blancco

GUIDE BOOK

The Ultimate Guide to Data Retention

November 2016
Updated January 2019

CONTENTS

Introduction

"Data retention" is now everyone's concern, and its scope goes far beyond what data to retain and for how long.

Not long ago, data retention programs were the province of a handful of specialists in the legal and compliance departments. Organizations knew they had to retain certain documents for a specified number of years to meet legal and regulatory obligations and that was about it.

The situation is completely different today. New legal and business requirements mean that a cross-functional team is needed to create and enforce data retention policies. The CIO and CISO must help align data retention policies with organization-wide initiatives.

Many large enterprises are appointing a full-time or part- time Data Protection Officer (DPO) to comply with the EU's General Data Protection Regulation (GDPR).

Why the dramatic change? Driving factors include:

- The rising tide of legal and regulatory requirements for preserving documents and files of many kinds.
- The growing awareness that data retention is a cybersecurity issue—that erasing data no longer needed by the business reduces the likelihood that data can be stolen by cybercriminals and hacktivists.
- Privacy legislation and changing public expectations about privacy place choices about information retention and erasure in the hands of customers and third parties outside of the organization.

This guide is designed to help organizations wrestling with these challenges. It answers key questions about data retention policies and programs such as:

- ✓ How does the concept of "data lifecycle" help you shape data retention and protection policies?
- ✓ Why is data erasure suddenly so important, and why are so many organizations weak in this area?
- ✓ Who should be on the team to build a data retention policy and how should it be enforced?

If these questions are important to you, please read on.

Chapters at a Glance

CHAPTER 1

Data Retention: A Critical Part of Security

> **Our memory... is like a dispensary or chemical laboratory in which chance steers our hand sometimes to a soothing drug and sometimes to a dangerous poison."**
>
> **– Marcel Proust**

"Data Retention" is About Much More than What to Keep

In the good old days, most organizations had a conceptual view of data retention that was pretty simple (Figure 1-1). A limited set of electronic and hard copy documents and files had to be retained for a specified period of time (or in special cases, indefinitely). These documents and files had to be identified, protected and monitored for the designated time period and then destroyed.

Other documents and data were outside the purview of the data retention program and were handled according to the data management practices of individual employees and hundreds of different applications.

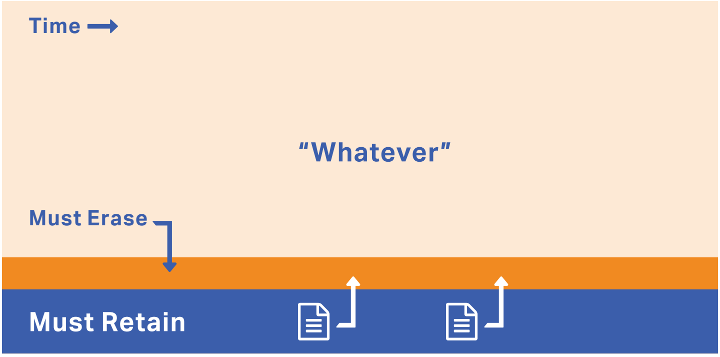
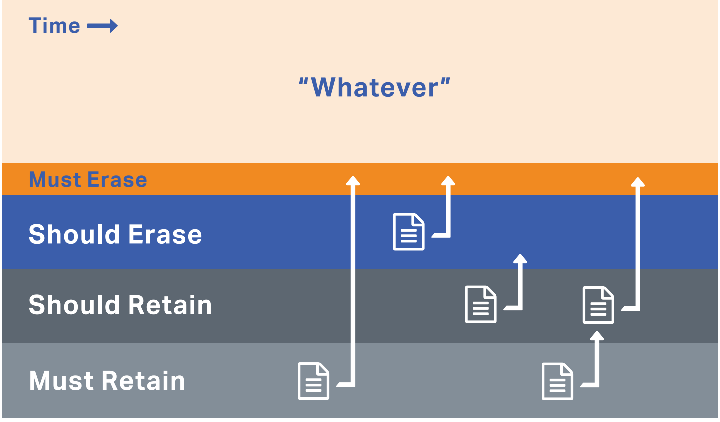


Figure 1-1: The old view of data retention

Of course, the reality was more complex and implementation could be demanding. However, most CIOs felt comfortable leaving data retention policy creation and enforcement in the hands of a few legal and compliance experts, or perhaps a consultant.

