

Insights on IT risk
Business briefing
October 2011


Data loss prevention

Keeping your sensitive data
out of the public domain



Contents

Introduction	1
Understanding the problem.....	2
Challenges	6
Employing a holistic approach	10
Data governance	13
Data loss prevention controls.....	16
Supporting information security processes.....	17
Using technology to support the DLP program.....	18
Ernst & Young insights and lessons learned	20
Don't be a victim	21



Data loss prevention (DLP) is the practice of detecting and preventing confidential data from being “leaked” out of an organization’s boundaries for unauthorized use. Data may be physically or logically removed from the organization either intentionally or unintentionally.

Introduction

Over the last few years, companies in every industry sector around the globe have seen their sensitive internal data lost, stolen or leaked to the outside world. A wide range of high-profile data loss incidents have cost organizations millions of dollars in direct and indirect costs and have resulted in tremendous damage to brands and reputations. Many different types of incidents have occurred, including the sale of customer account details to external parties and the loss of many laptops, USB sticks, backup tapes and mobile devices, to name just a few. The vast majority of these incidents resulted from the actions of internal users and trusted third parties, and most have been unintentional.

As data is likely one of your organization’s most valuable assets, protecting it and keeping it out of the public domain is of paramount importance. In order to accomplish this, a number of DLP controls must be implemented, combining strategic, operational and tactical measures.

However, before DLP controls can be effectively implemented, your organization must understand the answer to these three fundamental questions:

1. What sensitive data do you hold?
2. Where does your sensitive data reside, both internally and with third parties?
3. Where is your data going?

This paper explores these questions and the challenges organizations face in relation to business drivers and regulatory obligations for protecting this data. We will share our point of view and approach to data loss prevention, along with insights and lessons learned from our experiences working with some of the most advanced companies in the world on data loss prevention practices.



Understanding the problem

Common data loss vectors

- ▶ Email
- ▶ Webmail
- ▶ Instant messaging
- ▶ File transfer protocol
- ▶ Blogs
- ▶ Social media
- ▶ Web pages
- ▶ Removable media
- ▶ Cameras
- ▶ Hard copy

Recent highly publicized events, such as the leaking of government and corporate data to Wikileaks and the sale of customer banking records to tax authorities, have demonstrated that it is more difficult than ever to protect your organization's internal data. Advances in technology and productivity tools have made collaboration in the workplace easier, while also creating new vectors for data to leave the organization. Likewise, business demands to embrace new technologies such as social media and mobile devices have made it impossible for most organizations to simply build and rely on a strong perimeter for adequate protection.

Economic pressures on individuals and the monetization of data on the black market have created an environment where people with access to information can convert data into cash. Employees also find the lines between personal and business system use blurred in the modern workplace, resulting in many situations where users unintentionally leak internal data.

In the context of this document, data loss is the extraction and/or dissemination of sensitive data of an organization that can intentionally or unintentionally put an organization at risk. The term "data leakage" is also commonly used to refer to the same idea.

The changing data loss risk landscape

In addition to obvious data loss methods such as the loss of physical assets such as laptops, many data loss incidents are due to accidental disclosure through electronic transmissions. In most cases, end users do not realize the risks associated with sending sensitive data through unencrypted emails, instant messages, webmail and file transfer tools.

Technological development has caused data volumes to rise rapidly, and the increased use of mobile devices heightens the risk that unauthorized parties could gain access to sensitive data. The embedding of technological user-friendliness and access to data has become so intertwined that it has become relatively easy to engage in the unintentional spreading of confidential data.

The current use of information technology and the internet has increased the capabilities and connectivity of users and is constantly evolving. This evolution is constantly increasing the IT risk spectrum. IT risks are impacted heavily by a number of significant trends – so-called megatrends.

Wikileaks and internal security

The recent exposure of Wikileaks-related incidents has shown that internal security is at least as important as external threats. In one incident, a disgruntled (ex)-employee of a Swiss bank handed over the bank account data of more than 2,000 prominent individuals to Wikileaks, potentially exposing tax evasion. This incident emphasizes once more that employees with access to critical, restricted information can put organizations at risk by disclosing the information to the public. This risk has recently been fueled by a rise in rogue or disgruntled employee behavior as a consequence of the financial crisis, or from a sense of acting in the public interest. In practice, many firms are struggling with providing the right access to information to the right people in their organizations.



For a better understanding of the way to address IT Risk and developing an effective IT Risk management function, please refer to Ernst & Young's Insights on IT Risk paper, *The evolving IT risk landscape*, published in June 2011.

An overview of recent megatrends included in this paper shows that data protection will continue to be a significant challenge for organizations. Four out of six megatrends discussed are linked to the risk category "data," highlighting the fact that many of the technology trends observed in the market result in increasing data risk.

Megatrend	Business benefit	Business/IT risks	Categories of IT Risk Universe affected
Emerging consumerization	<ul style="list-style-type: none"> ▶ <i>Mobile computing</i>: Anytime and anywhere connectivity/ high-volume portable data storage capability ▶ <i>Social media</i>: New and advanced information sharing capabilities such as crowdsourcing 	<ul style="list-style-type: none"> ▶ Increased vulnerability due to anytime, anywhere accessibility ▶ Risk of unintended sharing, amplification of casual remarks and disclosure of personal and company data. The availability of this data on the web facilitates cyber attacks. ▶ Employees may violate company policies in terms of data leakage 	<ul style="list-style-type: none"> ▶ Security and privacy ▶ Data ▶ Legal and regulatory ▶ Infrastructure
The rise of cloud computing	<ul style="list-style-type: none"> ▶ Lower total cost of ownership ▶ Focus on core activities and reduction of effort spent on managing IT infrastructure and applications ▶ Contribute to reduction of global carbon footprint 	<ul style="list-style-type: none"> ▶ Lack of governance and oversight over IT infrastructure, applications and databases ▶ Vendor lock-in ▶ Privacy and security ▶ Availability of IT to be impacted by the use of the cloud ▶ Increased risk to regulatory noncompliance (e.g., SOX, PCI). The cloud also brings about challenges in auditing compliance. ▶ The cloud may impact the agility of IT and organizations; the platform dictated by the provider may not align with software development and strategic needs of the user 	<ul style="list-style-type: none"> ▶ Security and privacy ▶ Data ▶ Third-party suppliers and outsourcing ▶ Applications and databases ▶ Infrastructure ▶ Legal and regulatory
The increased importance of business continuity	<ul style="list-style-type: none"> ▶ 24/7/365 availability of IT systems to enable continuous consumer support, operations, e-commerce, and other functions 	<ul style="list-style-type: none"> ▶ Failure of the business continuity and disaster recovery plans causing financial or reputational loss 	<ul style="list-style-type: none"> ▶ Infrastructure ▶ Applications and databases ▶ Staffing ▶ Operations ▶ Physical environment
Enhanced persistence of cybercrime	<ul style="list-style-type: none"> ▶ N/A 	<ul style="list-style-type: none"> ▶ Spread of malicious code in company systems, causing system outages ▶ The risk of theft of personal, financial and health information ▶ Loss of confidential data due to external vulnerabilities ▶ Financial loss due to unauthorized wire transfers 	<ul style="list-style-type: none"> ▶ Security and privacy ▶ Data
Increased exposure to internal threats	<ul style="list-style-type: none"> ▶ N/A 	<ul style="list-style-type: none"> ▶ Assigning access rights that are beyond what is required for the role by employees or contractors ▶ Failure to remove access rights to employees or contractors who leave the organization 	<ul style="list-style-type: none"> ▶ Data ▶ Applications and databases
The accelerating change agenda	<ul style="list-style-type: none"> ▶ Fast adoption of new business models or reducing costs provides organizations with competitive advantage 	<ul style="list-style-type: none"> ▶ Failure to deliver IT projects and programs within budget, timing, quality and scope causing value leakage 	<ul style="list-style-type: none"> ▶ Programs and change management



The rising cost of data loss incidents

According to a 2010 Ponemon Institute study, the average total cost per data breach has risen to \$7.2 million, or \$214 per record lost. The Forrester Research institute calculated the cost per record as shown in the table below.¹

In addition to the costs of incidents increasing, the number of leaks appears to be increasing year on year. We expect this to continue to increase significantly over the next several years.

