The Governance of Long-Term Digital Information

IGI 2016 BENCHMARK

information governance initiative

SUPPORTED BY

Preservica
Active digital preservation

TABLE OF CONTENTS

BENCHMARK QUICK TAKES: FIVE KEY INSIGHTS

> "The critical role of digital . . .archives in ensuring the future accessibility of information with enduring value has taken a back seat to enhancing access to current and actively used materials. As a consequence, digital preservation remains largely experimental and replete with the risks . . . representing a time bomb that threatens the long-term viability of [digital archives]."
>
> DIGITAL PRESERVATION: A TIME BOMB FOR DIGITAL LIBRARIES[1]

1. **We have a problem.** Virtually every organization surveyed (98 percent) has digital records and information it must keep (or wants to keep) for longer than ten years. Digital information asset protection and access over the long term is a universal problem for public and private organizations—both large and small—across a wide swath of verticals.
2. **It is a technology problem.** Shared network drives are the most common repository for the storage of information ***we know must be protected and accessed for at least ten years*** (68 percent identified it as a storage location—the top response). Every day at the IGI, we are exposed to the maladies that afflict IG programs, but this result surprised even us. Shared network drives are the nicotine of IT infrastructure: easy to access, highly addictive, and incredibly dangerous over the long term. We should know better. There are many better alternatives that replicate the convenience of shared drives but radically improve governance. This addiction to shared drives must end, particularly for digital information we want or need to keep for longer than the next tech update cycle.
3. **It is a business problem.** We see a tendency among business leaders to view the problem of long-term protection and access as an academic one or one owned by museums and national archives. This is demonstrably untrue. In fact, 86 percent of our survey respondents said they have responsibility for ensuring the protection and access for ***business records for longer than ten years,*** not just archival or historical information. Further, the line between these categories is blurring, as you will see in our Snapshot on the Associated Press below.
4. **It is a legal problem.** Legal requirements are by far the number one reason that organizations are keeping digital information for ten years or longer (89 percent said it was a driver, and it was the top category in our results). These statutory, regulatory, and other legal obligations are not theoretical nor are they going away. In fact, the trend is moving in exactly the opposite direction, toward greater regulation of information, broader retention, and more prescriptive and, in some cases, even longer retention periods. It is not unusual for a single multinational corporation to maintain a records retention schedule that incorporates over 8,000 individual legal recordkeeping requirements. One provider of legal information services maintains over 10,000 citations from over 30 countries globally. Moreover, these requirements are proliferating, with one provider estimating that its legal citation database grows by 6 percent or more annually.
5. **We know what we must do, but are we doing it?** 97 percent of our survey respondents told us that they are *"aware that technology (hardware and software) obsolescence could mean that long-term digital records and information are at risk of not being readable or useable in the future."* This is great news—awareness is very high. The bad news? The number one solution to this problem currently being undertaken by our industry: *"we are currently considering our approach."* (44 percent) The second most common approach? *"We have no comprehensive strategy."* (31 percent). Only 16% are actually transferring this critical long-term information to a standards-based digital preservation system. The contrast between awareness and action is disappointing, but not unexpected. We have identified some of the perceptual factors at work, but another factor has been that, until relatively recently, there has not been a practical and systemic way to tackle this problem.

THE GOVERNANCE OF LONG-TERM DIGITAL INFORMATION

IGI 2016 BENCHMARK

> “We are moving into an era where much of what we know today, much of what is coded and written electronically, will be lost forever. We are, to my mind, living in the midst of digital Dark Ages . . .”
>
> TERRY KUNY, "DIGITAL DARK AGES?[2]

Introduction

Twenty years ago, a nonprofit representing hundreds of universities, national archives, museums, and other cultural institutions across the globe produced a landmark examination of the threat that digital transformation represented to our ability to capture, preserve, and provide access to our most important information. The report called for a global effort to design and develop “national information infrastructure to ensure that longevity of information is an explicit goal.”[3]

Today, no such global infrastructure exists. And, although significant progress has been made to address the challenge by industry bodies, individual institutions, and providers of digital preservation technology, the existential and commercial threat represented by our accelerating and deepening reliance on digital information has only grown exponentially in the intervening 20 years.