Today, a "data retention" program must be about much more than retention (Figure 1-2). As before, some documents and files must be retained and protected for specified periods. But organizations also need to think systematically about what items should be retained and which items should be erased, even when there is no absolute legal or business requirement. And today there are reasons why many more items must be erased.

Figure 1-2: A more complex view of data retention



Organizations also need to create policies and processes that handle documents and files appropriately as they migrate across categories. As files reach the end of required retention periods, should they be retained longer or erased immediately? For sensitive documents with no statutory retention period, how long should they be retained and when should they be erased? How should the organization handle requests from third parties like customers to delete personal information?

A data retention program also needs to ensure that intentions are carried out effectively. Are all sensitive files really destroyed beyond recovery when servers and personal computers are discarded or sold? If customers ask to be "forgotten," is their information actually erased everywhere it has been stored?

We will be looking at these issues in Chapters 2 and 3.

Retention, Security and Privacy

Organizations are taking a broader view of data retention programs because they realize the programs can have a major impact on data security and on meeting customer (and government) expectations about privacy.

From the perspective of cybersecurity, to state the matter plainly: information that has been erased can't be stolen and sold by hackers, and can't be used against the organization by hacktivists, hostile lawyers or anyone else. The possible business value of storing data indefinitely must be weighed against the risk of losing control over it.

Privacy has become a "hot button" issue on two fronts. Governments, particularly in the European Union (EU), have been raising the bar both for protecting customer data and for requiring that it be erased on demand. Customers have also become more sensitive about these issues. They are increasingly likely to look at privacy policies and security as reasons to do business with your organization – or with your competitors.

Activities: Classification, Monitoring and Enforcement

Data retention programs involve several major tasks. The first set of tasks revolves around determining legal, regulatory, business and security issues and requirements, and creating policies that address them.

But there are also a range of day-to-day activities that involve classifying documents and files, monitoring their use and storage, and enforcing policies for archiving and destruction. Documenting compliance with regulations and standards is also important.

We will examine these topics in Chapter 4.

Data Retention is a Team Sport

Defining data retention policies involves deciding what information must be retained (for legal, regulatory and business reasons), what information should be retained (typically for business reasons), what information should be erased (typically because of security and risk issues) and what information must be erased (primarily for privacy and security reasons). Implementing data retention policies requires knowledge of technologies and processes for storing, archiving and destroying data.

The wide range of knowledge and skills involved mean that data retention programs must be a team sport, with participation by legal and compliance experts, line of business managers, as well as IT and security staff. Third parties can also play an important role, for example, consultants and IT asset disposition (ITAD) firms offer tools for "erasing" data on devices you are selling or discarding.

Upper management guidance and support are also critical for the success of data retention programs, both to keep processes on track and to arbitrate the inevitable conflicts between wanting to save data "just in case," and to destroy it to minimize the impact of possible data breaches. Some organizations appoint a Data Protection Officer (DPO) to provide these management functions.

CHAPTER 2

What You Need to Keep: Data Retention and Protection

Reasons to Retain Data

Data retention programs have always been centered on "what you need to keep," and that continues to be the case. But today there are even more reasons to retain documents and files.

Regulatory Compliance

There are many far-reaching government laws and industry standards that require or strongly recommend retaining and protecting documents and files for specific lengths of time. A few of the most prominent include:

- PCI DSS (Payment Card Industry Data Security Standard)
- MoReq2010 (Model Requirements for the Management of Electronic Records)
- US FATCA (Foreign Account Tax Compliance Act)
- EU GDPR (General Data Protection Regulation)
- ISO 27001
- HIPAA (Health Insurance Portability and Accountability Act)

There are also bewildering multitudes of data retention regulations in hundreds of national and local jurisdictions. Some of these apply to virtually all enterprises, such as requirements to protect employment, payroll, accounting and legal records. Others are more targeted, covering documents and files specific to healthcare, manufacturing, technology, retail, energy, transportation and almost every type of business, not to mention government agencies (some of whom have to keep some files in perpetuity).

