![IDC - Analyze the Future]

Mining for Insight: Rediscovering the Data Archive

Sponsored by: Iron Mountain

Laura DuBois   Sean Pike
June 2015

EXECUTIVE SUMMARY

In the past, the main drivers for data archiving centered on legal/compliance aspects and business continuity, including eDiscovery, records retention, audits, and disaster recovery. But as more companies take advantage of Big Data and analytics – a fundamental component of what IDC refers to as the 3rd Platform of computing – to gain competitive advantage in their markets, archiving has taken on a new role. More companies are now interested in using their data archives as a source for analytics and are finding new ways to monetize archived information.

> The *3rd Platform* is IDC's term for the current generation of computing built on a foundation of cloud, mobile, social, and Big Data technologies.

A recent data archiving study, developed by IDC and sponsored by Iron Mountain, found that today's most successful organizations are using data archiving for more than traditional legal, compliance, and regulatory purposes. They are mining data archives to gain key insights that improve business outcomes across a range of areas, including improving customer service and enhancing revenue streams. The top 15% of companies surveyed generated more than $10 million in additional revenue as a direct result of mining their data archives over the past year.

Though data archives can add significant business value, much of this value remains untapped because many companies do not have the structure in place to fully monetize their archives. 57% of companies maintain six or more archives and are archiving a broad range of information, including files, email, security logs, and call center records. However, only 12% of companies have a uniform process to identify what gets archived for all categories of data, meaning 88% of companies may face difficulties identifying and accessing important information when they need it.

Moreover, there are major organizational disconnects related to archive accountability and lines of authority, particularly between IT and the legal and compliance departments. Legal and compliance are heavy users of data archives for critical purposes such as responding to regulatory audits, accessing data for eDiscovery and early case assessment, and finding information to substantiate legal positions. The same groups, however, are critical of their company's archiving strategies, see less business value in the archives than IT and lines of business (LOBs), and tend to be dissatisfied with IT's overall management of the archives.

There are many challenges preventing companies from realizing the full business value of archived data. Archiving schemes require attention to structured and semistructured data, as well as unstructured content. Often, organizations have different levels of maturity in each area. To become successful, enterprises must develop competencies in managing all data types.

About This Study

This global study, conducted by IDC and sponsored by Iron Mountain, is based on a survey of 1,011 members of senior and executive management involved in data archiving in organizations having more than 500 employees across a broad range of industries. The survey focuses on eight countries: the United States, Canada, the United Kingdom, France, Spain, the Netherlands, Germany, and Australia. Respondents were from IT departments and legal/compliance departments, as well as lines of business. They were asked about their current archiving strategies and processes, the business value of their data archives, and how satisfied they are with their access to archived information for a variety of business uses.

For the purposes of this study, data archive is defined as the long-term retention of fixed-content data. Data archives include data that is determined by an organization to be important and possibly necessary for future reference and can include active and inactive data. The data may be retained to meet legal or regulatory compliance requirements, to satisfy eDiscovery requests, or for future business reuse.

For a more detailed description of the survey methodology, see the Appendix.

SITUATION OVERVIEW

Archiving Is Chaos

In today's information-driven economy, success or failure often comes down to how effectively businesses manage data (i.e., how much subject data is available, whether the organization can find the most relevant information and how quickly the organization can act on that information). In this environment, access to archived data has become a significant contributor to business success. While this has generated renewed interest in data retention, few companies have the structure in place to fully manage and monetize their data archives.
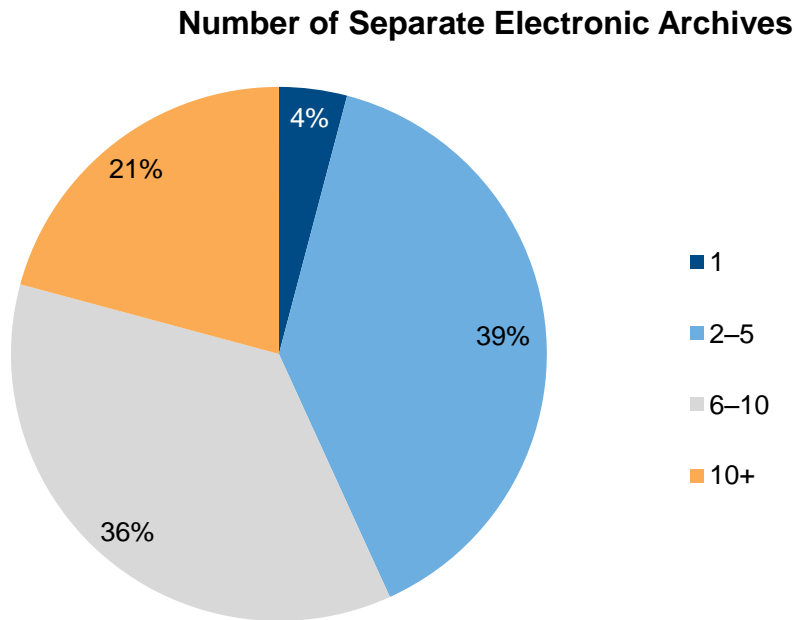
Companies Archive a Range of Data in Multiple Electronic Archives

Most midsize to large businesses (57%) maintain six or more electronic archives that contain both structured data and unstructured data(see Figure 1). The types of data go well beyond email and text files and include the following:

- 68% of companies archive application, database, and other structured files.