But good statistics on this phenomenon are very hard to get, and the figures available will never represent the actual situation because many more leaks and data breaches go unreported. There is not a finite number that can be reported with certainty, because there is no single repository for incident tracking and these statistics only include incidents that reach the media or are self-reported by companies. The only certainties are that leakage happens and that it is a growing problem. It is not only the number of incidents but also their magnitude and the increased attention of the general public that increase the impact of an incident for any organization.

The cost of a breach, broken out for three sample companies

Category	Description	Cost per record		
		Company A: Low-profile breach in a non-regulated industry	Company B: Low-profile breach in a regulated industry	Company C: High-profile breach in a highly regulated industry
Discovery, notification and response	Outside legal counsel, mail notification, calls, call center and discounted product offers	\$50	\$50	\$50
Lost employee productivity	Employees diverted from other tasks	\$20	\$25	\$30
Opportunity cost	Customer churn and difficulty in getting new customers	\$20	\$50	\$100
Regulatory fines	FTC, PCI, SOX	\$0	\$25	\$60
Restitution	Civil courts may ask to put this money aside in case breaches are discovered	\$0	\$0	\$30
Additional security and audit requirements	The security and audit requirements levied as a result of a breach	\$0	\$5	\$10
Other liabilities	Credit card replacement costs. Civil penalties if specific fraud can be traced to the breach	\$0	\$0	\$25
Total cost per record		\$90	\$155	\$305

Source: The Forrester Wave™: Information Security and Risk Consulting Services, Q3 2010, Forrester Research, Inc., 2 August 2010.

¹ *Calculating The Cost Of A Security Breach*, Khalid Kark, Forrester Research Inc, April 10, 2007.



Several high-impact incidents have occurred recently that have resulted in high costs and extreme media attention for the affected companies:

Overview of important incidents

Web technology firm	On its official weblog, a web technology firm published a message that it had uncovered a ploy to collect user passwords, likely through phishing. This ploy affected the personal accounts of hundreds of users, including among others, senior US Government officials, Chinese political activists, officials in several Asian countries (predominantly South Korea), military personnel and journalists.
Public health corporation	A public health corporation had to notify 1.7 million patients, staff, contractors, vendors and others about a reported theft of electronic record files that contained their personal information, protected health information or personally identifiable employee medical information. The information included Social Security numbers, names, addresses and medical histories.
International oil and gas company	An international oil and gas company lost a laptop which contained personal information for 13,000 individuals including names, Social Security numbers and addresses. The laptop was not encrypted, and the information lost was for claimants against the company.
US public agency	Personal details for 3.5 million teachers and other employees of a US public agency were accidentally published on the Internet. Information released included names, Social Security numbers and birthdates. This data had been posted on the internet for more than a year without the organization realizing it.
National retail bank	Two thousand customer records from a national retail bank were stolen by employees prior to leaving and joining a competitor firm. Records included customer bank account numbers, Social Security numbers and other highly sensitive personal data such as tax returns and pay statements.
Online storage provider	According to a blog post, an online storage provider explained that due to an authentication bug, all accounts were at risk of a data breach. As soon as the bug was discovered, as a precaution all logged in sessions were disconnected. The bug was active for almost four hours and took five minutes to fix.

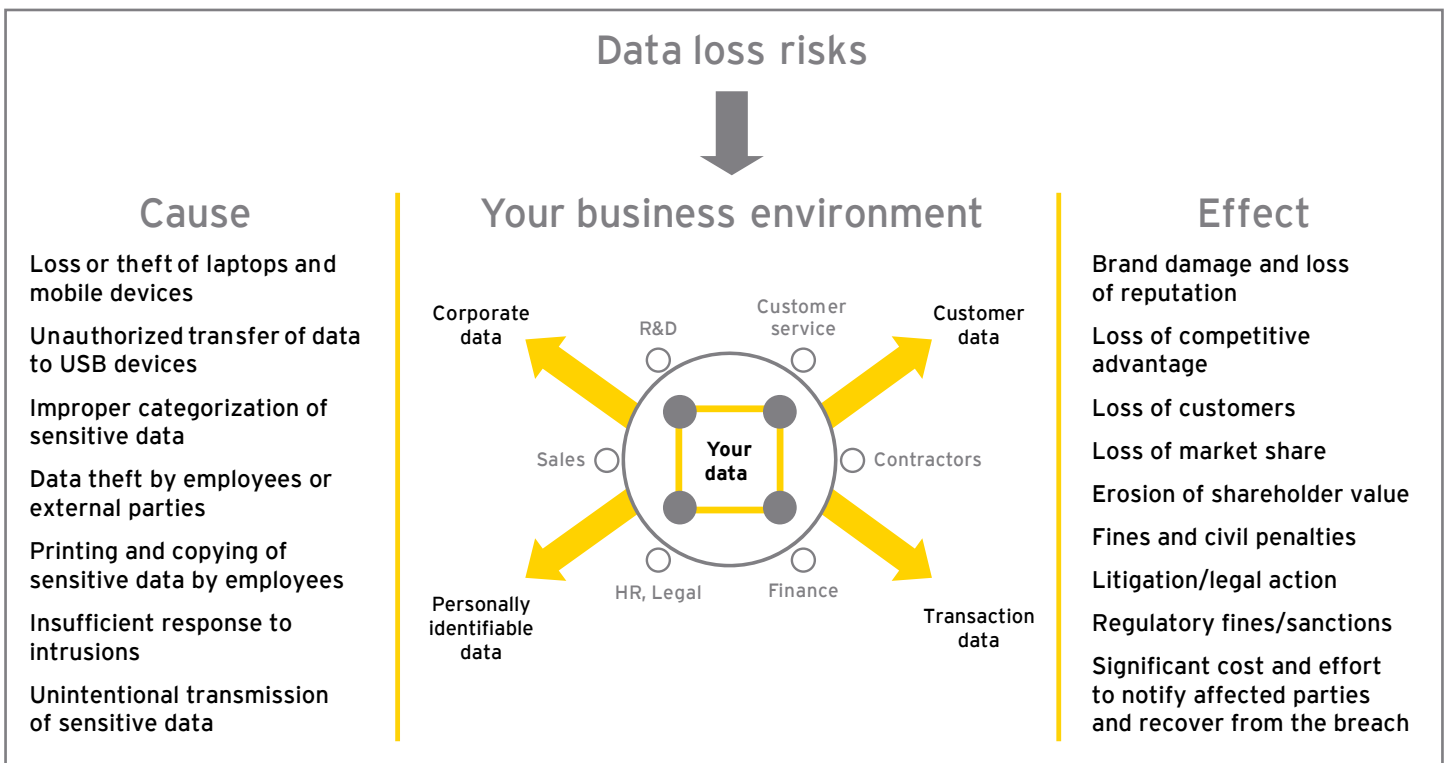
So, what is new? Threats of data loss from internal users have always been a risk. To sum up the changing landscape and increasing risk:

1. There are now many more ways data can leave an organization (i.e., data loss vectors).
2. Storage is cheap. Many gigabytes of data can walk out of the door on an employee's keychain or smartphone or be sent through online systems such as Dropbox.
3. Data is everywhere. Decentralized systems and work collaboration tools make it much more difficult for organizations to track and control information within the business.
4. Data has value in the real world, including from seemingly legitimate sources.
5. The most recent generation of workers to join companies has grown up with openness and information sharing as a cultural norm.
6. It is easier than ever for data to cross borders, and demand for sensitive information is coming from all over the world as companies (and nations) try to gain competitiveness in the global marketplace.
7. The sheer volume of data is increasing as never before.

In addition to business risks, regulatory risks are also increasing. The volume, impact and visibility of incidents has resulted in renewed focus from regulators. Data protection requirements, particularly breach notification rules, for organizations are becoming more strict, and enforcement penalties are on the rise. From a company's perspective, reducing the risk of data loss reduces regulatory risk and helps to protect the company's brand, strategic business data and intellectual property.