Archivists, historians, and librarians—among many others—have been sounding the alarm about an impending "digital dark age" and taking action to protect their digital information for decades.[4] However, for most corporations and organizations not explicitly engaged in historical preservation, this threat largely seems to have been relegated to the domain of academic specialists perceived as isolated from the prosaic demands of everyday commerce. Compounding the problem is the obvious human inclination to simply ignore problems for which there seems to be no easy or immediate solution.

However, this concern is neither academic nor theoretical. In fact, it is a problem shared equally by historians, by anyone taking a digital photograph, and by all organizations, large and small, who have replaced paper with digital in their businesses. In short, it is a problem we all share.

In the specialized world of archives, this problem is known as "long-term digital preservation." The word "preservation" is used here to denote a set of activities that go beyond simply storing a piece of information, but rather ensuring that the information remains accessible, trustworthy, secure, and authentic through its entire existence—even if that existence is forever.

A core part of our mission at the Information Governance Initiative (IGI) is to drive awareness and adoption of information governance (IG) as deeply as we can into the practices of public and private institutions around the globe. In fulfilling that mission, we are constantly seeking ways to "de-jargonize" information governance and its domains. In our experience, the term "preservation" is one of several that causes managers and executives to reflexively gaze down at their mobile devices and zone out until that part of the discussion is over. Further, it is our hope that this Benchmark will serve as an accessible introduction to the problem of long-term digital preservation for all audiences, not just those who already recognize it as a problem begging for a solution.

For this reason, throughout this Benchmark, we have adopted the phrase, "long-term protection and access." This phrase not only fairly captures the primary concerns of this domain, but also puts the focus on activities that are most relatable and top-of-mind for the managers and executives, i.e., those people who ultimately have the greatest influence on our ability to solve this problem simply because they control the money. "Protection" resonates because there is clearly a heightened and growing awareness of the need to invest in information security to confront the baseline threat that now exists in the digital world. "Access" is personally relatable to any executive who has been on the job for more than a few years and who has inevitably experienced the frustration (and fear) of not being able to locate and use an aging document vital to their job.

But, how long is "long-term?" At the IGI, we have yet to see a records retention schedule from a large organization that does not have several "PERMANENT" categories, even if those are just foundational corporate legal and financial documents. But even outside of this permanent category, most organizations have vast amounts of data that must be kept for periods longer than ten years (98 percent of them, in fact, as you will soon see).

This begs the question: in the digital world is there a material distinction between the need to keep something permanently and the need to keep something for at least ten years? We believe the answer is no. The inherent challenges of digital information (i.e., its ephemeral nature; proprietary data formats; proprietary software; software and hardware obsolescence; short-term thinking on IT architecture and infrastructure; storage media longevity; threats arising from complexity and volume; and so on) are essentially the same once you move out even a few years. For this reason, we have somewhat arbitrarily (but logically) chosen ten years as the practical equivalent to "very long" or even "permanent." Further, our ability to imagine keeping information for eternity is roughly equivalent to our ability to imagine infinity, i.e., very poor and difficult to act upon.

The IGI and its Supporters like Preservica are dedicated to advancing our understanding of this problem and its solutions. We share a vision with Preservica that this is a solvable problem. And, as you will see throughout this Benchmark, in addition to sharing our quantitative research, this Benchmark also includes snapshot stories of organizations and their visionary IG leaders who have done just that. This combination of data and anecdote provides a powerful message that we hope will play even a small role in helping organizations fulfill their responsibility to protect and provide access to their most critical digital information over the long term. Today, there is no difference between the digital world and the "real world." The time for short-term thinking is over. Let's take action.

Barclay T. Blair
Executive Director and Founder
Information Governance Initiative

ABOUT THE GOVERNANCE OF LONG-TERM DIGITAL INFORMATION: AN IGI 2016 BENCHMARK

“We are nonchalantly throwing all of our data into what could become an information black hole without realizing it . . . documents or presentations that we’ve created may not be readable by the latest version of the software. So even if we accumulate vast archives of digital content, we may not actually know what it is.”

VINT CERF, INTERNET PIONEER; CHIEF INTERNET EVANGELIST AT GOOGLE; DISTINGUISHED VISITING SCIENTIST, NASA JET PROPULSION LABORATORY[5]

The Governance of Long-Term Digital Information: An IGI 2016 Benchmark is based on quantitative, survey-based research conducted by the IGI in Spring 2016 that was distributed to our community of IG professionals. Nearly 400 professionals completed the survey in whole or in part. Respondents were a mix of both IG providers (i.e., people who work for organizations that provide IG products and/or services) and IG practitioners (people charged with doing IG at and for the organization where they work).