- 64% of companies archive images.

- 58% of companies archive video (surveillance or other).

- 51% of companies archive Web 2.0 content (corporate blogs, corporate wikis, RSS, etc.).

- 49% of companies archive voice data (call center records, voicemail, etc.).

- 45% of companies archive instant messages (Lync, etc.).

FIGURE 1

Most Businesses Maintain More than Six Archives

Number of Separate Electronic Archives

Legend:
- 1
- 2–5
- 6–10
- 10+

Pie chart values: 4%, 39%, 36%, 21%

n=1,011

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study,* April 2015

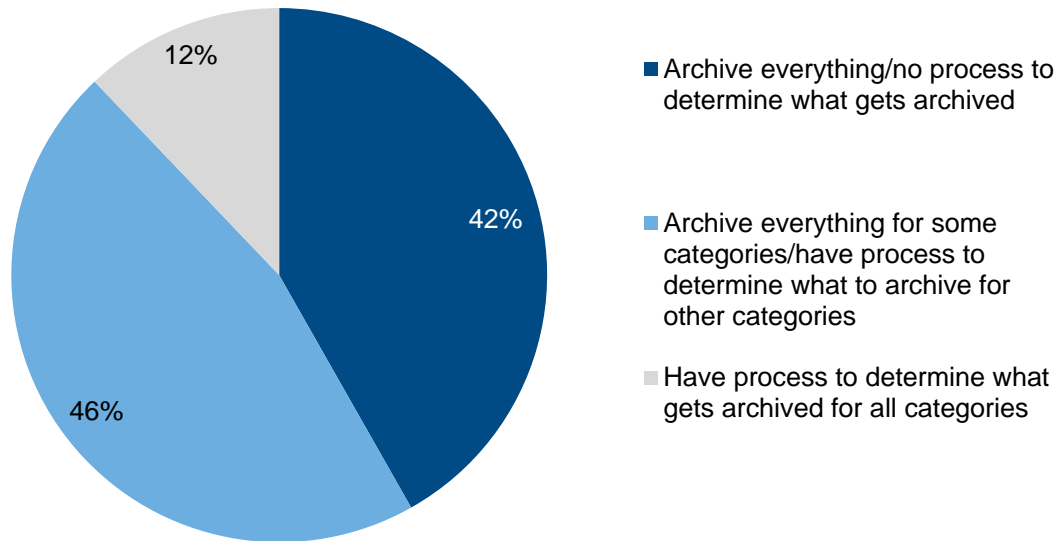## *The Inevitable Progression from "Data Lakes" to "Data Swamps"*

Despite the amount of archiving that companies perform, the vast majority (88%) lack a consistent approach to archiving across all data types. More than 40% of organizations simply archive everything, lacking the wherewithal to define and sift through data to determine what *is* important and what *is not* important (see Figure 2). Much has been written about companies' desire to maintain data lakes — easily accessible and inexpensive repositories designed to retain all data attributes — from which they can pull necessary information to make critical business decisions. However, this archive-everything approach leads to companies creating murky data swamps — holistic data repositories that can make information hard to find because the information is unclassified and treated inconsistently.

> **More than 40% of organizations simply archive everything to avoid spending time up front sifting through data.**

FIGURE 2

88% of Companies Lack a Consistent Approach to Archiving Across All Data Types

Overall Approach to Data Archiving

n = 1,011

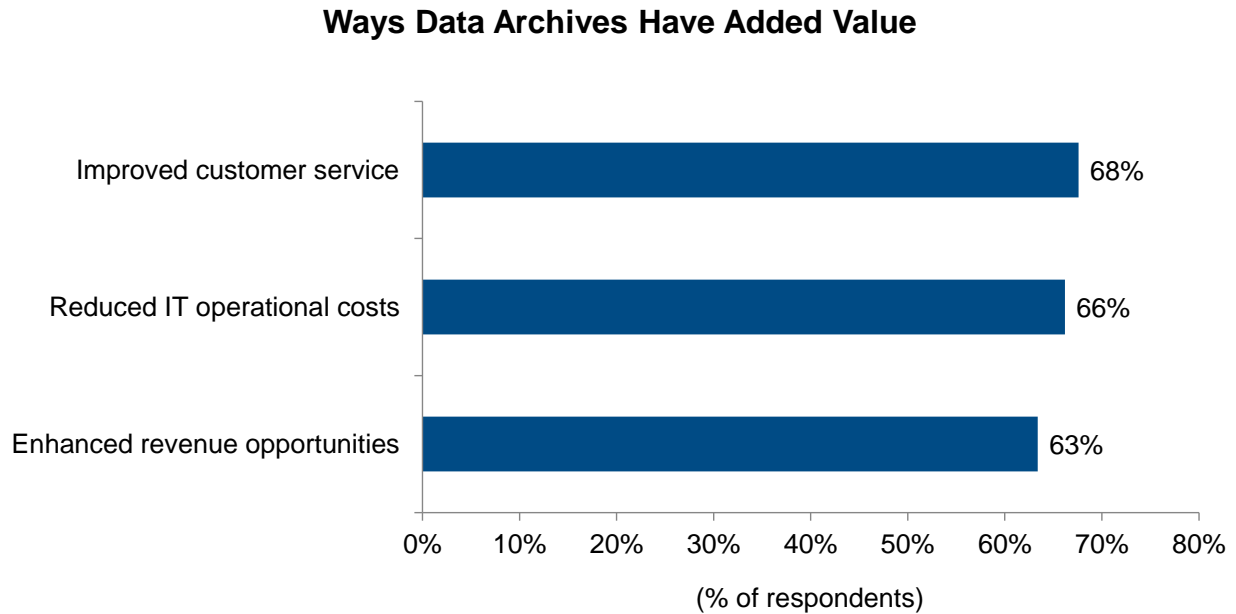Source: IDC and Iron Mountain's *Archiving Thought Leadership Study*, April 2015