Litigation

Organizations have a legal obligation to protect documents that are reasonably likely to be relevant to future litigation. Once litigation has commenced, they must prevent the destruction of any information that is likely to lead to the discovery of admissible evidence. Failure to fulfill these obligations can be interpreted as obstruction of justice.

Contracts

Many organizations are bound by agreements with customers, suppliers and other third parties to retain documents, either for a specified period or for the duration of a contract. These sometimes include sales records, warranty and service records, design documents, legal documents, among many types of records.

Internal Business Processes

Most documents and data generated by business processes can be stored (and deleted) based on the policies and procedures of the applications that create them. However, some types of materials should explicitly be covered within a data retention program, because their availability cannot be guaranteed by the organization's base level storage and archiving procedures and there is a reasonable probability they will be needed at a future time.

The same attention is appropriate for files containing confidential or sensitive information. These files may need to be given extra levels of protection, and possibly also destroyed, based on carefully enforced policies. In your organization these might include executive emails, documents from the legal and investor relations departments and business plans.

Protection and the Data Management Lifecycle

Okay, we have been a bit inconsistent until now. Sometimes we have said "data retention" and sometimes "data retention and protection." The truth is, any data worth retaining in a systematic way is also worth protecting in a systematic way.

Most of the regulations mentioned above focus mostly on information protection. Regulators don't just want you to retain employment records, customer data, and ePHI (electronic protected health information) for a certain number of years; they want you to make darn sure it doesn't fall into the hands of cybercriminals.

While you are taking the trouble to analyze what classes of documents and files are important and how they should be stored, you should also define the security measures and monitoring that should be applied to those valuable assets at each step of the information management lifecycle (Figure 2-1).

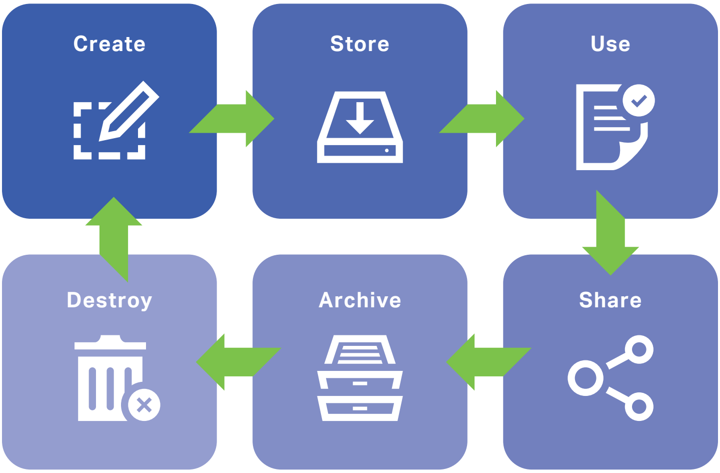


Figure 2-1: The information management lifecycle

Create

When documents and files are created, they first need to be classified (we will discuss classification in Chapter 4). Once they are classified, you can apply appropriate policies so they will be stored in locations with strong security defenses, including antimalware software, intrusion prevention systems and active monitoring. Policies can also dictate who will have access to the files (access control lists) and minimum levels for authentication (such as password standards and multi-factor authentication).

Store

Data retention policies should specify appropriate measures for protecting “data at rest.” This includes not only access controls and authentication processes, but also data encryption and minimum backup frequencies (including measures to encrypt and protect the backups). Variations in these policies may be needed for data stored on servers in the data center, on cloud storage services, on laptops and personal computers, and on mobile devices. You may even want to block sensitive files from being stored on public cloud services or smartphones.

Use

Additional controls can be deployed to protect data that is being accessed, viewed and processed. These might include access controls, encryption of “data in motion,” and digital rights management (DRM) solutions that prevent copyrighted and sensitive documents from being copied or redistributed by unauthorized users. Valuable information should also be protected by data loss prevention (DLP) software that can block files from leaving the network based on factors like their type (e.g. spreadsheets), their source (e.g. the CFO), and keywords and number patterns (e.g. “confidential” and XXX-XX-XXXX).

Share

Have you tried the “Anyone with the link” option lately?

Email, office applications, and workgroup collaboration packages make it incredibly easy to share documents. Sensitive files should be protected by disabling some of the sharing options in those applications. DLP solutions can also be used to limit sharing outside of the organization, and to prevent files from being copied to removable media.

