Data classification for cloud readiness

Microsoft Trustworthy Computing

Legal disclaimer

This document is for informational purposes only. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED, OR STATUTORY, AS TO THE INFORMATION IN THIS DOCUMENT.

This document is provided "as-is." Information and views expressed in this document, including URL and other Internet website references, may change without notice. You bear the risk of using it.

Microsoft and Windows Azure are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

The names of actual companies and products mentioned herein may be the trademarks of their respective owners.

Acknowledgments

Authors

Frank Simorjay

Contributors and Reviewers

Kellie Ann Chainier
Carlene Heath
Jim Pinter
Tim Rains
Kurt Dillard
Greg Lenti
Shont Miller (LCA)
Steve Wacker
Chris Hale (LCA)
Michael Mattmiller
Sian Suthers

Introduction

Data classification provides one of the most basic ways for organizations to determine and assign relative values to the data they possess. The process of data classification allows organizations to categorize their stored data by sensitivity and business impact in order to determine the risks associated with the data. After the process is completed, organizations can manage their data in ways that reflect its value to them instead of treating all data the same way. Data classification is a conscious, thoughtful approach that enables organizations to realize optimizations that might not be possible when all data is assigned the same value.

Data classification has been used for decades to help large organizations such as Microsoft, governments, and military entities manage the integrity of their data. This paper provides readers with an introduction to the fundamentals of data classification and highlights its value, specifically in the context of cloud computing. Organizations that are assessing cloud computing for future use or organizations that are currently using cloud services and seeking ways to optimize data management will benefit most from this paper.

Although risk assessments are sometimes used by organizations as a starting point for data classification efforts, this paper does not discuss a process for a formal risk assessment. Organizations are strongly encouraged to consider identified risks that are specific to their business when developing a data classification process.

Who should read this paper

This paper is primarily intended for consultants, security specialists, systems architects, and IT professionals who are responsible for planning application or infrastructure development and deployment for their organizations. These roles include the following common job descriptions:

- Senior executives, business analysts, and business decision makers (BDMs) who have critical business objectives and requirements that need IT support
- Architects and planners who are responsible for driving the architecture efforts for their organizations
- Consultants and partner organizations who need knowledge transfer tools for their customers and partners

Data classification fundamentals

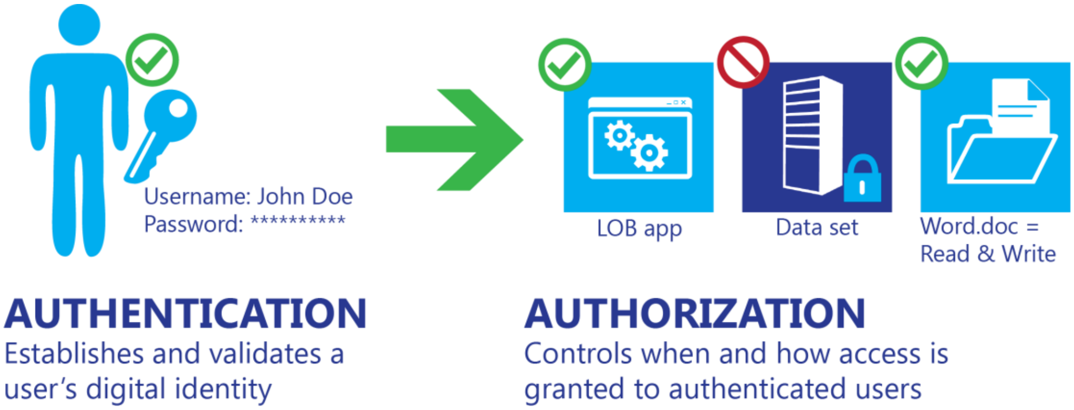
Successful data classification in an organization requires broad awareness of the organization's needs and a thorough understanding of where the organization's data assets reside.

Data exists in one of three basic states: at rest, in process, and in transit. All three states require unique technical solutions for data classification, but the applied principles of data classification should be the same for each. Data that is classified as confidential needs to stay confidential when at rest, in process, and in transit.

Data can also be either structured or unstructured. Typical classification processes for the structured data found in databases and spreadsheets are less complex and time-consuming to manage than those for unstructured data such as documents, source code, and email. Generally, organizations will have more unstructured data than structured data. Regardless of whether data is structured or unstructured, it is important for organizations to manage data sensitivity. When properly implemented, data classification helps ensure that sensitive or confidential data assets are managed with greater oversight than data assets that are considered public or free to distribute.