Because we believe this data to be the most insightful and revelatory of current industry perceptions, throughout this Benchmark we have chosen to primarily report on data drawn exclusively from IG practitioners who completed the entire survey, a population of 196.

About two-thirds of respondents in that population were from the USA, with the remainder split nearly evenly between Canada, the UK, and a group of other nations. By vertical, survey respondents were diverse, with about a quarter from Government and Military, 15% from Financial Services, and the majority of the rest from Legal, Healthcare, Utilities, Education, Manufacturing, and Pharma (ranked in descending order).

Organizations, both large and small, were also well represented, with about a third from large organizations (i.e., 5,001 or more), a third from mid-sized organizations (i.e., 501-5,000), and a third from small organizations (i.e., from 1-500 employees).

The majority of respondents identified their primary IG role as records and information management, which is in line with our expectations given the focus of the Benchmark. There was also strong representation from respondents focused on electronic discovery, data governance, legal, compliance, risk management, IT management, privacy, information security, and business management (in descending order).

In summary, we were very pleased with the survey response rates and diversity. We believe this data provides a very strong and deep insight into current attitudes and activities from practitioners who are well qualified to represent their organizations’ attitudes and activities regarding long-term protection and access.

IG SNAPSHOT

Preserving State History and Ensuring Citizen Access to Digital Government Records Using the Cloud

> "Most records today are born digital. As the official archive of state government we need to retain many of these records permanently. To meet this challenge we invested in expanding and enhancing our digital preservation capabilities, which was a significant undertaking. I'm confident that our approach will ensure that these essential government records remain accessible long into the future."
>
> JELAIN CHUBB, TEXAS STATE ARCHIVIST

The Texas State Library and Archives Commission (TSLAC) is taking action to ensure that critical digital records are properly governed and preserved in fulfillment of its mission to "safeguard government and historically significant records and to provide information services to support research, education, and individual achievement."[6]

TSLAC, with over 160 employees, was established in 1909. It supports a state government that has an annual budget of over $200 billion and employs more than 200,000 people.[7] Texas, if it were a country, would have the world's 12th largest economy.[8] TSLAC faces a massive ongoing deluge of digital information that must be governed and preserved in accordance with its legal obligations and agency mission.

One recent challenge for TSLAC was taking ownership of over 7 terabytes of digital records created by an outgoing gubernatorial administration which consisted of policy documents, press releases, and correspondence in a number of different file types (including digitized audio, still images, and video). On top of this, TSLAC had already created 26 terabytes of digital surrogates that required management and long-term preservation. In addition to preserving this information, TSLAC's mandate includes ensuring that both government users and the public at large have ready and secure access to records in its custody (as required by law). TSLAC also faces budgetary constraints and the pressure "to do more with less," just like many other organizations in both the public and private sectors.

To address its governance, access, and cost requirements, TSLAC developed a set of clear system requirements that it used to evaluate and select the tools and systems it needed.

Critical evaluative criteria for TSLAC included:

- **Cloud delivery**. TSLAC had concluded that cloud delivery was the best fit for the organization given the potential for lower acquisition, operational costs, and maintenance costs.
- **Support for standards.** Support for relevant standards such as the Open Archival Information System (OAIS) reference model (ISO 14721).
- **Migration.** Automated migration of records into new file types for long-term preservation to fulfill its mandate to ensure access for the entire life of the record (and in some cases, forever).
- **Integration.** Ability to function alongside and integrate with existing content and records systems.
- **Sector-specific expertise.** TSLAC concluded that it was important to select a provider with demonstrable understanding of unique governmental requirements.
- **Secure and reliable cloud infrastructure.** In particular, TSLAC was drawn to the AWS GovCloud, which was designed to support governmental use cases and requirements including, for example, encryption of records both in transit and at rest.

