Data remediation: Once you have discovered and classified your content, you then need to decide what you want to do with that data. You might decide to keep, tag, migrate, quarantine, or delete given pieces of information depending on factors such as age or redundancy. No matter what you decide, you must ensure there is a defensible audit trail upon disposition.

Ongoing governance: Information is constantly flowing in and out of an organization, meaning that governance is not a "one and done" type of project. There needs to be a process, one that's embedded into the culture of an organization. Establish a program to continuously identify and delete content that no longer provides value to the business, and find ways to get the end users excited about participating in your governance program, increasing chances of long-term success.

Where and when to start?

Start with the proverbial low-hanging fruit. Take bite-sized chunks of data—maybe that's a single data repository or business unit to begin with—and then build from there. Decide your success metrics ahead of time, whether that's efficiency gains, risk mitigation, or cost savings. If you're able to demonstrate your success, it will be easier to gain stakeholder support for larger projects.

Information governance provides insight into workflows and business processes, which can then be used to optimize workflows and processes, leading to even more efficiency and risk mitigation. By beginning small, measuring success, and communicating often, effective information governance can be achieved.

And remember, the data breach is inevitable. Additional privacy regulations will be signed into law. Security threats will evolve. It takes time to plan out an information governance program that will be effective for your specific organization, gain stakeholder approvals, find and deploy the tools needed to execute the initiative, and train the business on proper governance practices. Don't wait to start the process as a reaction to an event. Be proactive about establishing and implementing a robust data governance program so that when an event occurs, your organization is prepared.

As the saying goes, "The best time to plant a tree was 20 years ago. The second-best time is now."

Take a coordinated approach

Effective data and information governance is possible, but it needs to work alongside cybersecurity to minimize the impacts of data breaches. It also requires a coordinated approach among interested stakeholders. When implementing an information governance program, start by asking yourself:

- Do we understand what data we have, where it is, and how many copies we have of it?
- Are we safeguarding our most valuable and sensitive data and prioritizing data protection and stewardship?
- Are we positioned to confidently respond to our organization's business, legal, or regulatory objectives and obligations?
- Can we ensure that information that is no longer useful to the organization is deleted in a defensible manner?

Satisfactory resolution of these questions will help to mitigate the risks associated with a security incident.

About the author

As a former litigator, general counsel, and associate general counsel, **Dean Gonsowski** is an industry-recognized evangelist, thought leader, and speaker who generates compelling content, delivers it dynamically, and builds communities. Dean has a juris doctorate from the University of San Diego School of Law and a bachelor of science degree from the University of California, Santa Barbara.

Takeaways

- Unstructured data is the hidden threat in digital businesses and one that cannot be ignored.
- Information governance and cybersecurity need to work together to minimize the impacts of security incidents and data breaches.
- Data remediation, which doesn't just mean deletion, is key to addressing numerous problems caused by storing and collecting too much data.
- Create a community of information governance advocates and stakeholders to support information governance programs and efforts.
- Information governance needs to be embedded within a company culture, which means you need to ensure the conversation is happening at an executive level too.

1 Dwight Davis, "AI Unleashes the Power of Unstructured Data," CIO, July 9, 2019, https://bit.ly/3pc1xUI.
2 IBM Security, *Cost of a Data Breach Report 2020*, July 2020, https://ibm.co/2SLhUf9.
3 Catalin Cimpanu, "Microsoft discloses security breach of customer support database," ZDNet, January 22, 2020, https://zd.net/3fZyakC.