Controlling access to data

Authentication and authorization are often confused with each other and their roles misunderstood. In reality they are quite different, as shown in the following figure.



Authentication

Authentication typically consists of at least two parts: a username or user ID to identify a user and a token, such as a password, to confirm that the username credential is valid. The process does not provide the authenticated user with access to any items or services; it verifies that the user is who they say they are.

Authorization

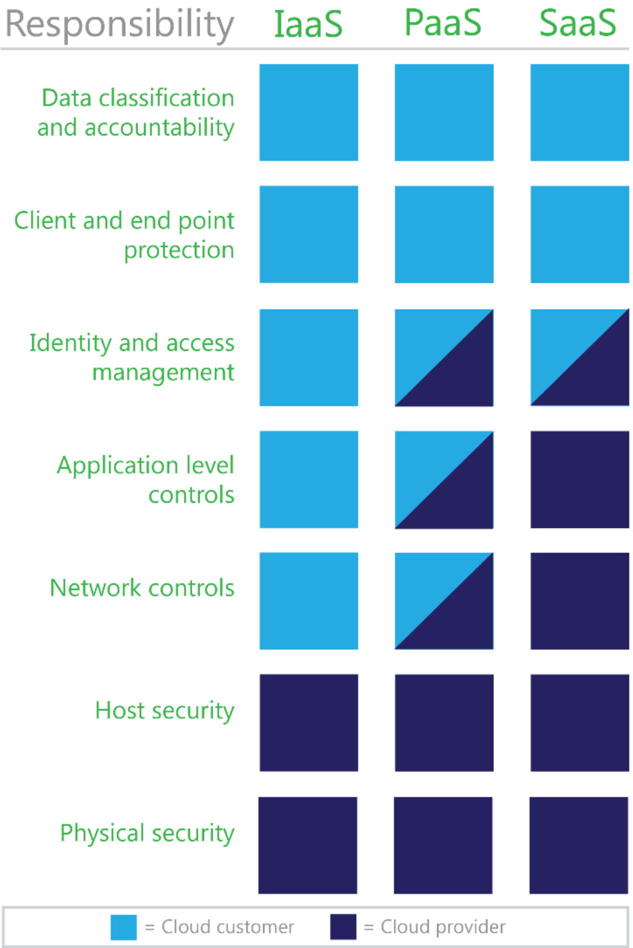
Authorization is the process of providing an authenticated user the ability to access an application, data set, data file, or some other object. Assigning authenticated users the rights to use, modify, or delete items that they can access requires attention to data classification.

Successful authorization requires implementation of a mechanism to validate individual users' needs to access files and information based on a combination of role, security policy, and risk policy considerations. For example, data from specific line-of-business (LOB) applications might not need to be accessed by all employees, and only a small subset of employees will likely need access to human resources (HR) files. But for organizations to control who can access data, as well as when and how, an effective system for authenticating users must be in place.

Roles and responsibilities in cloud computing

Authorization requires an essential understanding of the roles and responsibilities of organizations, cloud providers, and customers. Cloud providers must have operational practices in place to prevent unauthorized access to customer data; it's also important to note that any compliance requirements a customer organization has must also be supported by the provider. Although cloud providers can help manage risks, customers need to ensure that data classification management and enforcement is properly implemented to provide the appropriate level of data management services.

Data classification responsibilities will vary based on which cloud service model is in place, as shown in the following figure. The three primary cloud service models are infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS). Implementation of data classification mechanisms will also vary based on the reliance on and expectations of the cloud provider.



Responsibility	IaaS	PaaS	SaaS
Data classification and accountability	🟦	🟦	🟦
Client and end point protection	🟦	🟦	🟦
Identity and access management	🟦	◩	◩
Application level controls	🟦	◩	⬛
Network controls	🟦	◩	⬛
Host security	⬛	⬛	⬛
Physical security	⬛	⬛	⬛

🟦 = Cloud customer ⬛ = Cloud provider

Although customers are responsible for classifying their data, cloud providers should make written commitments to customers about how they will secure and maintain the privacy of the customer data stored within their cloud. These commitments should include information about privacy and security practices, data use limitations, and regulatory compliance. In addition, cloud providers should make certifications and audit reports that demonstrate compliance with standards such as the International Organization for Standardization (ISO) and controls such as the American Institute of CPAs Service Organization Controls (SOC1 and SOC2) available so customers can verify the effectiveness of their cloud provider's practices. Having this information will help customers understand whether the cloud provider supports the data protection requirements mandated by their data classification. Customers should not migrate data to a cloud provider that cannot address their data protection needs.