Challenges

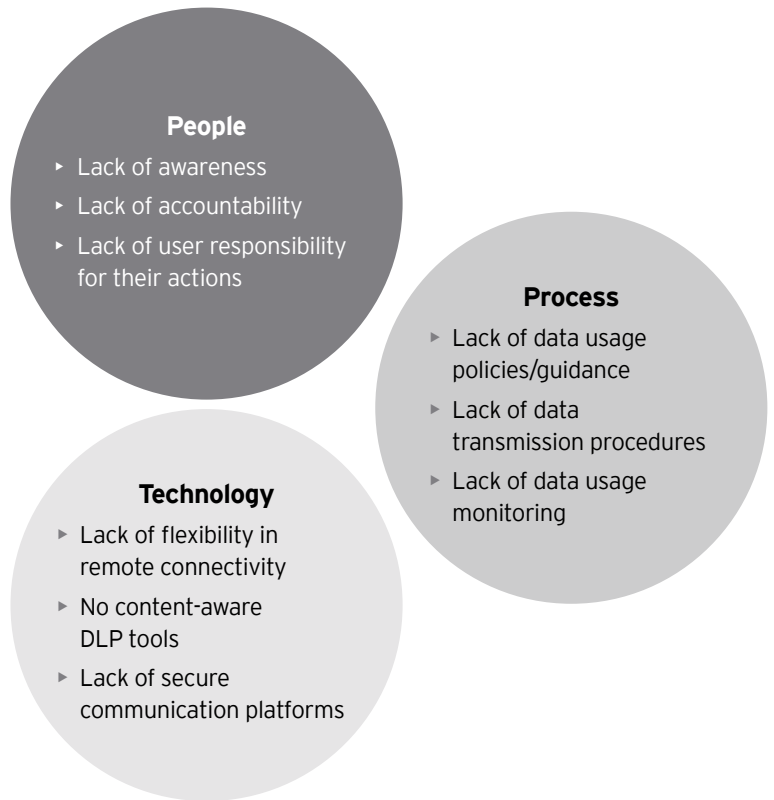
From our experience, one of the greatest challenges in managing data loss is that there are so many reasons why data loss can occur, numerous data loss scenarios to account for and many different controls that must be effective in order to manage the problem. There is no simple solution or tool that can be implemented to address the variety of data loss risks that organizations face. In order to address the pervasive issue that data loss risks pose, a comprehensive solution that includes people, processes and technology needs to be implemented.





Common unintentional data loss themes

In the data loss assessments we have performed, common, recurring root causes for data loss have become evident. These themes often capture the reasons data loss occurs, particularly unintentional data loss.



People	Process	Technology
<p>Employees do not clearly understand or feel accountable for the protection of sensitive data.</p>	<p>Data protection, data classification and acceptable use policies do not clearly articulate:</p> <ul style="list-style-type: none"> ▶ The controls that should be implemented for securely sending sensitive data to third parties ▶ Whether employees may send sensitive data to home computers and personal email accounts ▶ The specific data that is considered sensitive and requires data protection controls 	<p>Current remote access tools are not flexible enough to support the business, resulting in users employing alternative approaches, such as emailing documents to their personal email accounts, to enable working from home and remote locations.</p>
<p>Training and awareness programs do not focus enough on protecting sensitive data, appropriate use of email and the internet, use of security tools such as file encryption and each employee's personal responsibility for complying with information security/data protection policies.</p>	<p>Process owners have not assessed their methods in which sensitive data is shared with third parties to evaluate information security risks.</p>	<p>Content-aware email encryption tools are not effectively used to automatically require encryption of emails containing sensitive data, such as account numbers.</p>
<p>Employees feel that there is no risk involved in breaking the rules (i.e., "no one is watching so I will not be caught").</p>	<p>Without an ongoing DLP monitoring program, policy violations cannot be identified efficiently, and the success of policy communications, training and awareness programs and technical controls is not measurable.</p>	<p>Secure links between the company and its third parties are not in place to enable encrypted email or other secure transmission methods.</p>



Examples of malicious activity

Below is a summary of actual cases we have investigated to help illustrate the risks that organizations are facing from malicious users. The challenge facing organizations is how to control the actions of trusted internal users while still allowing those users to perform their jobs. These examples demonstrate that total prevention of data loss is very difficult. However, effective risk management activities can bring data loss risks to acceptable levels. It is also notable that not every case requires sophisticated procedures to be carried out by IT specialists. Simple screenshots or pictures taken from a camera phone can be used for data theft.

Root cause for data loss	Data category	Case description
Inappropriate access rights to applications with sensitive data	Customer data	A frustrated staff member used the standard data export procedures to export sensitive data and copied it to a CD.
Exploitation of weaknesses in a database's development environment	Personally identifiable data	A database administrator with an understanding of test procedures was able to reverse engineer a sanitized process by referencing hidden tables.
Breach of trust between developers	Transaction data	An experienced IT developer was able to reconstruct transaction data by gaining access to confidential data from an inexperienced developer unaware of the company's access policies and restrictions.
Unsupervised front office	Corporate data	A call center staff member provided screenshots of internal systems to fraudsters to help them reverse engineer an application.
Employee discontent	Corporate data	An employee leaving the company came in over a weekend prior to resigning on Monday, accessed the customer master file and exported it to an Excel file. This file was then emailed to the employee's personal email account.
Insider trading	Corporate data	An employee with access to prereleased financial information fed information to an external analyst, resulting in improper stock trades for both the employee and the analyst.



Growth in data volume and importance to the business

Organizations are increasingly investing in the collection and management of data to improve business performance and leverage their competitive advantage. As most business communications between employees and with third parties now take place electronically, the potential for sensitive data exposure continues to rise. According to IDC's 2010 Digital Universe Study: *Digital Universe Decade - Are You Ready?*, digital information will grow by a factor of 44 globally over the next 10 years to a total of 35 zettabytes (35 trillion gigabytes!). As data production increases, so will the number of data loss incidents. We expect a continued increase in accidental data loss due to the misuse of technology available to users, including social media and new tools that have not even been invented yet.

User-driven cross-border data transmissions

The global adoption of digital communications has resulted in an environment where data is flowing around the world, often without users even realizing it. Webmail providers and online storage systems in particular often have data centers located in countries across the globe, resulting in users having no idea of where their data is actually stored. Assessments we have performed have shown patterns of internal users sending business data, including sensitive data, to their personal email accounts to allow them to work from home. Users are also sending data to personal devices such as iPads so they can work on the device of their choice. In these scenarios, we have noted many examples where users have sent data to "themselves" and in the process unknowingly sent their sensitive business data to servers managed by third parties in foreign countries.

DLP and privacy

Data loss is a particularly stressful problem for global companies that hold a large volume of personally identifiable information in different legal jurisdictions. For global companies, there are a host of privacy issues to deal with that not only differ by country, but even within regions, states and by line of business. In Europe, companies must understand local laws enacting the EU Data Protection Directive. While the common principles are consistent, interpretation and enforcement varies significantly by country.

In the United States, privacy laws vary by data type and state. For example, the US Health Insurance Portability and Accountability Act (HIPAA) applies to patient health information, and states such as Massachusetts have enacted rules covering all companies that hold personally identifiable information of state residents. Outside of the US and EU, the picture does not get any simpler. Some countries have no privacy regulations, some have adopted legislation consistent with the EU, and some have come up with unique requirements. It all adds up to a very confusing picture for privacy and compliance officers of global companies.

When implementing a DLP tool, accessing data that may have privacy implications for individuals is unavoidable. In order to evaluate incidents, organizations must be able to view the sensitive data being transmitted. For companies operating in multiple jurisdictions, it is good practice to consult with professionals with local regulatory experience to ensure compliance with local privacy regulations before implementing DLP tools and processes.