Archive

Your policies should also establish standards for protecting data when it is archived for long-term retention. Encryption is essential, including control and preservation of the keys. You should also make sure there are security standards for choosing and monitoring third parties involved in archiving, such as transportation services (if you are using tape or other physical media) and cloud storage services.

Destroy

The destruction of data that reaches end-of-life raises many questions relating to policy and implementation. We discuss these issues in the next chapter.

CHAPTER 3

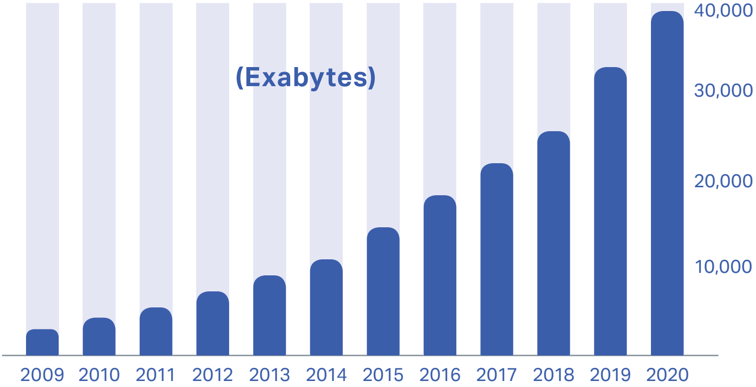
What You Can’t Afford to Keep: Data Erasure and Privacy

> **"Diligent as one must be in learning, one must be as diligent in forgetting."**
>
> **– Albert J. Nock**

The Case Against "Just in Case"

Computers and inexpensive storage have made hoarding the easiest it has ever been in the history of humankind. They accentuate the natural human bias towards saving everything until our closets are full to overflowing. Even worse, the amount of information available to store has been skyrocketing and is projected by IDC to exceed 40,000 exabytes (billion gigabytes) by 2020 (see chart).

Figure 3-1: Data growth in the digital universe. Source: IDC, The Digital Universe in 2020



Should we keep those outdated documents and files nobody has touched in years? Well, it is possible, conceivable, not completely beyond imagining, that somebody, someday, may want one of them. So let's save them all, "just in case." Besides, data storage is incredibly cheap these days.

But in reality, "just in case" is a terrible guide to policy. A single unnecessary document or an email could contain:

- A "smoking gun" that can be used against you in court
- Data that a hacker can use to attack your organization or your customers
- Dormant malware, waiting to be triggered as part of an advanced attack (try Googling "Regin malware")

Data retention programs provide the forum organizations need to weigh the potential value of storing data against the risks and costs of retaining it.

Legal exposure

If an organization is involved in a lawsuit, opposing lawyers can make their case by taking out of context information from documents and emails. To avoid this, you should erase files when they serve no business purpose, provided there are no legal or regulatory requirements for their retention.

Discovery costs

In the course of litigation, lawyers can force opponents to produce huge volumes of documents and emails. The discovery process involves both finding archived materials and examining them to ascertain which meet criteria supplied by the court. In one famous case, a pharmaceutical company realized that discover costs could reach $17 million and was forced to settle with the plaintiff. In another case an energy company settled because producing the required evidence would have taken six months and $6.2 million.[1] Eliminating unnecessary files reduces the potential costs of an involved discovery process.

Data breach costs

Every enterprise today must be concerned with the potential costs of data breaches. According to the Ponemon Institute, the average cost of a data breach reached $4 million in 2016, and

[1] *Thomas F. LINNEN, et als v. A.H. ROBINS COMPANY, INC. and Murphy Oil USA, Inc. v. Fluor Daniel, Inc.*

the average cost incurred for each record containing sensitive information totaled $158. Data breaches also affect customer trust and revenue. For example, 48 percent of U.S. consumers would consider changing healthcare providers if their medical records were lost or stolen.[2] Eliminating unnecessary files reduces the potential damage from breaches.

Data protection costs

While the costs of data storage have dropped, the cost of protecting data has not. This includes security software tools, as well as experienced staff to manage backup and archiving and watch for cyberattacks on the stored files.

Triggers for Data Erasure

There are actually a range of events and circumstances that can trigger a need for erasing data. All of them should be addressed in a data retention program.