To add further complication, companies are using a range of different archiving approaches to store unstructured data. Examples of unstructured archives include cloud services (used by 54% of companies), backup applications to retain long-term data on tape (51% of companies), and standalone archiving software/tools used for specific applications like email, file, and database archiving(45% of companies). About half of the companies (53%) are also using a structured data warehouse.

There Is Significant (Untapped) Value in the Archive

Historically, the primary value proposition of archiving data was for use in eDiscovery, records retention, audits, and business continuity. These remain important uses of the archive, but today, successful organizations are finding new ways to tap archives to add business value. This creates two very distinct opportunities: One is to leverage the archive for IT efficiency and risk reduction, and the other is to use the archive to gain business insight. Our study shows that most companies are finding value in each of these areas. Two-thirds of companies are reducing their IT operational costs associated with locating and retrieving data. A similar number of companies use archives to improve customer service and to enhance revenue opportunities (see Figure 3).

FIGURE 3

Companies Are Finding Significant Business Value in Archives

Ways Data Archives Have Added Value

Improved customer service — 68%

Reduced IT operational costs — 66%

Enhanced revenue opportunities — 63%

(% of respondents)

n = 1,011

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study,* April 2015

The results are significant –companies that leverage their archives see millions of dollars in benefits. Just over half of the companies (53%) saw savings of $1 million or more over the past year from risk mitigation/avoidance of litigation, with the top 21% of companies reporting savings of over $10 million. 44% of companies saw savings of $1 million or more from reducing operational or capital costs, with the top 18% of companies reporting savings of over $10 million.

Perhaps most striking is the impact data archives are having on revenue. More than a third of companies (39%) saw $1 million or more in additional revenue during the past year from monetizing their archives, with the top 15% of companies realizing over $10 million in additional revenue.

Average savings and revenue increases seen in the past year across all companies that are monetizing their data archives are shown in Table 1.

The top 15% of companies generated more than $10 million in additional revenue from their data archives in the past year.

TABLE 1

Average Values Realized from the Use of Data Archives Over the Past Year

Operational savings	
Risk mitigation/avoidance of litigation	$11.4 million
Reduced operational or capital costs	$9.4 million
Enhanced revenue	
Additional revenue	$7.5 million

n = 791

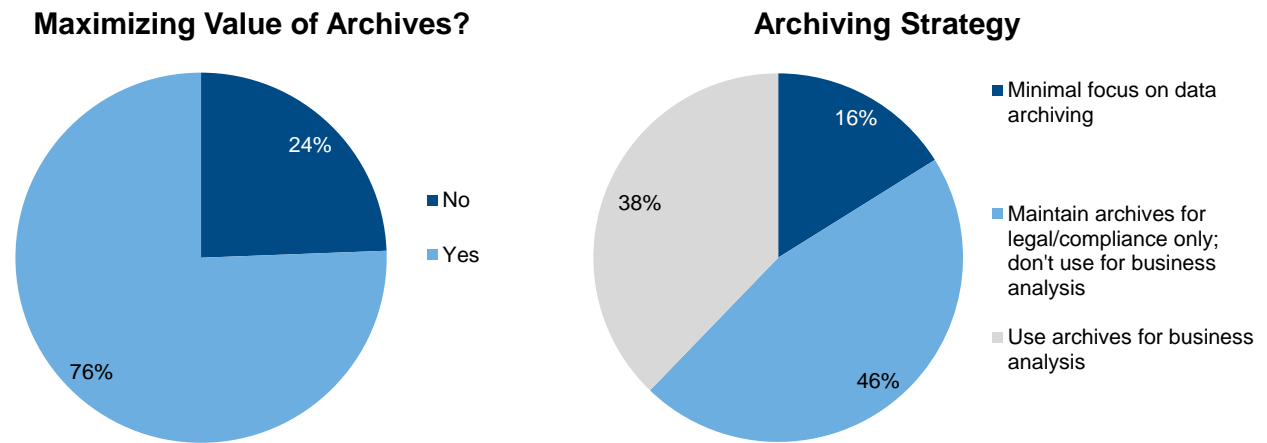Source: IDC and Iron Mountain's *Archiving Thought Leadership Study,* April 2015

Despite such impressive numbers, many companies aren't maximizing the value of their data archives. Worse, this missed opportunity is a blind spot because organizations believe they are doing a better job realizing value from the archives than they actually are. Most companies (76%) believe they are maximizing the value of their archives, yet only 38% of companies are using archives for business analysis (see Figure 4), which is critical to a business' ability to drive additional revenue by better understanding markets and improving products, service delivery, and customer service.

While most companies are blind to the value they are missing, 24% of companies acknowledge that they could be doing more. Nearly three in four of these companies (73%) believe they could be extracting two times or more value than they realize today (see Figure 5), with 27% of the group believing they could extract five times or more value.