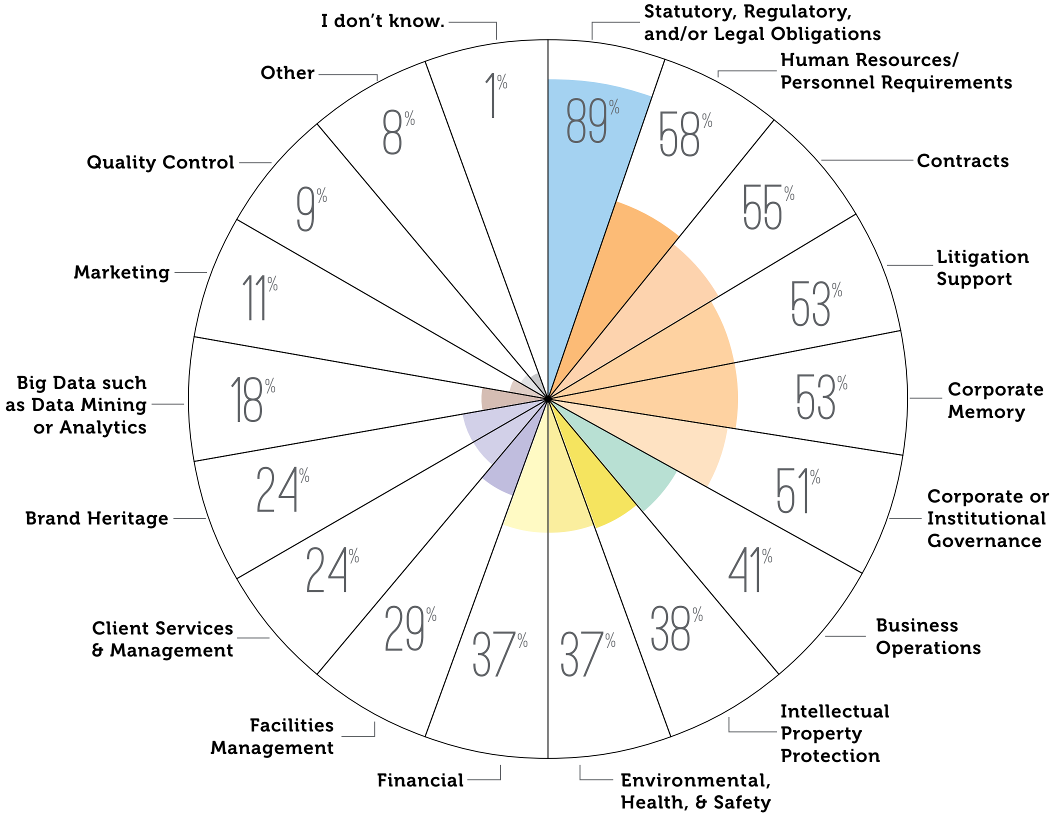
To meet these requirements, TSLAC selected and deployed a cloud-based solution from IGI Supporter Preservica. Preservica's service now also powers the recently launched Texas Digital Archive, which provides access to the publicly available electronic records collections of the TSLAC.

DEEPER ANALYSIS

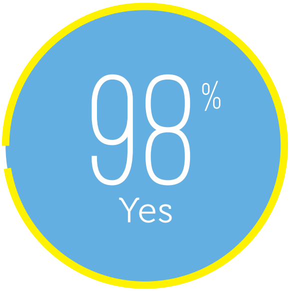
WHY DO WE NEED LONG-TERM PROTECTION AND ACCESS?

MOST ORGANIZATIONS HAVE DIGITAL RECORDS AND INFORMATION THEY KEEP LONG TERM BECAUSE OF THEIR IMPORTANCE

Organizations Report a Variety of Reasons Why They Keep Digital Information



The vast majority of practitioners (98%) report that their organizations have digital records and information they keep or need to keep for more than 10 years.



We asked practitioners whether or not their organizations had digital records and information they *keep or need to keep* in excess of 10 years. As the infographic shows, an overwhelming majority of respondents (98 percent) reported that they do.

These results are not surprising and are consistent with our anecdotal experience of organizational behavior—many organizations do keep records long term. The results are also consistent with preliminary research IGI conducted as part of our **2015-16 Annual Survey**. In that research, a majority of practitioners (91 percent) reported that their organization's records retention policies and schedules included permanent records, and 89 percent said they had *digital records* that they must retain in excess of 10 years.

What are the digital records and information that organizations keep? We asked practitioners to tell us the reasons why they are keeping digital records and information for more than 10 years and to select all that applied. As the infographic shows, most organizations are keeping them for a range of important reasons (e.g., six of the responses were selected by over half of respondents).