- **IaaS providers**. From a data classification perspective, IaaS provider requirements are limited to ensuring that the virtual environment can accommodate data classification capabilities and customer compliance requirements. IaaS providers have a smaller role in data classification because they only need to ensure that customer data addresses compliance requirements.

However, providers must still ensure that their virtual environments address data classification requirements in addition to securing their data centers.

- **PaaS providers**. Responsibilities may be mixed, because the platform could be used in a layered approach to provide security for a classification tool. PaaS providers may be responsible for authentication and possibly some authorization rules, and must provide security and data classification capabilities to their application layer. Much like IaaS providers, PaaS providers need to ensure that their platform complies with any relevant data classification requirements.
- **SaaS providers** will frequently be considered as part of an authorization chain, and will need to ensure that the data stored in the SaaS application can be controlled by classification type. SaaS applications can be used for LOB applications, and by their very nature need to provide the means to authenticate and authorize data that is used and stored.

Compliance considerations

In addition, organizations that are considering cloud solutions and need to comply with regulatory requirements can benefit by working with cloud providers that comply with regulations such as FedRAMP, U.S. HIPAA, EU Data Protection Directive, and others listed in Appendix 1. However, to achieve compliance, such organizations need to remain aware of their classification obligations and be able to manage the classification of data that they store in the cloud. For example, the Cloud Security Alliance identifies the following data classification control requirement in its Cloud Control Matrix:

The Cloud Security Alliance's Cloud Control Matrix question on Data Governance – Classification (CCM 1.0 - DG-02 | CCM 3.0 - DSI-03)

Data Governance – Classification (control) from the CCM states that:

Data, and objects containing data, need to be assigned a classification based on data type, jurisdiction of origin, jurisdiction domiciled, context, legal constraints, contractual constraints, value, sensitivity, criticality to the organization and third party obligation for retention and prevention of unauthorized disclosure or misuse.

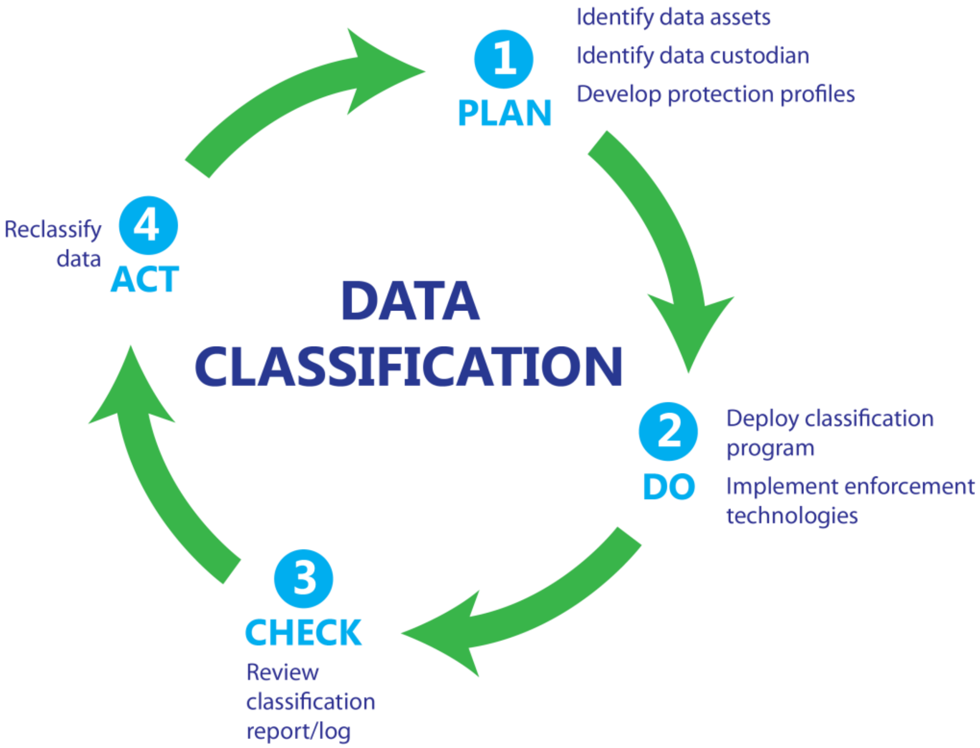
https://cloudsecurityalliance.org/research/ccm/

Classification process

Many organizations that understand the need for data classification and want to implement it face a basic challenge: where to begin?

