

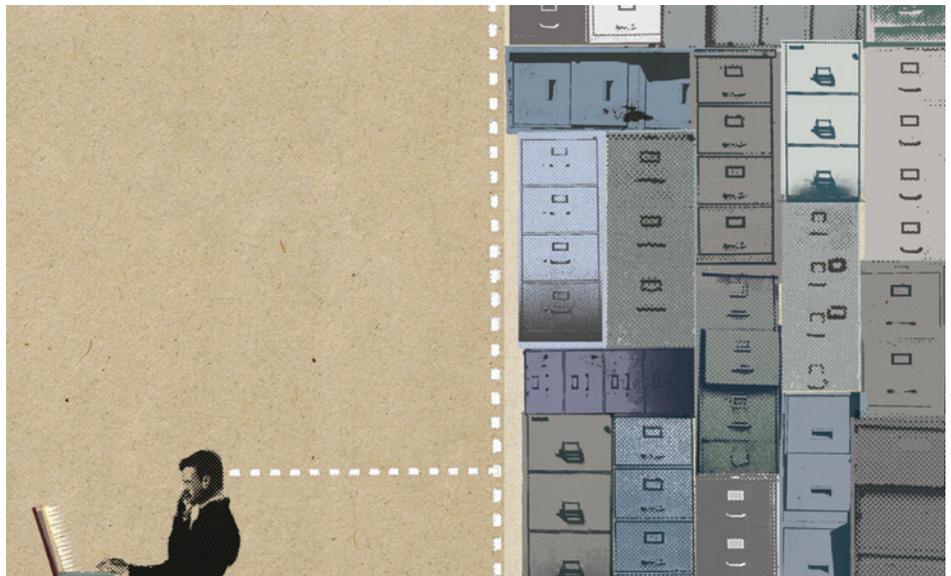
HARNESSING BIG DATA

Organizations can derive value from their information if they know where to start.

BY JOE GARBER

YOU ALREADY KNOW sound information governance practices can protect your organization from risk that can manifest itself in the form of fines, sanctions, and lost shareholder value. These benefits are real and have been proven by many enterprises worldwide. There is another aspect of information governance that has, to date, been discussed a little less in the boardroom but is equally important: leveraging these same technologies to derive value. This is largely because corporate leaders acknowledge the importance of leveraging their data to turn information into actionable insights that positively affect the bottom line, yet very few know where to start.

One option to get started is to jump right into analytics. To be sure, there are many excellent analytics technologies in the market that can help organizations separate signal from noise, and thus mine data for value. But there are downsides to such a strategy as well, if not approached correctly. The most common is casting the net too narrowly. Analytics applied to only a portion of an organization's data often only delivers a partial view and may miss some important context hiding in the rest of the information. Alternatively, analytics can be applied wholesale to all enterprise data, which routinely bogs down the system with such large volumes.



Another emerging option is to proactively take control of your data with information governance as a foundational aspect of Big Data exploration, which can streamline downstream analytics. This approach allows organizations to take into account 100 percent of their data in their analysis (regardless of source), but by valuing, categorizing and applying policy to data in advance, organizations can quickly apply analytics to the most relevant subset of data to drive superior business outcomes in an efficient and timely manner. In order to do so,

companies must have the right policies, procedures and processes in place to manage the information.

THE DATA-DRIVEN ORGANIZATION

Enterprise data is growing, and that has created an information governance problem for all organizations. Billions of documents, emails, text messages, IM's and other content are created and moving throughout hundreds—if not thousands—of systems that are accessed by an incalculable amount of users. With this reality, it's no surprise that organizations



LEGAL OPERATIONS | INFO GOV AND BIG DATA

are struggling to access, make sense of, classify and secure their data.

The most advanced organizations are proactive about how they manage and secure their data, but in today's environment, the big question is, what's worth keeping and what should be jettisoned? A large percentage of stored and retained information is redundant, obsolete, and trivial, and weighs down storage and serves as an impediment to analytics. Other information is "dark data" (data that no one is particularly aware of, which presents significant potential risk and untapped benefits) living in enterprise systems, essentially using up storage incurring management costs without offering any value.

"ORGANIZATIONS ARE STRUGGLING TO ACCESS, MAKE SENSE OF, CLASSIFY AND SECURE THEIR DATA."

The ability to access, understand, value and categorize enterprise information in an automated fashion is the Holy Grail of these projects. Accessing and understanding enterprise information is difficult in itself. Connectors to data repositories must be in place to bridge data silos, and only when you can truly understand information can you value and categorize it.

This seems simple on the surface, but it's much more difficult when you dig one level deeper. Take unstructured information like email and files that aren't nicely organized into rows and columns. Understanding the "aboutness" and relative importance of a given object, without manually reading each one at great expense, requires context beyond simple keywords. For example, an email talking about a bridge cannot simply be categorized as being about a structure over water or a canyon. What if the file is talking about dental work

or a card game - two other forms of bridge? Governing information accurately and in an automated fashion thus requires technology that provides deep information insight based on its context.

Fortunately, information governance technologies have evolved over time. Many technologies can deliver deep insight based on context. More importantly, no longer do organizations need to deploy a "boil the ocean" approach to governance that requires a large technology deployment over many years. The time-to-value is simply too long. Instead, modular portfolios—including file analytics, information archiving, records management and

e-discovery technology—are available to unlock data's potential and meet business objectives.

TAKING ON BIG DATA

While there isn't necessarily a one-size-fits-all solution or approach for every business, there are common tips to consider when combating and controlling Big Data sprawl.

1. *Understand Your Data:* As a first step in developing your information governance strategy, consider using a file analysis tool to provide a snapshot of your information (i.e. data types, sizes, uses, etc.). This will help you access, understand, and classify information across the enterprise, and ultimately allow you to defensibly dispose of non-business data while identifying and migrating critical business information.

2. *Manage Your Data:* Managing, capturing and gaining deeper visibility to all necessary business

information is critical to success. Therefore, it must be managed effectively in order to mitigate risk and maximize operational efficiency while maintaining regulatory compliance. Consider employing an archiving or electronic document and records management system (EDRMS) solution to leverage all information and utilize proven technologies, such as analytics.

3. *Guide Your Data:* A successful information governance strategy involves continuous understanding and proactive control of your information, from creation to disposition. Focus on an end-to-end approach in which all solutions, such as archiving, EDRMS/records management, e-discovery, and supervision can operate together in a unified manner to deliver even greater value. The alternative is to utilize multiple point solutions, or legacy systems, which cannot interoperate with each other and provide limited capabilities to overcome Big Data challenges. The goal is to implement a holistic framework to mitigate risk, increase efficiency, and drive profitability.

Selecting a solutions provider with the flexibility, agility and experience to manage evolving challenges and requirements is the number one priority during the process, but organizations need to be smarter about their own business and must have a concrete understanding of their data in order to gain benefits.

The most successful businesses are data-driven organizations that harness 100 percent of their data and information governance serves as the foundation in today's reality, where data and information has become an organization's most valuable asset.

Joe Garber is vice president of marketing, enterprise information management & governance, Hewlett-Packard.