"Statutory, Regulatory, and/or Legal Obligations" led the way as the most common response (89 percent). This is consistent with other research by the IGI that shows reducing or responding to outside risks are common drivers of organizations' IG policies. Indeed, these may be drivers behind a number of the options practitioners selected, here, for why their organizations keep digital records and information long term.

But a number of the reasons organizations say they are keeping digital records and information long term may have another side to them—regardless of whether organizations *have* to keep them, those digital information assets are likely to be important to the day-to-day functioning of the organization, too. "Human Resources/Personnel Requirements," "Contracts," "Corporate or Institutional Governance" were each selected by more than half of respondents and "Business Operations" by more than 40 percent, for example.

Regardless of the reason, digital assets should be considered business-critical, warranting formal steps to ensure that they are findable, readable, usable, and trustworthy long into the future. To do that requires a commitment to providing long-term protection and access as an inherent and critical part of an overall IG program.

A Practical Approach to Governing 170 Years of Critical Corporate Records

"With digital-only records, a number of things can go wrong. We have to deal with playback media that degrades and file formats and software becoming obsolete, among other long-term access challenges. It was vital to protect our unique digital assets from these risks by using digital preservation techniques much more sophisticated than simply storing the 'bits and bytes.'"
VALERIE KOMOR, DIRECTOR, ASSOCIATED PRESS CORPORATE ARCHIVES

As one of the only truly global news reporting organizations, Associated Press (AP) has been bringing us the news for 170 years. With journalists in over 100 countries, AP has been at the center of history for nearly two centuries. In the process AP has become the custodian of a vast treasure-trove of irreplaceable and historically significant information in a dizzying array of formats.

The task of ensuring that vital digital information is protected, preserved, and accessible for the next 170 years falls to Valerie Komor and her team in AP's Corporate Archives group. In 2003, the Corporate Archives was established with the mission to acquire, organize, preserve, and make available the historically valuable records of the institution, which include corporate, news and administrative records as well as photograph, audio and video collections. Today, the Archives holds 4,000 linear feet of records and over 30 TB of digital files. As nearly every document is today born digital, Valerie's challenge has been growing not only by volume, but also by complexity–with no end in sight.

Valerie and her team took on this challenge by focusing on ways they could practically govern their information while minimizing the burden on the organization. Here are the steps they took:

1. **Pragmatic & risk/value focused.** Valerie and her team are responsible for a massive amount of information requiring governance. It cannot all be tackled at the same time, nor does all of it require the same level of governance. So, the team conducted a prioritization process and started with corporate records and information essential to documenting AP's business history in the event of a system failure or other disruptive event.
2. **Phase and iterate.** In addition to prioritizing IG activities based on a clear assessment of information risk and value, AP adopted a phased approach. This means they divided their information into chunks based on content, anticipated use, and physical condition. This was the only practical way to approach their project because the volume of records is too great to allow any other approach. Valerie started with full sets of annual reports and charters and bylaws and intends to bring in other collections as they are reviewed. These include vast amounts of original wire copy, the ephemeral sheets of news copy, which flowed off teletype machines from 1920 until 1986 and survive within bureau records and other files.

To support this strategy, AP selected Preservica's standards-based digital preservation system, an approach that will also enable them to automate the operational and technical aspects of the project while meeting AP's needs for IG and long-term accessibility of its one-of-a-kind corporate history.

WHAT TECHNOLOGIES ARE ORGANIZATIONS USING?

CURRENTLY USED STORAGE SOLUTIONS ARE PUTTING LONG-TERM DIGITAL RECORDS AND INFORMATION AT RISK

WHERE ARE DIGITAL RECORDS AND INFORMATION BEING STORED?	
Shared Network Drive	68%
Line of Business Applications (e.g. CRM, ERP, Manufacturing, HR Systems, etc.)	52%
Enterprise Content Management System (ECM)	47%
Disk or Tape Backup Systems	44%
Records Management System (e.g. EDRMS)	43%
Application-specific Archiving (e.g. email)	33%
Removable Media (e.g. CD or USB)	22%
Enterprise Information Archiving System (EIA)	14%
Purpose-built, Long-term Digital Preservation System	11%
Other	9%
Commodity Cloud Storage (e.g. Amazon)	8%
I don't know.	1%

Most organizations are not storing their long-term digital assets in a manner sufficient to ensure their long-term protection and accessibility. In fact, the top method is shared network drives. This option, like a number of the others listed (including ECM and EDRMS), even with additional backup or archiving, provides no inherent capability to address the unique requirements of this class of information. This exposes the organization to the risk of not being able to read and use these digital information assets in the future, for example, if your organization no longer supports or licenses a particular application or the file format becomes obsolete. In addition, shared network drives are notoriously insecure and nearly impossible to govern well, further exposing these assets to accidental or malicious tampering and deletion.