One effective and simple way to implement data classification is to use the PLAN, DO, CHECK, ACT model from MOF. The following figure charts the tasks that are required to successfully implement data classification in this model.

1. **PLAN**. Identify data assets, a data custodian to deploy the classification program, and develop protection profiles.
2. **DO**. After data classification policies are agreed upon, deploy the program and implement enforcement technologies as needed for confidential data.
3. **CHECK**. Check and validate reports to ensure that the tools and methods being used are effectively addressing the classification policies.
4. **ACT**. Review the status of data access and review files and data that require revision using a reclassification and revision methodology to adopt changes and to address new risks.


1
PLAN
Identify data assets
Identify data custodian
Develop protection profiles
2
DO
Deploy classification program
Implement enforcement technologies
3
CHECK
Review classification report/log
4
ACT
Reclassify data
DATA CLASSIFICATION

Select a terminology model that addresses your needs

Several types of processes exist for classifying data, including manual processes, location-based processes that classify data based on a user's or system's location, application-based processes such as database-specific classification, and automated processes used by various technologies, some of which are described in the ”Protecting confidential data” section later in this paper.

This paper introduces two generalized terminology models that are based on well-used and industry-respected models. These terminology models, both of which provide three levels of classification sensitivity, are shown in the following table.

Note: When classifying a file or resource that combines data that would typically be classified at differing levels, the highest level of classification present should establish the overall classification. For example, a file containing sensitive and restricted data should be classified as restricted.

Sensitivity	Terminology model 1	Terminology model 2
High	Confidential	Restricted
Medium	For internal use only	Sensitive
Low	Public	Unrestricted

- **Confidential (restricted)**. Information that is classified as confidential or restricted includes data that can be catastrophic to one or more individuals and/or organizations if compromised or lost. Such information is frequently provided on a "need to know" basis and might include:
    - Personal data, including personally identifiable information such as Social Security or national identification numbers, passport numbers, credit card numbers, driver's license numbers, medical records, and health insurance policy ID numbers.
    - Financial records, including financial account numbers such as checking or investment account numbers.
    - Business material, such as documents or data that is unique or specific intellectual property.
    - Legal data, including potential attorney-privileged material.
    - Authentication data, including private cryptography keys, username password pairs, or other identification sequences such as private biometric key files.

    Data that is classified as confidential frequently has regulatory and compliance requirements for data handling. Specifics of some of these requirements are listed in Appendix 1.
- **For internal use only (sensitive)**. Information that is classified as being of medium sensitivity includes files and data that would not have a severe impact on an individual and/or organization if lost or destroyed. Such information might include:

    - Email, most of which can be deleted or distributed without causing a crisis (excluding mailboxes or email from individuals who are identified in the confidential classification).
    - Documents and files that do not include confidential data.

  Generally, this classification includes anything that is not confidential. This classification can include most business data, because most files that are managed or used day-to-day can be classified as sensitive. With the exception of data that is made public or is confidential, all data within a business organization can be classified as sensitive by default.

- **Public (unrestricted)**. Information that is classified as public includes data and files that are not critical to business needs or operations. This classification can also include data that has deliberately been released to the public for their use, such as marketing material or press announcements. In addition, this classification can include data such as spam email messages stored by an email service.

Define data ownership

It's important to establish a clear custodial chain of ownership for all data assets. The following table identifies different data ownership roles in data classification efforts and their respective rights.

Note: This table does not provide an exhaustive list of roles and rights, but merely a representative sample.

Role	Create	Modify/delete	Delegate	Read	Archive/restore
Owner	X	X	X	X	X
Custodian			X		
Administrator					X
User*		X		X	

*Users may be granted additional rights such as edit and delete by a custodian.

- The **data asset owner** is the original creator of the data, who can delegate ownership and assign a custodian. When a file is created, the owner should be able to assign a classification, which means that they have a responsibility to understand what needs to be classified as confidential based on their organization's policies. All of a data asset owner's data can be auto-classified as for internal use only (sensitive) unless they are responsible for owning or creating confidential (restricted) data types. Frequently, the owner's role will change after the data is classified. For example, the owner might create a database of classified information and relinquish their rights to the data custodian.