Employing a holistic approach

An effective DLP program requires an understanding of:

- ▶ What data you have
- ▶ The value of your data
- ▶ Your obligations for protecting that data
- ▶ Where your data resides
- ▶ Who is accessing your data
- ▶ Where your data is going
- ▶ How you protect your data
- ▶ Your protection gaps and risks
- ▶ How to respond to data leakage events

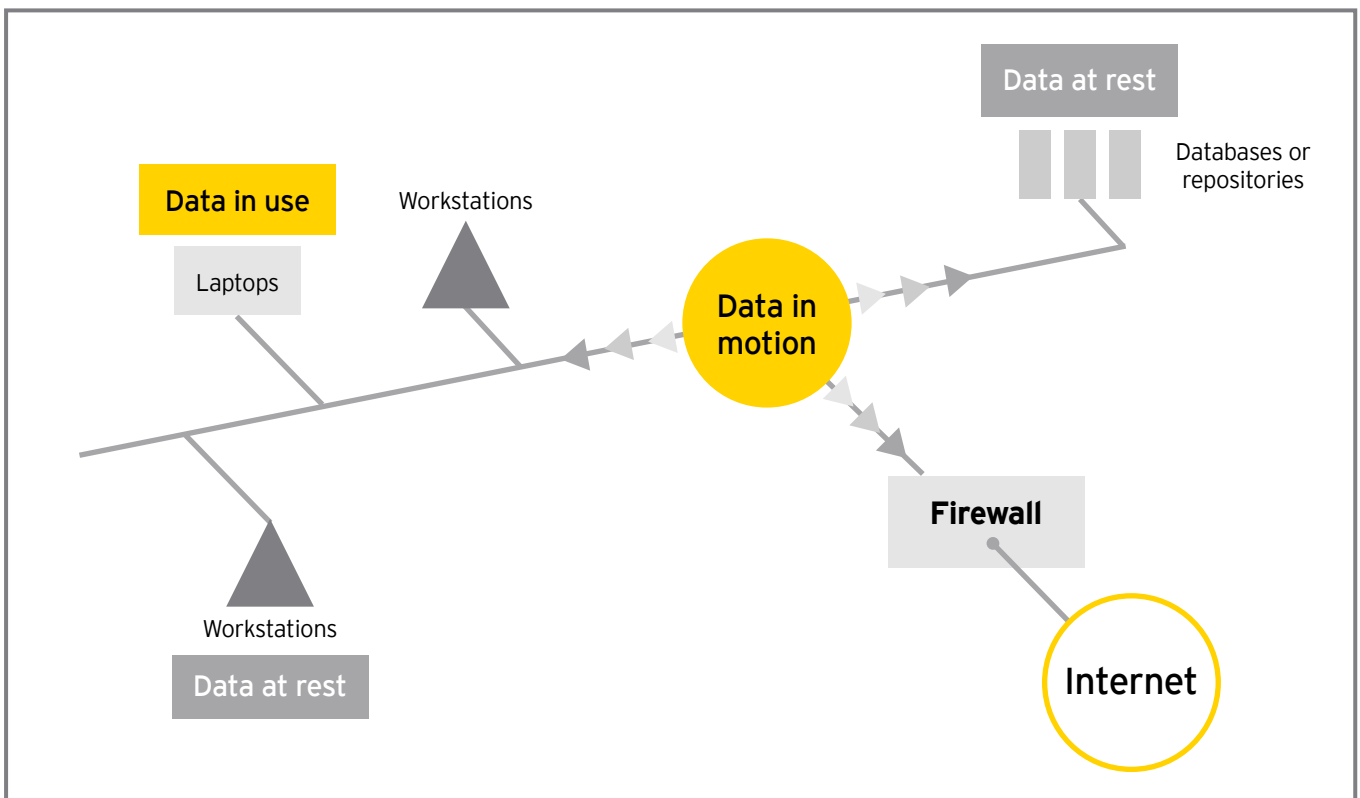
What can you do to prevent data loss?

To effectively manage data loss risks, clear business objectives should be defined to drive the DLP program. These objectives should cover the following items as a minimum:

- ▶ Prevent the intentional or unintentional disclosure of sensitive data at rest, in use or in motion to unauthorized parties
- ▶ Maintain adequate security and provide usability
- ▶ Protect customer data and brand reputation
- ▶ Protect personally identifiable information and intellectual property
- ▶ Reduce the organization's risk and cost of compliance

With your program objectives defined, you can develop a strategy and implement DLP controls to achieve your objectives. One of the key challenges to securing your critical data is the fact that there are so many ways for it to leave. In developing your DLP strategy, a holistic view should be taken to ensure that the combination of controls employed is geared to protect the most sensitive data that the organization holds.





The data lifecycle

From a data loss perspective, the industry has adopted three standard terms related to the states in the data lifecycle:

- ▶ **Data at rest** is data that is stored within the IT infrastructure and on media. Common components containing data at rest are servers, databases, file shares, intranet sites, workstations, laptops, mobile devices, portable storage, backup tapes, and removable media. Data at rest can also be stored externally with third parties or through external extensions of the IT infrastructure, such as cloud storage.
- ▶ **Data in motion** is data that is in transit, flowing across internal networks and to the outside world (i.e., data on the wire and in the air).
- ▶ **Data in use** is data that is being accessed or used by a system at a point in time. Examples include data in temporary memory on a local machine, an open report or running query on a workstation, an email that has been drafted but not sent, a file being copied to a USB drive, and data being copied and pasted from one local document to another.