Information end-of-life

As we have been discussing, many files are good candidates for erasure when they age past their retention period or reach the point where the risk of damage from their exposure exceeds the probable value of keeping them.

Customer demand

Customers are likely to make increasing use of their "right to be forgotten." Each request may involve digging out data on multiple systems and applications.

Equipment transfers and end-of-life

You must be careful that sensitive data cannot be read or recovered from disk drives and other media when computer equipment is transferred from one employee to another, and when

[2] Ponemon Institute: *Fifth Annual Study on Medical Identity Theft*

you sell or decommission used equipment. This applies to servers, laptops and personal computers, and mobile devices. This is a weak spot for many organizations.

Data migration

When organizations migrate data between servers, or between data centers, they have to be careful that the original copies were erased completely. This can be particularly challenging when data is stored on cloud platforms. Organizations need to ensure that sensitive data is not left behind on virtual machines or storage devices, where it might be accessible to other users of the cloud platforms.

Why Organizations are Failing at Data Erasure

The current state of data erasure in the enterprise is lacking. Most organizations show significant weaknesses in their tools, the reach of their programs, their monitoring and auditing, and a lack of awareness among IT staff as well as employees.

Ineffective tools and technology

The Delete and Factory Reset commands are not effective tools for data erasure. In most cases they simply remove pointers to the disk sectors where data resides. The data itself remains on the media and can be recovered by hackers and malicious insiders (for information on corporate servers), as well as by whoever takes possession of servers and devices when they are stolen or resold.

More sophisticated technologies also have shortcomings. Many mechanisms for reformatting and overwriting disks do not perform enough overwriting passes to ensure that information is unrecoverable, and do not provide erasure reports that meet regulatory requirements. When equipment reaches end of life, destruction or degaussing can render disk drives inoperable, but they eliminate any chance to reuse or resell them the systems.

Cryptographic erasure is a technique that involves destroying encryption keys, so encrypted files and documents cannot be decrypted and read. However, organizations have to be extremely scrupulous about encrypting all data before it is stored, and about managing the keys. If a hacker or insider steals the keys before they are destroyed, they can read the data. Computer speeds and cryptography are advancing so quickly that files encrypted a few years before often can be cracked. Also, encryption tools typically don't provide verification mechanisms or audit trails to prove that files have been erased completely enough to achieve compliance with key regulations.

Limited reach

Many organizations focus their data erasure activities entirely on corporate servers and storage systems. They often fail to address data stored on laptops and personal computers, tablets, smartphones and other mobile devices and removable media like USB drives.

Virtual machines and LUNs (logical unit numbers) pose additional challenges. Many data erasure technologies were designed to work at the level of physical devices and do not do well handling virtual and logical spaces that can migrate across physical environments.

Incomplete monitoring

Many IT groups do not regularly monitor data erasure activities across the enterprise. Many of the technologies they rely on for erasure do not provide the audit trails and reports required by some regulations (and by auditors who to see best practices employed). In addition, most organizations do not hold regular reviews of their data retention and erasure policies and processes, so that they fall behind new regulations and new technologies.

In the next two chapters, we will discuss how organizations can overcome these shortcomings with better data retention programs and policies, and with effective data retention and erasure tools.

CHAPTER 4

How to Build a Data Retention Program and Enforce Policies

"

Plans are only good intentions unless they immediately degenerate into hard work."

– Peter Drucker

Creating Data Retention Policies

Obviously, data retention policies will vary across different industries and sizes of businesses, but there are certain elements that should always be included.

Define the scope of the policy

The policy should always include a statement about its purpose and scope. It should describe the business reasons for the policy and list the major legal, regulatory and business requirements, including laws and standards that must be met. It should also specify the people affected by the policy (who may include third parties as well as employees) and the IT systems and equipment covered.

Classify the data

Good intentions degenerate into hard work at the point when you have to classify documents and files into categories for retention and erasure.

Most organizations start by identifying data that must be retained for specific periods, followed by data whose destruction is mandated by laws or regulations.

The next steps are to determine which documents and files fall into the "should be saved," "should be erased" and "whatever" categories (see Chapter 1) and to decide on retention and erasure rules for them. The team must weigh the possible future business value of the information against the risk of fines and costs that would result from a data breach.