    - **Note regarding personal data**: Data asset owners often use a mixture of services, devices, and media, some of which are personal and some of which belong to the organization. A clear organizational policy can help ensure that usage of devices such as laptops and smart devices is in accordance with data classification guidelines.
- The **data asset custodian** is assigned by the asset owner (or their delegate) to manage the asset according to agreements with the asset owner or in accordance with applicable policy requirements. Ideally, the custodian role can be implemented in an automated system. An asset custodian ensures that necessary access controls are provided and is responsible for managing and protecting assets delegated to their care. The responsibilities of the asset custodian could include:
    - Protecting the asset in accordance with the asset owner’s direction or in agreement with the asset owner
    - Ensuring that classification policies are complied with
    - Informing asset owners of any changes to agreed-upon controls and/or protection procedures prior to those changes taking effect
    - Reporting to the asset owner about changes to or removal of the asset custodian’s responsibilities
- An **administrator** represents a user who is responsible for ensuring that integrity is maintained, but they are not a data asset owner, custodian, or user. In fact, many administrator roles provide data container management services without having access to the data. The administrator role includes backup and restoration of the data, maintaining records of the assets, and choosing, acquiring, and operating the devices and storage that house the assets.
- The **asset user** includes anyone who is granted access to data or a file. Access assignment is often delegated by the owner to the asset custodian.

Implementation

Management considerations apply to all classification methodologies. These considerations need to include details about who, what, where, when, and why a data asset would be used, accessed, changed, or deleted. All asset management must be done with an understanding of how an organization views its risks, but a simple methodology can be applied as defined in the data classification process. Additional considerations for data classification include the introduction of new applications and tools, and managing change after a classification method is implemented.

Reclassification

Reclassifying or changing the classification state of a data asset needs to be done when a user or system determines that the data asset’s importance or risk profile has changed. This effort is important for ensuring that the classification status continues to be current and valid. Most content that is not classified manually can be classified automatically or based on usage by a data custodian or data owner.

- **Manual data reclassification**. Ideally, this effort would ensure that the details of a change are captured and audited. The most likely reason for manual reclassification would be for reasons of sensitivity, or for records kept in paper format, or a requirement to review data that was originally misclassified. Because this paper considers data classification and moving data to the cloud, manual reclassification efforts would require attention on a case-by-case basis and a risk management review would be ideal to address classification requirements. Generally, such an effort would consider the organization's policy about what needs to be classified, the default classification state (all data and files being sensitive but not confidential), and take exceptions for high-risk data.
- **Automatic data reclassification** uses the same general rule as manual classification. The exception is that automated solutions can ensure that rules are followed and applied as needed. Data classification can be done as part of a data classification enforcement policy, which can be enforced when data is stored, in use, and in transit using authorization technology.
    - **Application-based**. Using certain applications by default sets a classification level. For example, data from customer relationship management (CRM) software, HR, and health record management tools is confidential by default.
    - **Location-based**. Data location can help identify data sensitivity. For example, data that is stored by an HR or financial department is more likely to be confidential in nature.

Data retention, recovery, and disposal

Data recovery and disposal, like data reclassification, is an essential aspect of managing data assets. The principles for data recovery and disposal would be defined by a data retention policy and enforced in the same manner as data reclassification; such an effort would be performed by the custodian and administrator roles as a collaborative task.

Failure to have a data retention policy could mean data loss or failure to comply with regulatory and legal discovery requirements. Most organizations that do not have a clearly defined data retention policy tend to use a default "keep everything" retention policy. However, such a retention policy has additional risks in cloud services scenarios. For example, a data retention policy for cloud service providers can be considered as "for the duration of the subscription" (as long as the service is paid for, the data is retained). Such a pay-for-retention agreement may not address corporate or regulatory retention policies. Defining a policy for confidential data can ensure that data is stored and removed based on best practices. In addition, an archival policy can be created to formalize an understanding about what data should be disposed of and when.