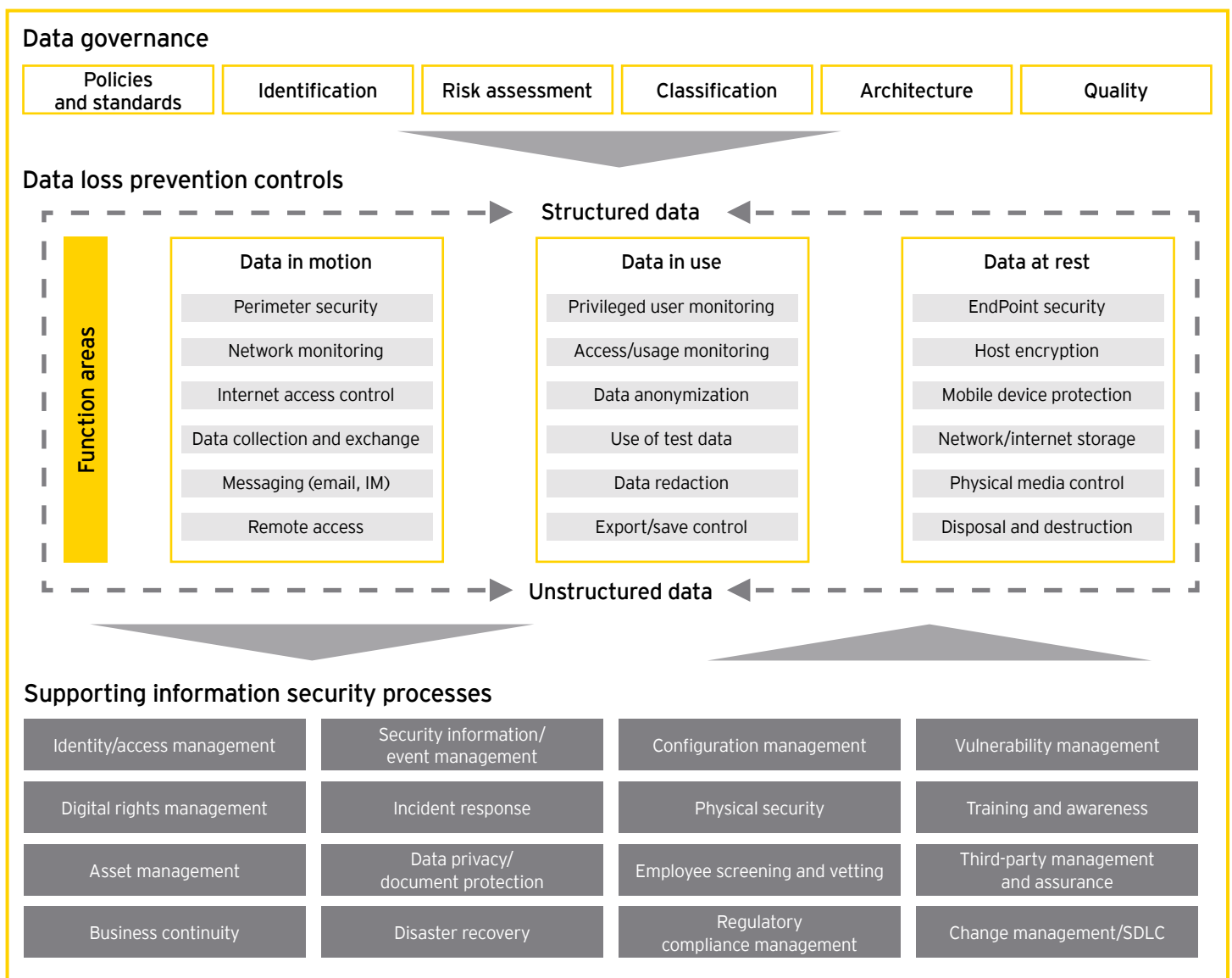


DLP conceptual model

The following diagram presents a holistic view of data security with a focus on data loss prevention controls. In this model, data governance ensures alignment to the business and drives DLP control requirements. DLP controls link in to the numerous information security processes that must be effective to ensure the confidentiality, integrity and availability of data within the organization.

Throughout the remainder of this section, we will explain how this model can be used to help you think through your DLP strategy and risks – tackling the three components of the model:

- ▶ Data governance
- ▶ Data loss prevention controls
- ▶ Support for information security processes





Data governance

Data governance activities help to answer some of the key questions that must be addressed prior to implementing DLP controls, including:

1. What sensitive data do you hold – what is your most important data?
2. Where does your sensitive data reside, both internally and with third parties?
3. Where is your data going?

What is your most important data?

Data classification and risk assessment

Data must be classified from a business perspective. This means that in addition to applying labels such as confidential, top secret, sensitive and public, an organization must understand the specific types of data that it holds so that DLP controls can be customized for its specific business needs.

In general, there is no one-size-fits-all description of what your most important data is, because this depends largely on the type of organization you are. However, by considering what data is vital to the core of your operations, it becomes clear which data you cannot afford to lose.

When classifying data from a DLP perspective, for each data type, you must consider the consequences of this data becoming available to unauthorized parties (e.g., business disruption, damage to reputation and brand, regulatory violations and fines, loss of competitive advantage, direct financial losses). The following categories depicted below are a good starting point for identifying common types of sensitive data.

Additionally, it is helpful to assess the risk of the relevant data types that you identify. Examples of considerations for assessing the risk of each data type include:

- ▶ Whether the data is protected by regulations
- ▶ Relative value of internal data (e.g., board papers versus corporate customer lists)
- ▶ Direct impact to customers and business partners
- ▶ Potential impact on brand and reputation
- ▶ Potential loss of competitive advantage in the market

Going through this exercise will help you to prioritize DLP activities so that the highest risk data is protected first.

Corporate data	Transaction data	Customer data	Personally identifiable data
<ul style="list-style-type: none"> ▶ Price/cost lists ▶ Target customer lists ▶ New designs ▶ Source code ▶ Formulas ▶ Process advantages ▶ Pending patents ▶ Intellectual property ▶ Unreleased merger/acquisition plans and financial reports ▶ Legal documents ▶ Employee personal data 	<ul style="list-style-type: none"> ▶ Bank payments ▶ B2B orders ▶ Vendor data ▶ Sales volumes ▶ Purchase power ▶ Revenue potential ▶ Sales projections ▶ Discount ratios 	<ul style="list-style-type: none"> ▶ Customer list ▶ Spending habits ▶ Contact details ▶ User preferences ▶ Product customer profile ▶ Payment status ▶ Contact history ▶ Account balances ▶ Purchase/transaction history ▶ Payment/contract terms 	<ul style="list-style-type: none"> ▶ Full name ▶ Birthday, birthplace ▶ Biometric data ▶ Genetic information ▶ Credit card numbers ▶ National identification number, passport numbers ▶ Driver's license number, vehicle registration number ▶ Associated demographics ▶ Preferences



Where does your sensitive data reside?

Identification and architecture

Depending on the way you conduct business, data can reside anywhere in your IT infrastructure. Sensitive data may be stored on servers, workstations, shared through the internal network, sent out through the internet, carried around on laptops and mobile devices and stored in databases, file shares and the cloud. This means that even within your organization, your data is virtually everywhere within the IT infrastructure.

Once you have determined what your most sensitive data is, you can start to determine where it sits within your IT infrastructure. There are two high-level ways data is stored within an organization:

- ▶ **Structured repositories:** Data is organized into structured repositories, such as relational databases, that are typically supported and controlled by the IT organization.
- ▶ **Unstructured repositories:** Data is generally end-user driven and stored in less controlled repositories such as network shares, SharePoint sites and workstations.



Strategies for identifying sensitive structured data in each type of location include working with the business and IT to identify repositories that hold sensitive data, as defined through the data classification exercise. This can be accomplished through a combination of:

- ▶ Business process walk-throughs
- ▶ Use of questionnaires sent to business process owners, analysts, database administrators, application developers, business intelligence/report developers and other relevant parties.
- ▶ Reviews of existing documentation (e.g., data flow diagrams, application descriptions, information held in IT asset inventories).

Unstructured sensitive data, by definition, will reside in unpredictable locations, as well as on commonly known network shares, internal sites and end-user repositories. As is the case with structured data, inquiry of business users and IT personnel can provide useful information about commonly used data stores. However, due to the nature of unstructured data, these sources are not likely to be complete. For this reason, the use of tools should be considered to assist in data discovery. DLP discovery tools have advanced significantly over the last two years and can provide the following features that aid in data discovery:

- ▶ Scanning of known network shares, intranet sites, wikis and databases with different levels of access privileges
- ▶ Scanning of network segments to identify undocumented share drives, databases and servers
- ▶ Scanning of user workstations to identify sensitive data stored on local drives

These activities can be performed using rules designed to detect sensitive data. Rules should be customized for your organization based on the high-risk data types identified. Through the use of these tools, repositories containing sensitive data can be identified and then steps can be taken to ensure that these repositories are secure or to move sensitive data to more appropriate locations.



Where is your data going?

Identification and architecture

Understanding where you hold sensitive data internally is essential, but it does not provide a complete view of where your data resides. Organizations must also understand where data is going and how it flows from the inside out. Understanding data flows is a complicated task, and technology can help to provide transparency into network based data transfers. However, you must additionally understand third-party data access, business-driven data exchange and end-user data transmission capabilities.

Third-party data access

Data identification should include obtaining an understanding of sensitive data that is accessible by third parties. This includes data that is exchanged with third parties and third parties that have direct access to internal systems. Once an inventory of third parties with access to sensitive data exists, controls should be implemented to safeguard third-party access. Focus areas should include:

- ▶ Secure data transmissions
- ▶ Controlled access to company networks and data
- ▶ Monitoring of third-party access to company resources
- ▶ Third-party due diligence/information security assurance

What about policies and standards?

Once sensitive data is classified and identified, data protection policies should be developed and/or customized to document security requirements that are specific to the types of sensitive data held by the organization. A high-level policy specifying the requirements for protecting sensitive data should exist and clearly link to the data classification policy.

Detailed DLP standards can then be developed to further document requirements for protecting the organization's sensitive data. It is also common for organizations to embed DLP concepts into existing policies covering topics such as logical security and the acceptable use of email. In any case, policies must make data protection requirements clear to users, as most breaches directly relate to end user activities.