Maximizing business value from the archives requires companies to have the right archiving strategy and processes that enable LOB to easily find information and empower stakeholders.

FIGURE 4

Businesses Are Overestimating the Value They Are Extracting from the Archive
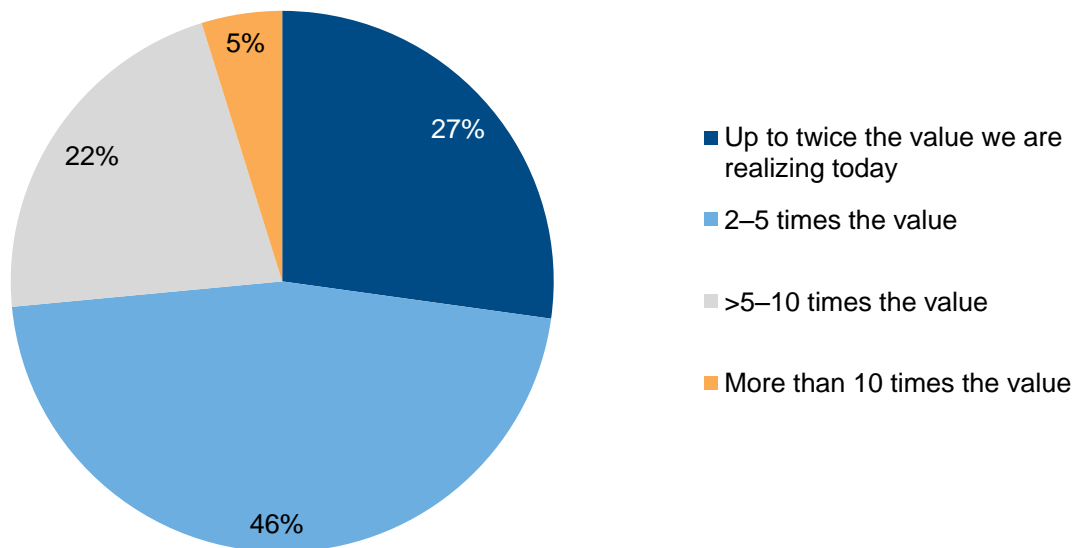
Maximizing Value of Archives?

Archiving Strategy

n = 1,011

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study,* April 2015

FIGURE 5

Significantly More Value Could Be Extracted from Archives

n = 229

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study,* April 2015

Legal and Compliance Are at Odds with IT and Lines of Business

Fully realizing the value of their data archives will also require companies to overcome important organizational challenges, beginning with the legal and compliance departments' frustration with the ability to access archives. Though the departments rely heavily on archived data to do their job, only 48% of legal and compliance respondents are satisfied with their ability to access archived information compared with 75% of LOB respondents and 78% of IT respondents. Time sensitivity for regulatory and legal requests fuels this frustration.
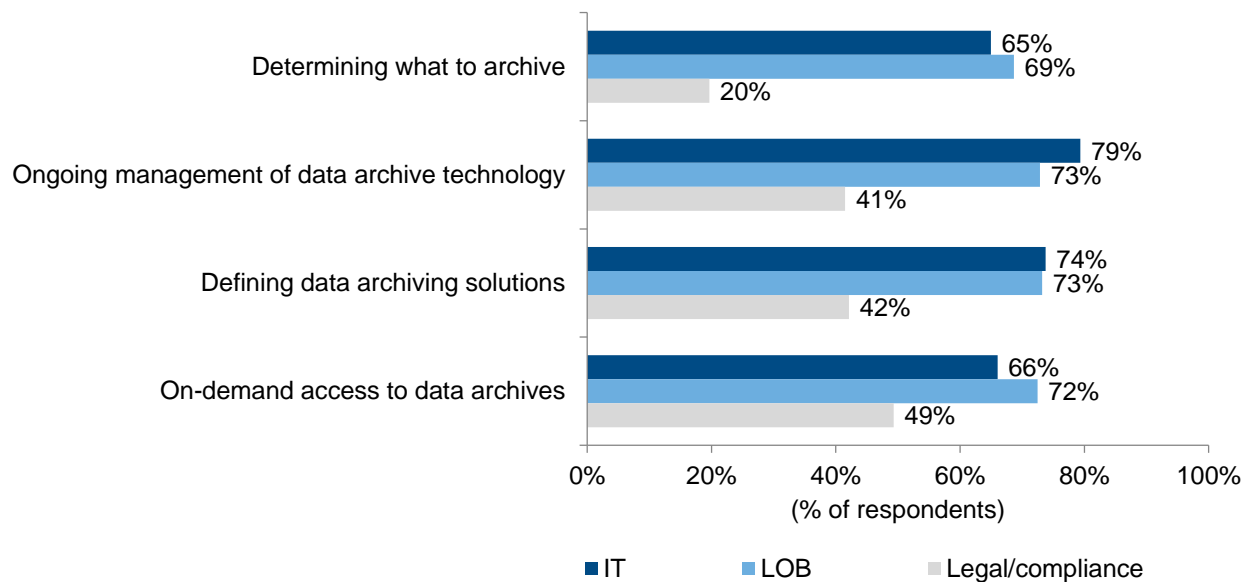
The frustration of legal and compliance goes beyond the ability to access information. The departments are critical of IT's management of many different aspects of the archiving process. Legal and compliance give IT low marks across a range of archiving activities, from determining what to archive to defining data archiving solutions. In contrast, LOB is largely on the same page as IT, with high satisfaction ratings across the board (see Figure 6).