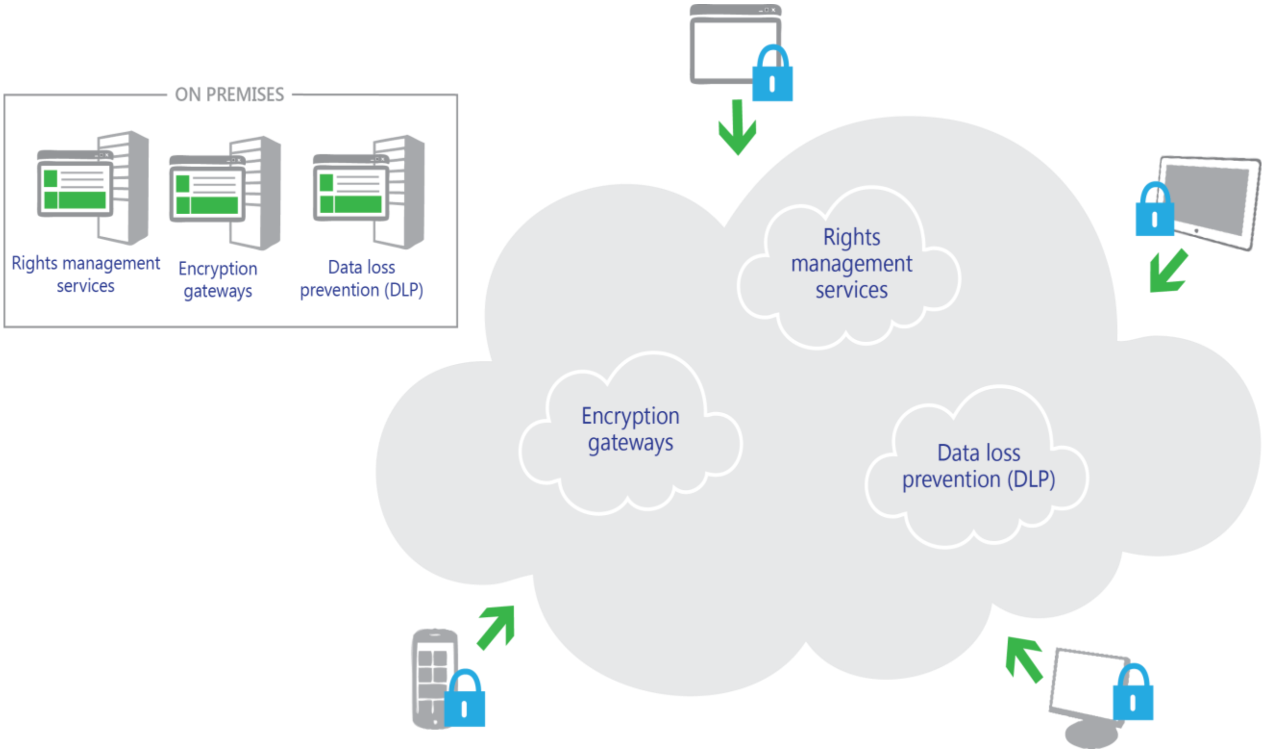
Data retention policy should address the required regulatory and compliance requirements, as well as corporate legal retention requirements. Classified data might provoke questions about retention duration and exceptions for data that has been stored with a provider; such questions are more likely for data that has not been classified correctly.

Protecting confidential data

After data is classified, finding and implementing ways to protect confidential data becomes an integral part of any data protection deployment strategy. Protecting confidential data requires additional attention to how data is stored and transmitted in conventional architectures as well as in the cloud.

This section provides basic information about some technologies that can automate enforcement efforts to help protect data that has been classified as confidential.

As the following figure shows, these technologies can be deployed as on-premises or cloud-based solutions—or in a hybrid fashion, with some of them deployed on-premises and some in the cloud. (Some technologies, such as encryption and rights management, also extend to user devices.)


ON PREMISES
Rights management services
Encryption gateways
Data loss prevention (DLP)
Rights management services
Encryption gateways
Data loss prevention (DLP)

Rights management software

One solution for preventing data loss is rights management software. Unlike approaches that attempt to interrupt the flow of information at exit points in an organization, rights management software works at deep levels within data storage technologies. Documents are encrypted, and control over who can decrypt them uses access controls that are defined in an authentication control solution such as a directory service.

Some of the benefits of rights management software include:

- **Safeguarded sensitive information**. Users can protect their data directly using rights management-enabled applications. No additional steps are required—authoring documents, sending email, and publishing data offer a consistent data protection experience.
- **Protection travels with the data**. Customers remain in control of who has access to their data, whether in the cloud, existing IT infrastructure, or at the user’s desktop. Organizations can choose to encrypt their data and restrict access according to their business requirements.
- **Default information protection policies**. Administrators and users can use standard policies for many common business scenarios, such as "Company Confidential–Read Only" and "Do Not Forward." A rich set of usage rights are supported such as read, copy, print, save, edit, and forward to allow flexibility in defining custom usage rights.