DLP concepts that should be documented in policies include:

- ▶ Transmission of sensitive data through email and the internet
- ▶ Storage of sensitive data on mobile devices, laptops, workstations and non-company owned equipment
- ▶ Storage of sensitive data on company file and document repositories (where it is acceptable and not acceptable to store sensitive data)
- ▶ Appropriate use of remote access technologies
- ▶ Use of technology not provided by the organization (such as work use of personal email accounts, portable devices, storage and media)
- ▶ User responsibilities for classifying data at the point of creation and ensuring that sensitive data users create is included in relevant data/information inventories

In addition, DLP principles should be used to drive security requirements in system development and change projects. Example principles include:

- ▶ Sensitive data may not be transmitted through public networks without adequate encryption
- ▶ Only company-approved technologies may be used to exchange data with third parties
- ▶ Access to sensitive data must be logged and monitored where appropriate
- ▶ Access to sensitive data stored on information systems must be restricted to those who require it to perform their job responsibilities
- ▶ Sensitive data may not be shared with third parties without sufficient contracts in place specifying information security requirements, their obligations to protect company data, their responsibilities for monitoring their own third parties and the company's right to audit and monitor
- ▶ Sensitive data must be anonymized before being stored in less controlled environments, such as test and development environments
- ▶ Sensitive data must be adequately protected through all stages of the data lifecycle and the systems development lifecycle (SDLC)



Data loss prevention controls

Key focus areas for DLP controls

The DLP conceptual model shown on page 11 includes the following DLP control focus areas to help you think through the controls that need to be in place to manage data leaving your organization.

However, as technology is always changing, organizations must continue to keep abreast of technology advancement to identify new ways data can be leaked to unauthorized parties. Furthermore, this list provides a starting point, but additional risks and control requirements may exist within your organization.

Data in motion		
Focus area	Example of control objective	Supporting technologies
Perimeter security	Prevent unencrypted sensitive data from leaving the perimeter	DLP technology, firewalls, proxy servers
Network monitoring	Log and monitor network traffic to identifying and investigate inappropriate sensitive data transfers	DLP technology
Internet access control	Prevent users from accessing unauthorized sites or uploading data through the web through personal webmail, social media, online backup tools, etc	Proxy servers, content filters
Data collection and exchange with third parties	Ensure that data exchange with third parties only occurs through secure means	Secure email, secure FTP, secure APIs, encrypted physical media
Use of instant messaging	Prevent file transfers to external parties through instant messaging and other non-web-based applications	Firewalls, proxy servers, workstation restrictions
Remote access	Ensure that remote access to the company network is secured and control the data that can be saved through remote facilities such as Outlook Web Access	Encrypted remote access, restrictions on use of remote access tools to prevent data leakage to non-corporate assets
Data in use		
Focus area	Example of control objective	Supporting technologies
Privileged user monitoring	Monitor the actions of privileged users with the ability to override DLP controls, perform mass data extracts, etc	Security information and event monitoring, operating database and application log files
Access/usage monitoring	Monitor access and usage of high-risk data to identify potentially inappropriate usage	Security information and event monitoring, operating database and application log files, endpoint DLP logs
Data sanitation	Sanitize/anonymize sensitive data when it is not required for the intended use	Data sanitation routines and programs
Use of test data	Do not use or copy sensitive data into non-production systems. Sanitize data before moving into test systems when possible	Data sanitation routines and programs
Data redaction	Remove sensitive data elements from reports, interfaces and extracts when they are not necessary for the intended use	Data redaction tools
Export/save control	Restrict user abilities to copy sensitive data into unapproved containers (e.g., email, web browsers) including controlling the ability to copy, paste and print sections of documents	Endpoint DLP technology, application controls



Data at rest		
Focus area	Example of control objective	Supporting technologies
Endpoint security	Restrict access to local admin functions such as the ability to install software and modify security settings. Prevent malware, viruses, spyware, etc	Operating system workstation restrictions, security software (e.g., A/V, personal firewall, etc.), endpoint DLP technology
Host encryption	Ensure hard disks are encrypted on all servers, workstations, laptops and mobile devices	Full disk encryption tools
Mobile device protection	Harden mobile device configurations and enable features such as password protection and remote wipe facilities	Built-in security features, third-party mobile device control products
Network/intranet storage	Govern access to network-based repositories containing sensitive data on a least privilege basis	Access control software and permission control in operating systems, databases and file storage systems
Physical media control	Prevent the copying of sensitive data to unapproved media. Ensure that authorized data extraction only takes place on encrypted media.	Endpoint DLP technology, endpoint media encryption tools, operating system workstation restrictions
Disposal and destruction	Ensure all equipment with data storage capabilities are cleansed or destroyed as part of the equipment disposal process (including devices such as digital copiers and fax machines)	Data erasure/data wiping software

Supporting information security processes

DLP controls cannot operate effectively in a vacuum. In order for a DLP program to be effective, the links to other information security processes must be understood so that multiple layers of defense are established and monitored. For example, effective logical access controls may be in place, but if physical controls fail and sensitive hard copy information is removed from your facilities, data loss still occurs. Likewise, if changes to your infrastructure are not carefully controlled, existing DLP controls can become ineffective. The areas listed in the “Supporting

information security processes” section of the DLP conceptual model will help you identify key controls outside of the DLP program that can impact your overall effectiveness in managing data loss risks.

It is essential that DLP controls and supporting information security controls are implemented and that the effectiveness of these controls is monitored over time. Having a structured data loss risk management program and a clear set of controls to mitigate data loss risks can provide a holistic view of data loss potential across your organization. The DLP conceptual model can also aid in building a customized data loss risk dashboard and performing current-state assessments.



Leveraging the power of built-in policies and rules

Market-leading DLP technologies include preconfigured policies and rules to enable you to implement DLP technology rapidly. A key part of a DLP technology implementation is evaluating, selecting and customizing the built-in policies and rules to provide the right coverage for your company and industry. Enabling all of the built-in policies can flood you with data and false positives, resulting in difficulties to spot the high risk incidents. However, not leveraging these policies may keep important events off of your radar. A balance must be achieved to provide adequate coverage without requiring a massive team of people to analyze the data.

Generally, policies that scan for specific data patterns, such as credit card numbers, will provide low false positives and false negatives, particularly when mathematical algorithm checking can be used for validation, in addition to simple pattern matching. Built-in policies covering data such as credit card numbers, Social Security/national ID numbers, international bank account numbers, and medical claim information will generally provide valuable results with reasonable numbers of false positives and are a good place to start in an implementation to test the waters. In contrast, rules that rely only on text search terms, such as scanning for the word "classified" in a document, will initially result in high numbers of false positives and require a significant time to tweak and perfect the policies and rules.

Types of built-in policies:

- ▶ PCI (credit card numbers)
- ▶ PII/Privacy (Social Security numbers, national ID numbers)
- ▶ Medical data/personal health information
- ▶ Acceptable use (traffic related to gambling, pornography)
- ▶ Financial reporting (income statements, balance sheets, annual financial reports)