Organizations should seek out technological solutions that are purpose-built for the unique requirements of long-term protection and access. Unfortunately today, only a small percentage of organizations (11 percent) are employing these systems, putting vast swaths of critical information across the globe at risk.

THE AWARENESS AND ACTION GAP

PRACTITIONERS KNOW DIGITAL RECORDS AND INFORMATION ARE AT RISK, BUT PRESERVATION STRATEGIES HAVE NOT CAUGHT UP

HOW ARE ORGANIZATIONS ADDRESSING THE CHALLENGE?	
We are currently considering our approach.	44%
Convert official records to formats like PDF, TXT, CSV, etc.	33%
We have no comprehensive strategy.	31%
Postpone action until required (such as on-demand conversion or migration)	16%
Transfer to a standards based digital preservation system	16%
Other	12%
Convert to analog format (paper or film)	10%
I don't know.	6%

Most practitioners (97%) are aware that technology obsolescence could put long-term digital records and information at risk of not being readable or useable in the future.

97%
Yes

Why aren't organizations doing more to protect their digital information assets? Awareness of the problem is very high—97 percent. Yet, many are failing to take definitive action to ensure that their critical information assets are protected and accessible over the long term.

We asked practitioners what their organizations were doing to address the unique challenge of safeguarding their long-term digital records and information and to select all that applied. While it is good news to see that 44 percent are currently considering what to do (as the infographic shows), only 16 percent report that they are transferring data to a standards-based digital preservation system. Further, nearly a third of our respondents (31 percent), report that their organizations do not have a comprehensive approach.

Sixteen percent report postponing action until it is required—a risky strategy. As discussed previously, if you delay the steps necessary to safeguard your information from the start, degradation, corruption, and obsolescence can happen in the meantime. You may find when you need digital records and information they are not fully intact or that the costs (time, money, and technical resources) necessary to access and read them are prohibitively high.

Finally, a third of respondents report that they are converting official records to a common file type (e.g. PDF, TXT, or CSV). While this approach might seem to work, for now, for certain types of documents, there is also the risk that the chosen file format itself might become obsolete. If you adopt a strategy of converting once (especially if you do not also retain the original format), you also risk losing your vital information should such obsolescence occur. To be effective, digital preservation needs to be an active process. In addition, these simplified formats do not really work for certain content. You can't preserve multimedia files (images, video, and audio, for example) this way. Further, other content, like websites, emails, spreadsheets, slide presentations, and maps, for example, lose their interactivity, context, and inherent value when saved this way.

Future-Proofing Critical Digital Data in an Increasingly Complex Global Regulatory Environment

> "We have a very large repository of physical and digital records that require long-term preservation and access. Critical digital information is also being created every day, at high volume. We needed a system that could help us govern information over the long-term and also integrate with our existing systems so we could achieve a single, cohesive view of our most important information assets."
>
> TINA STAPLES, HSBC GLOBAL HEAD OF ARCHIVES

HSBC, one of the largest financial services organizations in the world, was founded 150 years ago in Hong Kong with a mandate to finance trade between Europe and Asia. With a fascinating corporate history that is woven into the fabric of world history, itself, HSBC today serves nearly 50 million customers in 72 countries.

Along the way, the bank has accumulated a vast and fascinating archive that includes photos, letters, and bank notes as well as critical evidence of strategic decision-making at the bank. This information plays a vital role in enhancing brand value, supporting a wide variety of HSBC projects and events, and informing researchers, historians, and the general public. However, the challenge does not end at preserving and presenting history.

Tina Staples is global head of HSBC's Archives team, a group of twenty specialists located in London, Hong Kong, Paris, and New York. As the group's name suggests, Tina's team governs HSBC's historical information, but her mandate has expanded to governing the digital information that the bank creates every day—information of enduring historical value, that will provide essential evidence of the bank's activities and decision-making.