Legal and compliance aren't on the same page with IT regarding which department has primary responsibility for archiving. Both are much more likely to see themselves as responsible for many aspects of data archiving, from determining what to archive to ensuring data archives are secured (see Figure 7). IT is least likely to see legal and compliance as responsible for archiving, with LOB falling somewhere between the two.

FIGURE 6

Legal and Compliance Less Satisfied with IT Performance Across a Range of Archiving Metrics
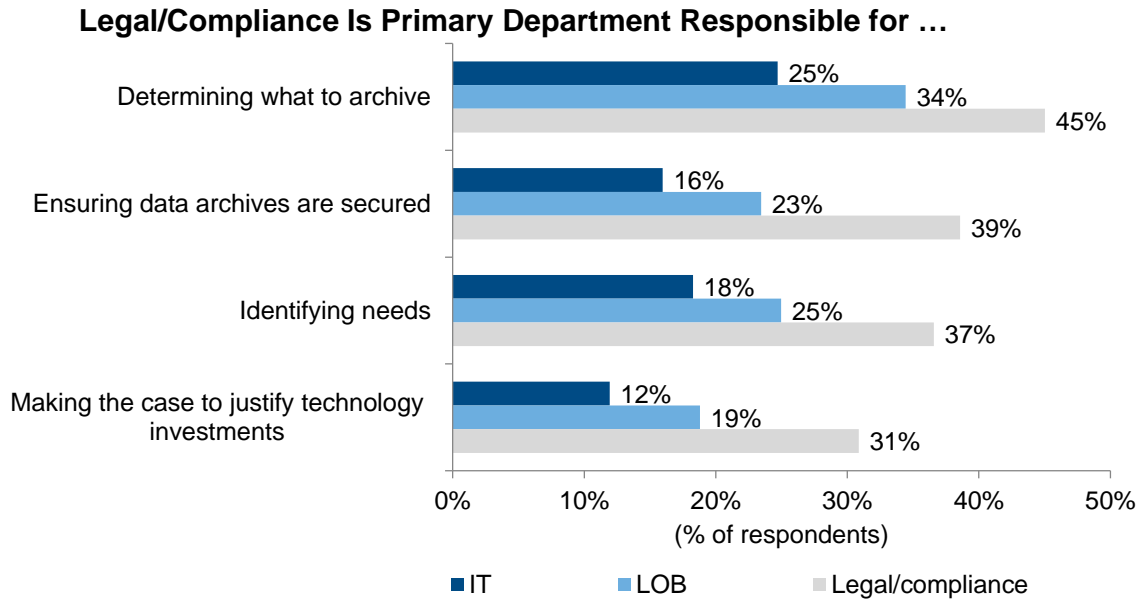
Satisfaction with IT Performance Across Archiving Metrics

n = 1,011

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study,* April 2015

FIGURE 7

Legal and Compliance More Likely to See Themselves as Responsible for Data Archiving

Legal/Compliance Is Primary Department Responsible for …

n = 1,011

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study,* April 2015

Legal and compliance respondents also see less value in leveraging the archive for legal and business purposes than LOB and IT respondents. For example:

- 38% of legal and compliance respondents see archives as reducing litigation costs compared with 63% of LOB respondents and 68% of IT respondents.

- 38% of legal and compliance respondents see archives as enhancing revenue compared with 60% of LOB respondents and 70% of IT respondents.

- 51% of legal and compliance respondents see archives as improving customer service compared with 69% each of LOB and IT respondents.

> **Only 38% of legal and compliance respondents see archives as enhancing revenue compared with 60% of LOB respondents and 70% of IT respondents.**

How can different departments have such dramatically different impressions of data archives? IDC believes the underlying reason for this disconnect stems from a fundamental difference in roles and objectives: Legal and compliance focus on risk mitigation and limiting access to information in the archives, while LOB wants to broadly disseminate information, including data from the archives. By sharing information throughout the organization and with customers, suppliers, and partners, LOB is looking to monetize the data and drive better business results.