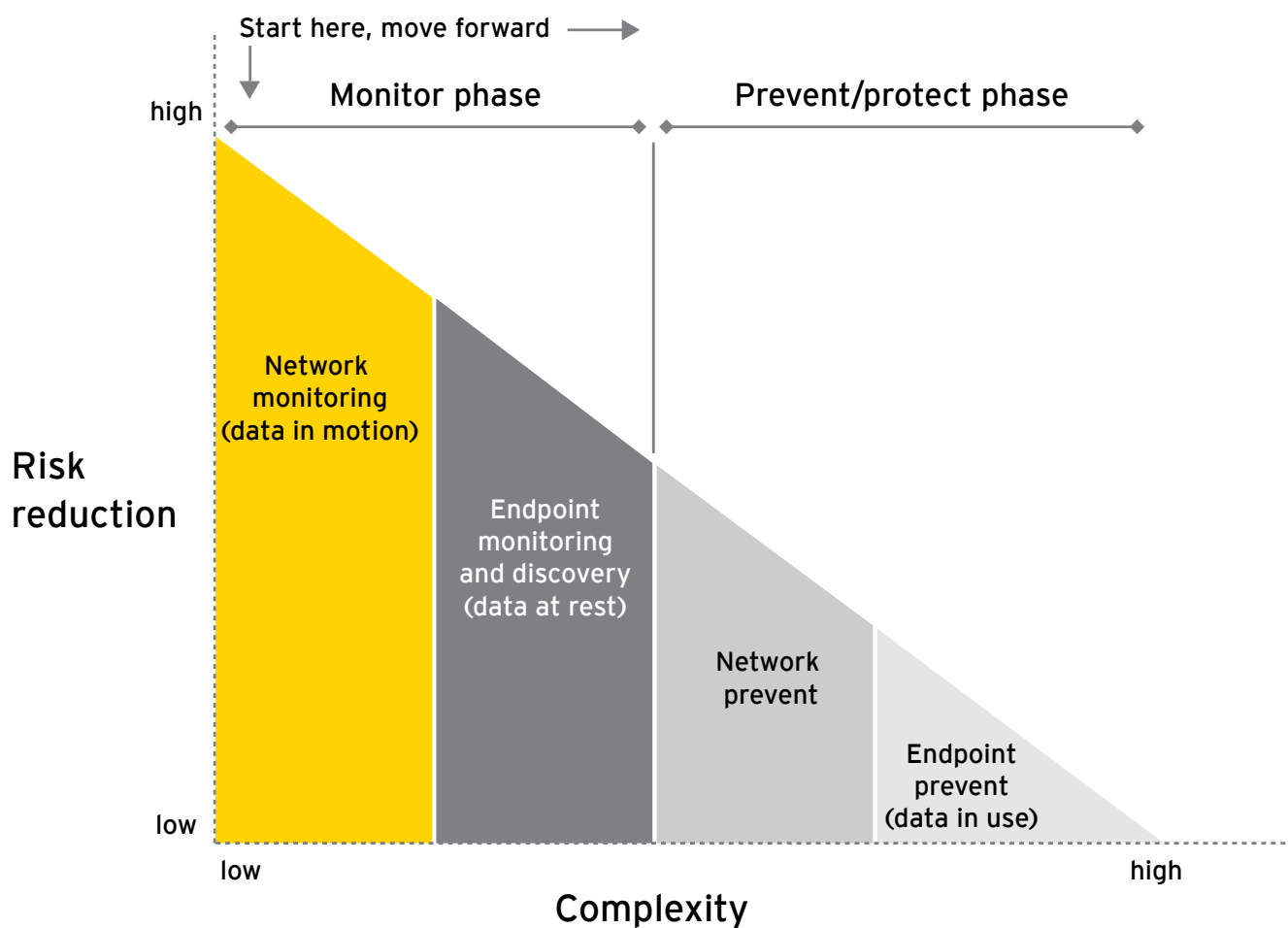
Using technology to support the DLP program

DLP technology is intended to detect and prevent the unauthorized use and transmission of sensitive data and can be set to identify, monitor and protect data in use, data in motion and data at rest.

Generally, the following types of DLP technologies are available:

Type	Purpose
Endpoint-based tools	Control user capabilities on endpoint systems
Network-based monitoring tools	Detect and report on sensitive data in motion
Network-based scanning tools	Scan the network and specific hosts/shares to identify and report (and potentially quarantine) unprotected sensitive data
Perimeter DLP prevent tools	Detect sensitive data flowing through endpoints and stop traffic that violates DLP rules

Endpoint technologies are more difficult and slower to deploy, but they provide more direct control over user actions and coverage for systems that are connected outside of the corporate network. Network-based technologies can be deployed very quickly and provide greater coverage on the internal network, but they are generally unable to scan encrypted traffic and can impact network performance when they are in "prevent" mode i.e., blocking inappropriate traffic. A hybrid approach provides multiple layers of defense with detect and prevent capabilities but is more costly than a single approach to deploy.



DLP technology and trends

In our experience working with a variety of clients, DLP tools and vendors over the last several years, DLP technology has made significant advancements and has matured as major vendors have purchased niche players and integrated DLP products with their other information security solutions. DLP tools now provide significant capabilities that can certainly help a company to establish multiple layers of defense and DLP capabilities.

However, technology alone does not provide a complete solution. Several challenges in tool implementation still exist, particularly for global organizations.

These include:

- ▶ Difficulty scaling to support many languages
- ▶ Limited effectiveness in identifying sensitive intellectual property
- ▶ Limited built-in support for standard data formats outside of the United States
- ▶ Limited deployment capabilities in different countries based on local privacy laws

These are among the reasons that effective people and processes are as important as ever in managing data loss risks. Users must be aware of the risks, DLP roles and responsibilities must be clearly defined and processes must be in place to properly configure DLP tools and to act efficiently on the output.



Ernst & Young insights and lessons learned

By rethinking the way you handle your data distribution and effectively implementing DLP controls, your organization can manage the right risks by:

- ▶ Establishing effective data governance
- ▶ Implementing DLP controls that address your organization's highest risks and most sensitive data
- ▶ Taking a holistic approach by understanding the impact of the broader information security agenda on your DLP program and monitoring the effectiveness of all the relevant controls that support data loss prevention

Lessons learned

From our involvement helping many clients with their DLP programs, we have identified the following practices that can help to make a DLP program successful:

- ▶ **Determine goals and objectives for your DLP program up front** – As with all change initiatives, DLP programs should help achieve strategic business objectives and provide benefits in return for the costs incurred. Clear goals and objectives based on the company strategy and mission should be determined up front as a baseline for your program. This will ensure that the program is focused on protecting the data that is most important to the business.
- ▶ **Address all aspects of people, process and technology** – As we have illustrated, a defense-in-depth approach must be taken, with clear roles and responsibilities for individuals, fit-for-purpose tools to identify and prevent data loss and effective processes to research and respond to incidents.
- ▶ **Establish ample executive support, understanding and participation** – Company-wide support and involvement from various business and operating units will create more user acceptance of the transition toward a more secure environment and will help to ensure that business input is provided at key stages.
- ▶ **Defining sensitive data is a fundamental requirement** – Implementing DLP technology and controls universally across an organization has an adverse and costly impact on the business. By defining sensitive data up front and aligning the program to protect their most sensitive data, organizations can ensure that resources are spent managing the highest risks.
- ▶ **Focus on a defense-in-depth approach, not just complying with legislation** – A DLP implementation should not be based solely on solving your compliance issues but should focus on the entire risk spectrum that affects your data.

- ▶ **Leverage the opportunity to reinforce awareness in the end user community** – The end user is mostly confronted by the limitations of DLP controls. Without an awareness and acceptance of the reasons behind additional security measures, employees will be more likely to seek “work-arounds” to bypass controls. Increased awareness will also help users to execute their responsibilities for classifying data at the point of creation.
- ▶ **Formally define and monitor the effectiveness of your DLP controls** – Once implemented, the DLP controls and their effectiveness in protecting your data assets should be monitored in order to feed back into an improvement cycle. A formal process can ensure that continuous improvement can be applied in line with the current and future threats to your data.

Common DLP program mistakes

Data protection efforts tend to focus on technology	<ul style="list-style-type: none"> ▶ Too much focus on technical implementation details ▶ Lack of integration with existing processes ▶ Lack of holistic data protection strategy
Poor information classification definition and user awareness	<ul style="list-style-type: none"> ▶ What is considered sensitive data? ▶ Are data elements only sensitive when combined? ▶ Where can I store sensitive data? File shares, SharePoint? ▶ Who has (or can have) access to my data?
No authoritative source exists for share ownership	<ul style="list-style-type: none"> ▶ Lack of centralized authoritative list of share owners ▶ Poor awareness for established share ownership
Access approval performed without proper context	<ul style="list-style-type: none"> ▶ Inappropriate personnel executing (e.g., IT) ▶ Limited understand and visibility of data, users and groups ▶ How and where inheritance is used ▶ How and where nested groups are used
Limited or non-existent entitlement review process	<ul style="list-style-type: none"> ▶ Access is often not revoked until a user separates from organization ▶ Over time, users have access to data that they no longer require
Ownership for non-user accounts	<ul style="list-style-type: none"> ▶ Service accounts ▶ Non-primary user accounts without documented owner
Lack of engagement from relevant stakeholders	<ul style="list-style-type: none"> ▶ Lack of coordination with business stakeholders ▶ Project viewed as an IT project only