It was critical that the bank's approach to IG addressed both the ***past*** and the ***future***. In order to future-proof and safeguard digital information the HSBC team realized they needed an approach that would not only provide long-term preservation of existing information, but one that would integrate with the HSBC cataloguing system to provide a unified view of the archive. This was a practical need that Tina's team knew was essential for both adoption and usability. However, this needed to be done in a way that addressed the compliance complexity inherent to an organization in a heavily-regulated sector, operating globally, and subject to the (sometimes contradictory) laws and regulations of numerous jurisdictions.

HSBC's legal and regulatory environment is incredibly complex, meaning that its information assets are subject to multiple overlapping privacy and security requirements. To achieve compliance, HSBC adopted a foundational IG approach focused on identifying and addressing interests, concerns, and requirements of critical stakeholders including HSBC's chief legal officer as well as senior representatives from RIM, IT, Legal, Compliance, and Risk. Making sure all relevant stakeholders were consulted during such efforts was a key to successful implementation and project success.

To address these needs as part of its overall IG program, HSBC opted for on-premise software from Preservica. The bank has already ingested many born-digital records from HSBC's more recent business activities and continues to develop and evolve its capabilities to ensure long-term preservation and access for its critical digital assets.

GOVERNING LONG-TERM DIGITAL INFORMATION: TAKING ACTION

What can you do, today, to help make sure that your organization's long-term digital records and information are protected? Here are our recommendations to help you get started.

1. Triage

You might have digital information in your organization right now that is in serious danger of being lost, damaged, or rendered inaccessible. This is not the time for careful deliberation or assessment. It is time for action. Perhaps, some repositories or information types immediately come to mind? The 10,000 backup tapes for the merger that seems like just yesterday but in fact will be ten years in June? The obsolete email archive filled with records you know you need to keep, but the system is moldering away in a forgotten data center somewhere? Talk to people responsible for IT storage infrastructure and also line-of business owners about their most immediate concern, and start there.

2. Assess

Once the most critical at-risk repositories and information types have been stabilized and addressed, it is time to conduct a formal assessment so that you can benefit from strategic planning and economies of scale. Do you have digital information that you need to keep longer than ten years? If so, where is it, what is it, and who had control of it? Is there a plan in place for its protection and access? Does your records retention schedule say that you are supposed to be keeping some records for ten years or longer? (Hint: This is likely the case). A critical first step is simply an assessment of the current state and visibility into your information environment.

An additional tip: if you are not already involved in electronic discovery (i.e., the process by which information is found, collected, and produced by your organization in the context of lawsuits and other formal proceedings), talk to the people who are as they often have a very comprehensive view of the information environment, and especially ancient data repositories that they have been required to produce data from. Another lesson to learn from these colleagues is pragmatism. These practitioners are often forced to accomplish complex information collection, categorization, processing, and management tasks under intense pressure and ridiculously short timeframes in incredibly high-stakes situations. In this environment, perfection is simply not possible, nor is it the goal. Rather, the standard is reasonable efforts and most importantly, progress and completion. All IG practitioners can and should learn from this as they approach long-term protection and access: focus on progress, pragmatism, and incremental improvement.

3. Address the Past, Protect the Future

Our massive stores of legacy information clearly must be brought under governance. However, legacy information may not be the right place for your organization to start (after you have triaged immediate risks as described above, that is). While you focus on the past, the present is conspiring to magnify and compound your IG problem. Every day your organization is creating new information—some of which likely requires protection and access over the long-term (as our research shows). Every day you fail to govern this new information is a day that only makes your future IG problem more difficult and expensive.

4. Catalog Consequences

Do you clearly understand the consequences of not being able to access, use, and rely upon your own records and information? Does your management? The consequences can be disastrous, and you need to assess, catalog, and rank these potential negative outcomes. What are the digital records and information your organization is keeping long term? Are they important or business critical? Knowing why they are of value to your organization can help you make the case for investing adequately in their preservation (e.g. fines for non-compliance, cost of legal challenge, reputational damage, failure to meet mandate, inability to leverage and re-use company knowledge, etc.).

5. Build Your Rules

Protection and accessibility of digital information over the long term must be a standardized part of your IG program. This means creating and enforcing rules. Do your existing IG policies and procedures address this need? If not, get to work. If you want to be sure your digital information assets will be available when you need them in the future, your policies, procedures, and systems must ensure that you can find, read, and use them.