Other IDC research shows that IT is evolving into a service organization, working closely with LOB to enable core business operations. With this focus, legal needs and requests may be seen as secondary to the core mission of IT instead of its top priority.
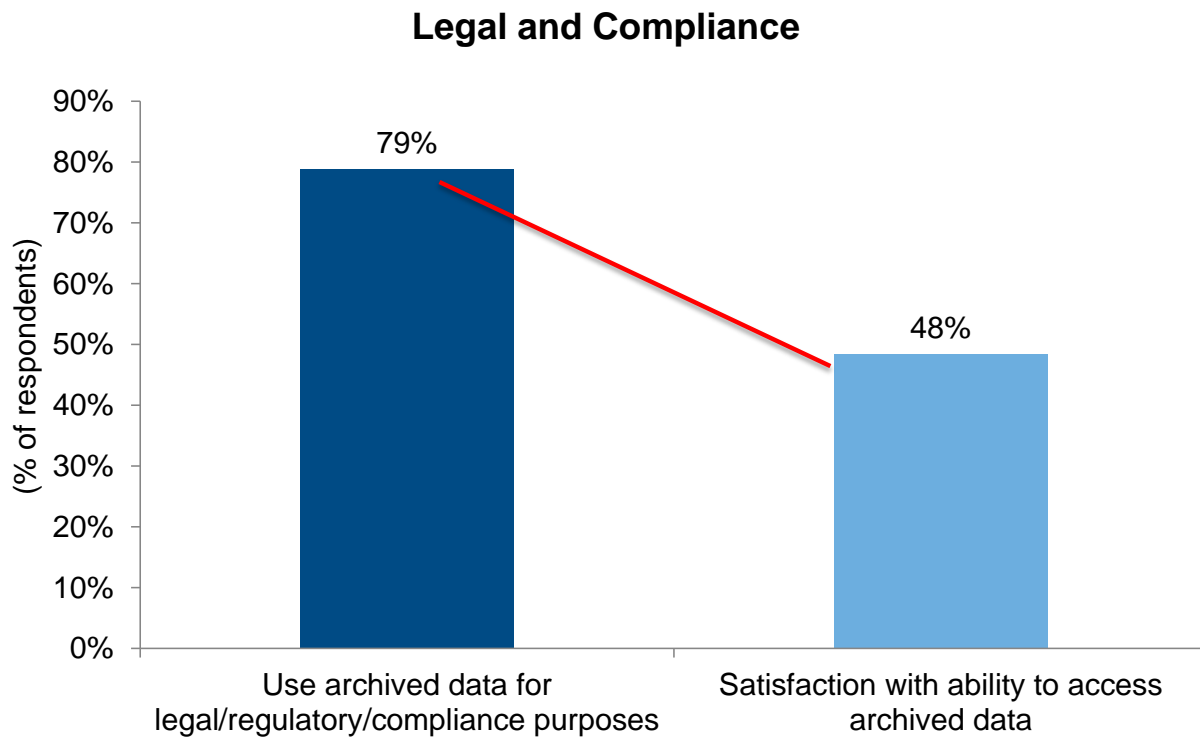
Companies Must Make Data More Accessible

Making archives fast and easy to access is critical for all departments using the archive. Our study shows most companies could do more to make their archives easier to use. Legal and compliance, in particular, are frustrated with their ability to access the information they need. Almost 80% of legal and compliance respondents use archived data, but less than half (48%) are satisfied with their ability to access this data(see Figure 8).

> **Almost 80% of legal and compliance respondents use archived data, but less than half are satisfied with their ability to access this data.**

FIGURE 8

Legal and Compliance Not Satisfied with Archive Access
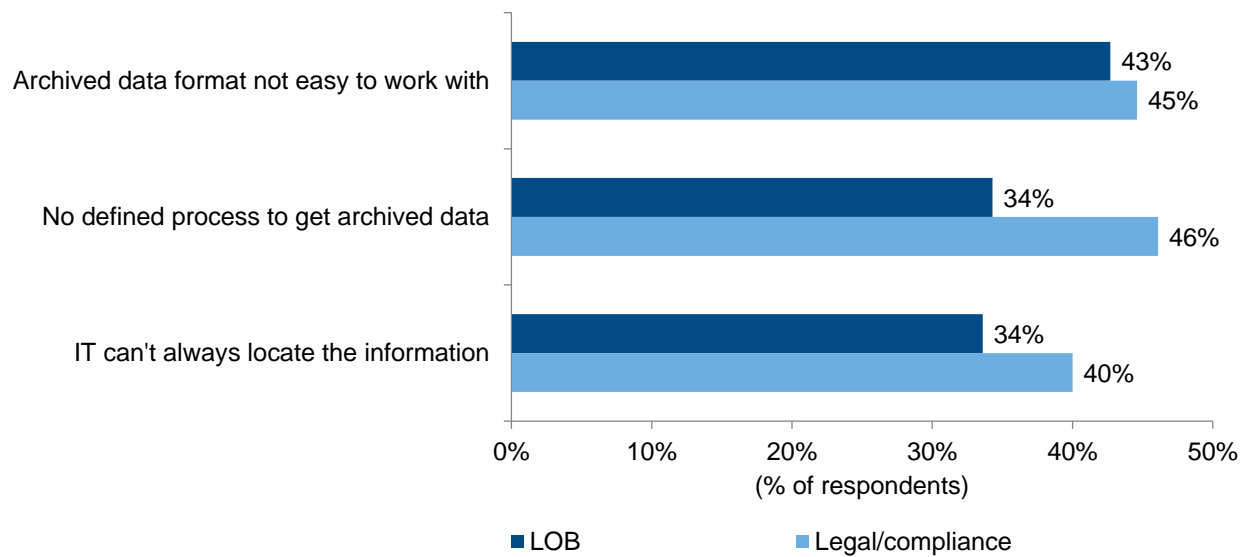

Legal and Compliance

n = 91

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study,* April 2015

Working with IT to get information from the archives is frustrating for both legal and compliance and LOB. More than 60% of LOB respondents and more than 70% of legal and compliance respondents experience frustrations. Receiving archived information in a format that is not easy to use is the greatest source of frustration, followed by the lack of a defined process to access archived information and IT's inability to locate information when it is needed (see Figure 9).

Over 60% of LOB respondents and 70% of legal and compliance respondents experience frustrations when working with IT to get archived information.

FIGURE 9

Largest Frustrations Working with IT to Get Archived Information

n = 378 for LOB, n = 93 for legal/compliance

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study,* April 2015

The challenges of accessing archived data impact LOB productivity: 49% of respondents believe lines of business lose significant productivity searching for information in data archives, and 48% of respondents say lines of business spend too much time searching for information that is difficult to access. Without a defined process to easily access archived information, LOB is wasting time and missing opportunities to monetize data in the archive.