More information about using rights management solutions in Microsoft environments is available in the following links:

- “The Desktop Files – Data Loss Prevention with Enterprise Rights Management” in TechNet magazine
- The Information Rights Management page on TechNet
- This Windows Azure Active Directory Rights blog post
- This Microsoft Rights Management blog post

Encryption gateways

Encryption gateways operate in their own layers to provide encryption services by rerouting all access to cloud-based data. This approach should not be confused with that of a virtual private network (VPN); encryption gateways are designed to provide a transparent layer to cloud-based solutions.

Encryption gateways can provide a means to manage and secure data that has been classified as confidential by encrypting the data in transit as well as data at rest.

Encryption gateways are placed into the data flow between user devices and application data centers to provide encryption/decryption services. These solutions, like VPNs, are predominantly on-premises solutions. They are designed to provide a third party with control over encryption keys, which helps reduce the risk of placing both the data and key management with one provider. Such solutions are designed, much like encryption, to work seamlessly and transparently between users and the service.

Data loss prevention

Data loss (sometimes referred to as data leakage) is an important consideration, and the prevention of external data loss via malicious and accidental insiders is paramount for many organizations.

Data loss prevention (DLP) technologies can help ensure that solutions such as email services do not transmit data that has been classified as confidential. Organizations can take advantage of DLP features in existing products to help prevent data loss. Such features use policies that can be easily created from scratch or by using a template supplied by the software provider.

DLP technologies can perform deep content analysis through keyword matches, dictionary matches, regular expression evaluation, and other content examination to detect content that violates organizational DLP policies. For example, DLP can help prevent the loss of the following types of data:

- Social Security and national identification numbers
- Banking information
- Credit card numbers
- IP addresses

Some DLP technologies also provide the ability to override the DLP configuration (for example, if an organization needs to transmit Social Security number information to a payroll processor). In addition, it's possible to configure DLP so that users are notified before they even attempt to send sensitive information that should not be transmitted.

A technical overview of the DLP features in Microsoft Exchange Server 2013 and Exchange Online is available on the Data Loss Prevention page on Microsoft TechNet.

Conclusion

Generally, the topic of data classification does not generate as much interest as other, more exciting technology topics. However, data classification can yield significant benefits, such as compliance efficiencies, improved ways to manage the organization's resources, and facilitation of migration to the cloud. Although data classification efforts can be complex undertakings and require risk assessment for successful implementation, quicker and simpler efforts can also yield benefits. Any data classification effort should endeavor to understand the needs of the organization and be aware how data is stored, processing capabilities, and how data is transmitted throughout the organization.

It’s important for management to support data classification efforts, and for IT to be involved as well. The concept of classification may seem primarily to be an auditing function, but many technology solutions are available that can reduce the amount of effort that is required to successfully implement a data classification model.

It’s also worth noting that data classification rules that pertain to data retention must be addressed when moving to the cloud, and that cloud solutions can help mitigate risk. Some data protection technologies such as encryption, rights management, and data loss prevention solutions have moved to the cloud and can help mitigate cloud risks.

Although this paper did not specifically discuss hybrid environments, a mixture of on-premises and cloud-based data classification technologies can help effectively reduce risk for organizations of any size by providing more control about where data is stored, which gives customers the option to keep highly sensitive data on-premises and under a different set of controls than data stored in the cloud. Indeed, hybrid environments are likely to be the way of the future, and the key to effective data management may well depend on effective data classification.

Appendix 1: Data classification regulations, compliance requirements, and standards

The following table identifies sample control objective definitions. This list is not complete or authoritative, and should only be used as a discussion point to consider when moving services to a cloud solution.