Note that some documents and files can be classified based on multiple criteria. For example, you may want to retain all spreadsheets created by the finance department and stored on the SharePoint server, or erase all emails more than three years old, containing the words “confidential” or “personal,” and not covered by a retention requirement.

You may also want to calculate the full cost of retaining data, which includes the expenses involved in protecting, managing and monitoring the files. This will help counter the “just in case, since storage is cheap” mentality.

Specify how data will be retained and protected

Outline the policies and procedures for retaining and protecting data. This includes:

- ✓ Retention periods for each data category
- ✓ Policies for protecting files during each phase of their life cycle (see Chapter 2)
- ✓ Steps for handling files at the end of the required retention period; should they be erased automatically, reclassified into another category or turned over to their “owners” for final disposition?

Remember that data retention policies have to address all of the locations where sensitive data might be stored. That includes not only servers in the corporate data center, but also laptops and PCs, mobile devices and virtual machines in the cloud.

In some cases, the policy may forbid storing data on certain devices, or require certain conditions. For example, there are mobile device management (MDM) products that can prevent files from being saved on smartphones, or can isolate corporate files in secure “containers” that are protected from malware.

Specify how data will be erased

In Chapter 3, we outlined several “triggers” for data erasure. You might need a different set of procedures for each of them, including:

- ✓ Technologies and processes that will be used to erase files when they reach their end of life
- ✓ Technologies and processes that will be used to destroy data on hard drives, SIM cards, and removable media when equipment reaches end of life (which might involve ITAD firms and other third parties)
- ✓ Procedures to follow when customers and other third parties request that their data be destroyed

As we discussed in Chapter 3, many “data erasure” technologies don’t really do the job. Your policy needs to be explicit on requirements. For example, they should specify that data be completely overwritten with enough passes to comply with industry standards (as well as the security needs of your organization). Regulations with explicit data erasure requirements include:

- ✓ HIPAA (Health Insurance Portability and Accountability Act)
- ✓ EU GDPR
- ✓ ISO 27001

Define roles and responsibilities

It is important to clarify the roles and responsibilities of different groups for tasks like:

- ✓ Defining and refining data retention policies
- ✓ Classifying and protecting files
- ✓ Ensuring the destruction of files

- ✓ Handling requests from customers and third parties
- ✓ Responding to litigation and discovery requests
- ✓ Monitoring retention and destruction processes and documenting the results for audit purposes and to make improvements

Enforcing Data Retention Policies

Enforce data retention policies

To ensure that data retention policies are applied consistently, as much as possible automate processes such as file classification and protection.

Data classification products are available for scanning documents and emails and assigning them to categories based on factors like file type, creator, location, keywords and tags. You may also be able to leverage capabilities already in place. For example, you can use an enterprise directory such as Microsoft Active Directory to manage permissions, giving you control over exactly who can access, create and delete files in specific network folders.

Enforce data erasure policies

You can automate data erasure tasks. For example, you can program Active Directory and systems management tools to erase temporary files, clean out recycle bins and shred free disk space on employee laptops and desktop computers.

Data erasure solutions are available that use advanced techniques to guarantee secure erasure of files across servers and storage environments, laptops and desktop computers, mobile devices, and virtual machines in cloud data centers. We will review some of these capabilities in the next chapter.

Monitor and report

It is important to monitor and report on data retention and erasure activities, both to satisfy auditors and regulators and to collect data to improve these activities. The information collected and reported should include details including:

- ✓ The classification category of each file, the reason for selecting its category and the retention period
- ✓ Where retained files were stored initially, and where they moved over time
- ✓ When files were erased, the method used and the reason for erasure
- ✓ Who performed or authorized each action
- ✓ Exceptions and failures to apply policies

Review and refine

Data retention policies, and the processes to enforce them, should be reviewed and refined at regular intervals. The data retention team should discuss new legal, regulatory and business requirements, and what adjustments are needed to address them.

The team should also conduct its own audits and spot tests to ensure that procedures are effective and produce the expected results.

CHAPTER 5

Selecting the Right Partners

"If you reveal your secrets to the wind, you should not blame the wind for revealing them to the trees."