ESSENTIAL GUIDANCE

With the important role archiving plays in the business, and its significant untapped value, companies can't afford to ignore frustrations over archive accessibility and disconnects regarding who has responsibility for the archives. IDC recommends companies take several steps, including:

- **Hire a chief data officer (CDO) to be directly accountable for all data issues.** This person should spearhead the development of an archiving strategy that takes into account the role of the archive and maximizes its value for all internal audiences. The CDO should lead a cross-functional team including IT and internal archiving customers (LOB and legal/compliance) to define the tools, processes, and procedures to ensure smooth and effective operation of the archive. The CDO should also tap into organizational best practices around process management (e.g., Six Sigma), leveraging internal experts when possible. Finally, the CDO should be equipped to work side by side with the organization's chief operating officer and chief information officer to define long-term business and data strategies.

- **Develop information maps of all data sources and repositories across the organization.** Information maps should be set up by business unit, group, and business process as a way to make archives more understandable and accessible to drive monetization of archived information. These maps should identify the types of information in the different archives and the value of the information to the organization.

- **Implement a holistic, consistent archiving approach across all areas of the business.** Organizations should take into account data retention schedules, use cases, the value of data, and necessary accessibility and archive costs. They should consider a technology refresh for any legacy tape or disk archives that are needed for Big Data and analytics and consider moving deep archive and unstructured data not needed for business analysis to offsite tape or disk solutions to free up internal storage and manage costs.

- **Consider working with a third-party partner.** Partners often have specific expertise that can help organizations optimize their final archiving solution, enabling data to be accessed faster and more easily, and a partner can free up an organization's internal IT resources to focus on innovation and strategy that support the bottom line.

CONCLUSION

Many of today's successful companies are extracting significant value from their data archives — not only mitigating legal and regulatory risk but also reducing operational costs and driving revenue. Unfortunately, many other companies cannot fully monetize their archives because they lack a rational, cohesive archiving strategy.
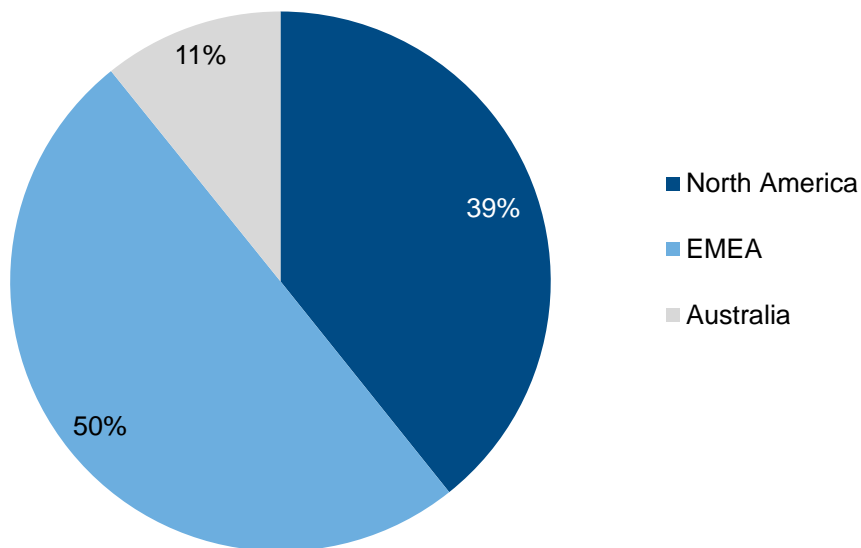
Key factors holding back companies include difficulty leveraging archives for business analysis, the mistaken belief that they are doing a better job monetizing their archives than they actually are, and considerable disconnects between legal/compliance and LOB/IT regarding archive value and accountability. Addressing each of these issues is key to truly unlocking the value of the archive.

Methodology

The information for this white paper came from the April 2015 global *Archiving Thought Leadership Study* developed by IDC and sponsored by Iron Mountain. IDC surveyed 1,011 members of senior and executive management involved in data archiving in organizations with more than 500 employees across a broad range of industries. Respondents were from IT departments and legal/compliance departments as well as lines of business. They were asked about their current archiving strategies and processes, the business value of their data archives, and how satisfied they are with their access to archived information for a variety of business use cases. More details on survey respondents can be found in Figures 10-12.

FIGURE 10

Respondents by Region

n = 1,011

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study,* April 2015

FIGURE 11

Respondents by Organization Size

500–1,499 26%
1,500–4,999 33%
5,000–9,999 22%
10,000+ 19%

n = 1,011

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study,* April 2015

FIGURE 12

Respondents by Industry

- Banking 17%
- Manufacturing 9%
- Healthcare 7%
- Retail/wholesale 9%
- Professional services 8%
- Broadcast/telecommunication 6%
- IT 17%
- Education/government 6%
- Construction/transport/energy/utility 13%
- Other 8%

n = 1,011

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study,* April 2015

Country-Level Data

Tables 2-9 contain key data and findings at the individual country level.