6. Assess the IT Environment

Do you have the systems and infrastructure in place to protect and ensure access to your digital information assets over the long term? Despite widespread reliance revealed here by our research, shared drives and other general-purpose storage repositories are generally insufficient to address these unique requirements, without specialized customizations or add-ons that can address preservation beyond simple bit-level protection. In this regard, adherence to open industry standards is critical as a means to avoid the risk of inaccessibility due to the obsolescence of a proprietary technology. Standards are also critical for ensuring that these systems for long-term protection and access can talk to and exchange data with line of business applications, electronic content management (ECM) systems, and other repositories where these assets are created or temporarily stored.

ENDNOTES

We have used the following numeric convention for survey data throughout this document: results that included a half percentage point or more were rounded up, and results below half a percentage point were rounded down. As such, in some cases aggregated results for particular questions do not add up to 100 percent.

This work should be cited as: Information Governance Initiative, "The Governance of Long-Term Digital Information: An IGI 2016 Benchmark" (Information Governance Initiative LLC, May 2016).

[1] Margaret Hedstrom, "Digital Preservation: A Time Bomb for Digital Libraries," Computers and the Humanities 31: 189-202, 1998. 1998 Kluwer Academic Publishers.

[2] Terry Kuny, "A Digital Dark Ages? Challenges in the Preservation of Electronic Information," delivered at the 63rd annual International Federation of Library Associations and Institutions Conference, September 4, 1997.

[3] "Preserving Digital Information," Report of the Task Force on Archiving of Digital Information. Commissioned by The Commission on Preservation and Access and The Research Libraries Group, May 1, 1996.

[4] See, e.g., the seminal paper referenced above: "Digital Dark Ages? Challenges in the Preservation of Electronic Information," delivered by Terry Kuny at the 63rd annual International Federation of Library Associations and Institutions conference on September 4, 1997.

[5] Ian Sample, "Google boss warns of 'forgotten century' with email and photos at risk," The Guardian, February 13, 2015.

[6] State of Texas Legislative Budget Board, "Fiscal Size-Up 2014-2014 Biennium," February 2014.

[7] State of Texas Legislative Budget Board, "Fiscal Size-Up 2014-2014 Biennium," February 2014.

[8] If Texas were a country, its economic output of $1.65 trillion would make it the world's 12th largest economy. Mark J. Perry, "If New York is Spain and California is Brazil, What is Texas?" Newsweek, June 22, 2015. Online at: http://www.newsweek.com/if-new-york-spain-and-california-brazil-what-texas-344702

ABOUT THIS PUBLICATION

This publication was created by the IGI as part of our ongoing work exploring issues, strategies, and techniques related to information governance. As part of our commitment to excellence and to maintain objectivity, the IGI does not recommend, evaluate, or endorse specific products, services, or providers. However, the IGI's work is made possible through the generous contributions of our supporters for which we are grateful. This publication was made possible by Preservica's support of the IGI.

About the Information Governance Initiative

The Information Governance Initiative (IGI) is a think tank and community dedicated to advancing the adoption of Information Governance (IG) practices and technologies through research, events, advocacy, and peer-to-peer networking. We are dedicated to the professionalization of IG and have called for the creation of a new kind of information leader called the Chief Information Governance Officer. Our Annual Report has become an industry standard reference guide for organizations benchmarking and building their IG programs. The IGI Community is where thousands of practitioners from cybersecurity, IT, analytics, privacy, legal, records management and the other facets of IG come together and learn from each other. We produce hands-on educational workshops and executive roundtables each year. The IGI was founded by recognized leaders in the field of IG, and is supported by leading providers of IG products and services. You can find us online at iginitiative.com. Join us.

About Preservica

Preservica is a world leader in digital preservation software, consulting and research with active preservation solutions used by businesses, archives, libraries, museums, and government organizations globally to safeguard and share valuable digital content, collections and electronic records, for decades to come. Customers include the European Commission, Texas State Archives, Wellcome Library, the Associated Press, and HSBC, to name a few. More information about Preservica can be found online at: www.preservica.com

The information contained in this publication has been obtained from sources the author believes to be reliable. The author disclaims all warranties as to the completeness, adequacy, or accuracy of such information and shall have no liability for errors, omissions, or inadequacies herein. The opinions expressed herein are subject to change without notice. Although the author may include a discussion of legal issues, the author does not provide legal advice or services, and its research should not be used or construed as such.

information
governance
initiative

2016 Information Governance Initiative LLC www.IGInitiative.com