– Kahlil Gibran

Organizations that want to improve and expand their data retention programs can benefit from outside partners. Here we make some suggestions for selecting the right advisors and technology providers.

Technology Partners for Data Erasure

The two types of technology partners enterprises engage most frequently for data erasure are ITAD (IT Asset and Disposition) companies and data erasure software solution vendors.

ITAD companies specialize in the secure disposition, and sometimes brokering, of equipment that has outlived its usefulness for the organization.

For organizations that dispose of a lot of equipment, ITAD companies can offload a tremendous number of headaches and produce some cash in the bargain. However, a security failure by the ITAD can result in a major data breach for the original owner of the equipment. Organizations must therefore select an ITAD only after verifying that it has solid processes, including the use of advanced data erasure technology with audit-ready reporting.

Enterprises should also have a data erasure software solution to use internally. Selection criteria include:

- ✓ Technologies that guarantee secure erasure, such as multiple random overwrites

- ✓ Ability to work with a very wide range of systems, including servers, laptops, PCs, and iOS, Android and Windows mobile devices

- ✓ Ability to work with many storage environments and disk types, including RAID systems, solid-state drives, virtual machines and LUNs (logical unit numbers)

- ✓ Features for local and remote data erasure

- ✓ Detailed reporting and comprehensive audit trails

- ✓ High performance and the ability to erase multiple devices in a short time

- ✓ Certification by major government and industry standards bodies

Recap

Law professor Jeffrey Rosen has said: "Privacy is not for the passive," and now you understand why the same statement applies to data retention and data erasure.

Today, organizations need well-designed programs and policies not only to deal with regulations mandating that specific document types be held for set periods, but also to address critical privacy issues and to reduce the cost of potential data breaches.

We have made a number of points in this guide that can help you address these issues, including:

- ✓ Data protection is an inseparable part of data retention and security measures must be applied across the entire information management lifecycle.

- ✓ "Just in case, because storage is cheap" is a terrible guide to policy – there are legal, security and other risks associated with retaining files that aren't needed.

- ✓ Privacy concerns, especially the EU GDPR, are forcing major changes because now outside parties like customers can trigger data erasure processes.
- ✓ Many organizations overlook some of the factors that should trigger data erasures and rely on ineffective erasure technologies.
- ✓ Every organization needs a cross-functional data retention team, ideally lead by a full-time or part-time DPO (Data Protection Officer).
- ✓ A data retention policy needs to define how data is going to be classified, how data will be retained and protected, when and how data will be erased and the roles and responsibilities of the team members.
- ✓ Automated tools and processes are essential for providing consistent, reliable enforcement of data retention and data erasure policies.
- ✓ Organizations can accelerate the upgrading of their data retention programs by using third party advisors and technology partners.

We hope that this guide has provided some practical advice and stimulated your thinking about data retention, privacy and data erasure.

About Blancco

Blancco is the industry standard in data erasure and mobile device diagnostic. Blancco data erasure solutions provide thousands of organizations with the tools they need to add an additional layer of security to their endpoint security policies through secure erasure of IT assets. All erasures are verified and certified through a tamper-proof audit trail.

Blancco data erasure solutions have been tested, certified, approved and recommended by 15+ governing bodies and leading organizations around the world. No other data erasure software can boast this level of compliance with the rigorous requirements set by government agencies, legal authorities and independent testing laboratories.

Blancco Mobile Diagnostics solutions enable mobile network operators, retailers and call centers to easily, quickly and accurately identify and resolve performance issues on their customers' mobile devices. As a result, mobile retailers can spend less time dealing with technical issues and, in turn, reduce the quantity of NTF returns, save on operational costs and increase customer satisfaction.

Additionally, Blancco Mobile Diagnostic solutions empower mobile processors, 3PLs, Recyclers and Repair & Refurbishment Operations to easily, quickly and accurately process used mobile devices to identify any issues and determine overall value. By incorporating Blancco Mobile Diagnostics, mobile processors automate processes, deliver intelligent routing based on device attributes and increase overall efficiency, while driving incremental revenue and profitability.

For more information, visit our website at www.blancco.com.

Contact Us

For Marketing, Please Contact:
Email: marketing@blancco.com

For Corporate Communications & PR, Please Contact:
Email: press@blancco.com