TABLE 2

Estimated Annual Savings Provided by Data Archives in Terms of Risk Mitigation/Avoidance of Litigation

	United States	Canada	United Kingdom	France	Spain	Netherlands	Germany	Australia
Under $10,000	1.4	.0	1.0	2.1	6.4	.0	.0	.0
$10,000 to less than $100,000	8.7	.0	12.6	15.6	17.3	19.4	14.1	7.5
$100,000 to less than $1,000,000	27.1	75.0	35.9	32.3	34.5	45.2	35.9	38.7
$1,000,000 to less than $10,000,000	33.2	25.0	27.2	32.3	16.4	32.3	30.8	39.8
$10,000,000 to less than $50,000,000	17.0	.0	12.6	7.3	19.1	3.2	11.5	8.6
$50,000,000 or more	9.2	.0	6.8	5.2	1.8	.0	7.7	2.2
Don't know	3.4	.0	3.9	5.2	4.5	.0	.0	3.2

n = 873

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study*, April 2015

TABLE 3

Estimated Annual Savings Provided by Data Archives in Terms of Reducing Operational or Capital Costs

	United States	Canada	United Kingdom	France	Spain	Netherlands	Germany	Australia
Under $10,000	3.0	.0	2.2	3.6	7.1	3.3	.0	.0
$10,000 to less than $100,000	12.6	.0	21.1	15.5	18.2	23.3	23.3	12.5
$100,000 to less than $1,000,000	33.3	100.0	28.9	38.1	29.3	40.0	31.5	30.0
$1,000,000 to less than $10,000,000	26.7	.0	27.8	25.0	14.1	26.7	23.3	40.0
$10,000,000 to less than $50,000,000	14.7	.0	10.0	11.9	18.2	3.3	8.2	11.3
$50,000,000 or more	6.3	.0	6.7	3.6	2.0	.0	5.5	1.3
Don't know	3.3	.0	3.3	2.4	11.1	3.3	8.2	5.0

n = 791

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study*, April 2015

TABLE 4

Estimated Additional Annual Revenue Provided by Data Archives

	United States	Canada	United Kingdom	France	Spain	Netherlands	Germany	Australia
Under $10,000	5.1	33.3	8.0	2.2	10.7	11.5	7.8	4.8
$10,000 to less than $100,000	19.2	.0	9.0	18.7	21.4	26.9	28.6	11.9
$100,000 to less than $1,000,000	23.4	66.7	31.0	33.0	24.3	26.9	26.0	25.0
$1,000,000 to less than $10,000,000	26.7	.0	24.0	29.7	13.6	26.9	16.9	34.5
$10,000,000 to less than $50,000,000	15.0	.0	8.0	6.6	16.5	3.8	5.2	11.9
$50,000,000 or more	3.6	.0	4.0	3.3	1.9	.0	2.6	4.8
Don't know	6.9	.0	16.0	6.6	11.7	3.8	13.0	7.1

n = 817

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study*, April 2015

TABLE 5

Are You Maximizing the Value of Your Data Archives?

	United States	Canada	United Kingdom	France	Spain	Netherlands	Germany	Australia
Yes	76.8	60.0	72.0	84.2	79.3	83.7	68.9	81.7
No	23.2	40.0	28.0	15.8	20.7	16.3	31.1	18.3

n = 1,011

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study*, April 2015

TABLE 6

Data Archiving Strategy

	United States	Canada	United Kingdom	France	Spain	Netherlands	Germany	Australia
Minimal focus on data archiving	12.5	20.0	18.6	16.7	14.0	39.5	21.4	23.9
Maintain the archives for legal/compliance only; don't use for business analysis	47.7	20.0	49.2	55.0	43.8	44.2	41.7	52.3
Use our data archives for business analysis	39.8	60.0	32.2	28.3	42.1	16.3	36.9	23.9

n = 1,011

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study*, April 2015

TABLE 7

Significantly More Value Could Be Extracted from Archives

	United States	Canada	United Kingdom	France	Spain	Netherlands	Germany	Australia
Up to twice the value we are realizing today	20.9	50.0	21.2	15.8	44.0	.0	37.5	40.0
Two to five times the value	45.1	50.0	51.5	63.2	32.0	57.1	43.8	35.0
Five times the value or more	34.1	.0	27.3	21.1	24.0	42.9	18.8	25.0

n = 229

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study*, April 2015

TABLE 8

Number of Separate Electronic Archives

	United States	Canada	United Kingdom	France	Spain	Netherlands	Germany	Australia
One	3.3	.0	4.2	7.5	2.5	16.3	5.8	3.7
Two to five	37.5	40.0	51.7	28.3	34.7	32.6	45.6	45.9
Six to ten	39.3	20.0	28.8	42.5	31.4	39.5	29.1	37.6
More than ten	19.9	40.0	15.3	21.7	31.4	11.6	19.4	12.8

n = 1,011

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study*, April 2015

TABLE 9

Overall Approach to Data Archiving

	United States	Canada	United Kingdom	France	Spain	Netherlands	Germany	Australia
Archive everything/no process to determine what gets archived	42.1	40.0	47.5	37.5	32.2	60.5	38.8	46.8
Archive everything for some categories of data/have process to determine what to archive for other categories	47.7	40.0	39.8	50.8	49.6	30.2	47.6	38.5
Have process to determine what gets archived for all categories of data	10.2	20.0	12.7	11.7	18.2	9.3	13.6	14.7

n = 1,011

Source: IDC and Iron Mountain's *Archiving Thought Leadership Study*, April 2015

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA01701
USA
508.872.8200
Twitter: @IDC
idc-insights-community.com
www.idc.com