Top practical tips to help optimize your data loss prevention program

- 1. Identify and classify your data.** A well-developed, granular data classification scheme will enable your company to design and implement the proper controls for different types of data. A data inventory, linking the data classification scheme to specific data held within the IT infrastructure and with external parties, will help appropriately scope your DLP program.
- 2. Be concerned about view-only access.** The use of data warehouses and demands for organization-wide reporting capabilities means it can be much easier for users to put together previously disparate data elements, resulting in wide access to very sensitive data. As many organizations have built security controls with the idea that view-only or read-only access is low-risk, often individuals can view (and extract) far more data than they truly need to do their jobs.
- 3. Implement a data management lifecycle.** A primary challenge organizations have with data protection is the improper definition, classification and storage of sensitive data at the point of creation. If data is not properly defined, classified and stored appropriately at the starting point, it can propagate throughout the organization, which can make protecting it later an even greater challenge. Additionally, compliance with data retention policies is a challenge, which increases the amount of data that needs to be managed and protected.
- 4. Do not allow unauthorized devices on your network.** Allowing non-corporate assets to access the internal network can lead to several risks. These include unauthorized parties with physical access to the company's premises accessing internal network resources, and internal users connecting personal devices to the corporate network. Personal devices will most likely lack corporate device protection measures and endpoint security controls.
- 5. Do not permit the copying of sensitive data to removable media.** Endpoints should be configured to disable writing to all removable storage devices, or alternatively, content-aware endpoint DLP technology should be in place to prevent sensitive data from being copied from the source. Mobile devices like laptops and mobile smartphones or PDAs should have full disk encryption, and your company should have the ability to erase them remotely if they are lost or stolen.
- 6. Improve authorization and access control measures.** Review and tighten data access controls and usage triggers. When identity and access management systems are initially deployed, user roles and user access are typically broadly defined. Now is the time to review those roles and access to make sure employees have access only to the data required to fulfill their responsibilities successfully and nothing more. Access restrictions should be augmented by monitoring to identify unusual or suspicious activity by those with access to sensitive data.
- 7. Understand data usage and flows and your data loss vectors.** Organizations need to understand how data is being used and how it can leave the organization. Tools should be implemented to monitor data traffic flows within your infrastructure. File access monitoring can increase your current knowledge of data usage within the organization, and network-based DLP tools should also be considered to gain visibility over data flowing across internal networks and to the internet.
- 8. Take a risk-based approach.** Not all data is created equally. Your DLP program should be designed to protect your most sensitive data (not *all* data). When considering DLP technology implementation across a complex environment, consider structuring the implementation in phases, with the initial roll-out addressing high-risk/high-value networks, hosts or physical locations. A phased approach will help to prove the technology, increase buy-in as benefits are realized, and gradually build internal skills, rule sets and data analysis capabilities.
- 9. Update your policies and create awareness.** You should improve your security awareness program to incorporate data loss awareness in line with company policies to ensure that everyone is aware of the potential data loss risks. Clear guidance should educate employees on what is expected from them when handling data. In addition, a clear and well-understood data protection policy will encourage proper behavior by employees and data owners with respect to data handling, storage and transfer.
- 10. Audit your own compliance.** Often awareness programs are not tested properly to validate an employee's knowledge and compliance with company policies. Organizations should test employee awareness through real-life examples of attacks to verify that employees know what to do under specific circumstances. Consider conducting social engineering and phishing tests to determine awareness levels.

Don't be a victim

The ever-evolving risk landscape is becoming more challenging to manage. With data loss, prevention is always better than recovering after a breach. Today's common threats are accelerating by technological evolution. Data loss through social media, consumerization, cybercrime and internal threats represent increasing business risks. An organization that knows which data

is most vital to its business operations, understands where that data resides and how that data is sent beyond its walls will achieve competitive advantages in the marketplace. It is important to use that knowledge to ensure that key business data is effectively protected and that the organization rapidly and efficiently responds to incidents that occur.

About Ernst & Young

Ernst & Young is a global leader in assurance, tax, transaction and advisory services. Worldwide, our 141,000 people are united by our shared values and an unwavering commitment to quality. We make a difference by helping our people, our clients and our wider communities achieve their potential.

Ernst & Young refers to the global organization of member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients. For more information about our organization, please visit www.ey.com.

About Ernst & Young's Advisory Services

The relationship between risk and performance improvement is an increasingly complex and central business challenge, with business performance directly connected to the recognition and effective management of risk. Whether your focus is on business transformation or sustaining achievement, having the right advisors on your side can make all the difference. Our 20,000 advisory professionals form one of the broadest global advisory networks of any professional organization, delivering seasoned multidisciplinary teams that work with our clients to deliver a powerful and superior client experience. We use proven, integrated methodologies to help you achieve your strategic priorities and make improvements that are sustainable for the longer term. We understand that to achieve your potential as an organization you require services that respond to your specific issues, so we bring our broad sector experience and deep subject matter knowledge to bear in a proactive and objective way. Above all, we are committed to measuring the gains and identifying where the strategy is delivering the value your business needs. It's how Ernst & Young makes a difference.

© 2011 EYGM Limited.
All Rights Reserved.

EYG no. AU0969



In line with Ernst & Young's commitment to minimize its impact on the environment, this document has been printed on paper with a high recycled content.

This publication contains information in summary form and is therefore intended for general guidance only. It is not intended to be a substitute for detailed research or the exercise of professional judgment. Neither EYGM Limited nor any other member of the global Ernst & Young organization can accept any responsibility for loss occasioned to any person acting or refraining from action as a result of any material in this publication. On any specific matter, reference should be made to the appropriate advisor.

About Ernst & Young

At Ernst & Young, our services focus on our individual clients' specific business needs and issues because we recognize that each is unique to that business.

IT is a key to allowing modern organizations to compete. It offers the opportunity to become closer to customers and more focused and faster in responses, and can redefine both the effectiveness and efficiency of operations. But as opportunity grows, so does risk. Effective ITRM helps you to improve the competitive advantage of your IT operations, by making these operations more cost efficient and managing down the risks related to running your systems. Our 6,000 IT risk professionals draw on extensive personal experience to give you fresh perspectives and open, objective advice – wherever you are in the world. We work with you to develop an integrated, holistic approach to your IT risk or to deal with a specific risk and information security issue. We understand that to achieve your potential you need tailored services as much as consistent methodologies. We work to give you the benefit of our broad sector experience, our deep subject-matter knowledge and the latest insights from our work worldwide. It's how Ernst & Young makes a difference.

For more information on how we can make a difference in your organization, contact your local Ernst & Young professional or a member of our team listed below.

Contacts

Global

Norman Lonergan +44 20 7980 0596 norman.lonergan@uk.ey.com
(Advisory Services Leader, London)

Paul van Kessel +31 88 40 71271 paul.van.kessel@nl.ey.com
(IT Risk and Assurance Services Leader, Amsterdam)

Advisory Services

Robert Patton +1 404 817 5579 robert.patton@ey.com
(Americas Leader, Atlanta)

Andrew Embury +44 20 7951 1802 aembury@uk.ey.com
(Europe, Middle East, India and Africa Leader, London)

Doug Simpson +61 2 9248 4923 doug.simpson@au.ey.com
(Asia-Pacific Leader, Sydney)

Naoki Matsumura +81 3 3503 1100 matsumura-nk@shinnihon.or.jp
(Japan Leader, Tokyo)

IT Risk and Assurance Services

Bernie Wedge +1 404 817 5120 bernard.wedge@ey.com
(Americas Leader, Atlanta)

Manuel Giralt Herrero +34 91 572 7479 manuel.giraltherrero@es.ey.com
(Europe, Middle East, India and Africa Leader, Madrid)

Troy Kelly +85 2 2629 3238 troy.kelly@hk.ey.com
(Asia Pacific Leader, Hong Kong)

Giovanni Stagno +81 3 3506 2411 stagno-gvnn@shinnihon.or.jp
(Japan Leader, Chiyoda-ku)