US regulation, requirement, or standard	Control details
NIST SP800-53 R3 National Institute of Standards and Technology	RA-2 Security Categorization AC-4 Information Flow Enforcement
PCI DSS v2.0 Payment Card Industry Data Security Standard	9.7.1 Classify media so the sensitivity of the data can be determined. 9.10 Destroy media when it is no longer needed for business or legal reasons. 12.3 Develop usage policies for critical technologies (for example, remote-access technologies, wireless technologies, removable electronic media, laptops, tablets, personal data/digital assistants (PDAs), e-mail usage and Internet usage) and define proper use of these technologies.
NERC CIP North American Electric Reliability Corporation Critical Infrastructure Protection	CIP-003-3 - R4 - R5 - Responsible Entities have minimum security management controls in place to protect Critical Cyber Assets.
FedRAMP Federal Risk and Authorization Management Program	RA-2 Security Categorization AC-4 Information Flow Enforcement
AICPA SOC2 American Institute of CPAs Service Organization Controls	(S3.8.0) Procedures exist to classify data in accordance with classification policies and periodically monitor and update such classifications as necessary. (C3.14.0) Procedures exist to provide that system data are classified in accordance with the defined confidentiality and related security policies.

International regulation, requirement, or standard	Control details
ENISA IAF European Union Agency for Network and Information Security – Information Assurance Framework	6.05.(c) Asset management - classification, segmentation Employees obliged to adhere to regulations on information security, data protection, adequate handling of customer data
ISO/IEC 27001-2005 International Organization for Standardization / International Electrotechnical Commission	A.7.2.1 Classification guidelines

Appendix 2: Glossary of terms

Archive and recovery. As discussed in this paper, the long-term storage of data and its retrieval when it needs to be returned to service. Archival and recovery methods must conform to the retention model that is used.

Authentication. A process that confirms that a user (identified by a username or user ID) is valid through use of a token or password. This process verifies that the user is who they say they are.

Authorization. A process that provides an authenticated user with the ability to access an application, data set, data file, or some other object.

Cloud. The NIST Definition of Cloud Computing (PDF) states:

> "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models."

Cloud Security Alliance. The Cloud Security Alliance (CSA) is a not-for-profit organization with a mission to promote the use of best practices for providing security assurance within cloud computing, and to provide education on the uses of cloud computing to help secure all other forms of computing. The Cloud Security Alliance is led by a broad coalition of industry practitioners, corporations, associations, and other key stakeholders.
www.cloudsecurityalliance.org

Cloud Control Matrix. The Cloud Security Alliance Cloud Controls Matrix (CCM) is specifically designed to provide fundamental security principles to guide cloud vendors and to assist prospective cloud customers in assessing the overall security risk of a cloud provider. As a framework, the CSA CCM provides organizations with essential structure, detail, and clarity with regard to information security as it relates to the cloud industry.
https://cloudsecurityalliance.org/research/ccm/

Data disposal. As discussed in this paper, the policies, timeframes, and methods for secure disposal of data. Disposal policy may require the destruction of data using strong deletion methods or shredding of disks. Data disposal policies require the same care as data retention policies. More information is available at www.microsoft.com/security/online-privacy/safely-dispose-computers-and-devices.aspx.

Data retention. As discussed in this paper, the policies, timeframes, and methods for storing, archiving, and retrieving data. Data retention policy should reflect the data classification model and data retention rules that apply to the data that is being retained. For example, highly sensitive data may be retained for a longer periods than data that is less sensitive. More information is available at http://technet.microsoft.com/en-us/library/jj574217.aspx.

Separation of duty. As discussed in this paper, the division of responsibilities in an IT environment that helps ensure that no one person can use IT resources for their personal benefit or cause IT-related outcomes that are detrimental to the organization. One of the most common ways to achieve separation of duty is to use a role-based access control system for authorization. More information is available at http://msdn.microsoft.com/en-us/library/windows/desktop/aa379318(v=vs.85).aspx.

Spam. Any kind of unwanted online communication. The most common form of spam is unwanted email, but text message spam, instant message spam (sometimes known as spam), and social networking spam also exist. Some spam is annoying but harmless, but sometimes spam is used in identity theft or other types of fraud. www.microsoft.com/security/resources/spam-whatis.aspx

Structured data. Data that is typically human readable and able to be indexed by machine. This data type incudes databases and spreadsheets. More information is available at http://msdn.microsoft.com/en-us/library/aa289148(v=vs.71).aspx.

Token. An item that is used to authenticate a username or user ID. A token can be something a user possesses, such as a card key, something that is biometrics-based, such as a fingerprint, retinal scan, or voice print, or something that is known, such as a password. More information is available at http://technet.microsoft.com/en-us/library/cc759267(v=WS.10).aspx.

Unstructured data. Data that is not human readable and is difficult to index. This data type includes source code, binaries, and documents, and can include such things as email because the data is typically randomly managed.