imperva

CAPABILITY BRIEF

Imperva Data Classification

Evaluate and classify sensitive data

Knowing where sensitive data resides is an organization's first step best security practice and an essential part of any security project. Without clear knowledge of the assets that need to be protected it can be difficult to create effective protection policies.

Imperva Data Classification identifies and classifies sensitive data so that you know what sensitive information assets reside in your enterprise databases. The challenge of data classification often lies in the complex mix of legacy, homegrown and third-party applications that run in your business and Imperva Data Classification makes it easy to uncover sensitive data.

Simplify the process

Imperva Data Classification comes pre-built with a number of definitions for commonly regulated data types, such as Personally Identifiable Information (PII), or financial information, and also allows customers to create their own definitions. The out-of-the-box predefined pattern templates quickly locate and identify a wide range of sensitive data, including but not limited to:

- Credit card numbers
- Birth dates
- Healthcare codes
- Identification numbers
- Social security numbers/ National ID
- Names
- Addresses
- Phone numbers
- Financial fields (account information, transactions)

Understanding the nature of your sensitive data and the context in which it resides enables you to put appropriate data privacy and security controls in place.

KEY FEATURES AND BENEFITS:

Protect your data more efficiently

Automate database discovery and location of sensitive data

Categorize data by type, sensitivity and value

Streamline regulatory compliance efforts

Reduce resource impact of database projects

CLASSIFIED DATABASE DATA

SCAN	DATA TYPE	DB	SCHEMA	TABLE	TABLE TYPE	TABLE STATUS	SITE	SERVER GROUP	SERVICE	DB	TABLE GROUP	TG STATUS	SENSITIVE	DATE	ACTION
DB2 Classification	Insurance Group Number	S...	TESTDB	CONT...	Table	New	ESXi Lab	DB2 WIN2...	DB2 10.1		ESXi Lab - DB2...	Existing	✓	09/27/2012 21:44	Accepted By User

COLUMN NAME	DB TYPE	LENGTH	DISCOVERY ACCURACY	DISCOVERY RULE NAME	DISCOVERY METHOD	SAMPLES
insurancegroupnum	VARCHAR	255	91%	Insurance Group Number	Content Based Search	

SCAN	DATA TYPE	DB	SCHEMA	TABLE	TABLE TYPE	TABLE STATUS	SITE	SERVER GROUP	SERVICE	DB	TABLE GROUP	TG STATUS	SENSITIVE	DATE	ACTION
Oracle Classification	Email Address	xe	APEX_040000	APEX_040000	View	New	ESXi Lab	Oracle WIN2...	Oracle 11gR2		ESXi Lab - Oracle	Existing	✓	12/18/2012 15:08	Accept

COLUMN NAME	DB TYPE	LENGTH	DISCOVERY ACCURACY	DISCOVERY RULE NAME	DISCOVERY METHOD	SAMPLES
email	VARCHAR2	240		Email Address Col. Name	Name Based Search	

SCAN	DATA TYPE	DB	SCHEMA	TABLE	TABLE TYPE	TABLE STATUS	SITE	SERVER GROUP	SERVICE	DB	TABLE GROUP	TG STATUS	SENSITIVE	DATE	ACTION
Oracle Classification	Address	xe	APEX_040000	WWV_...	Table	New	ESXi Lab	Oracle WIN2...	Oracle 11gR2		ESXi Lab - Oracle	Existing	✓	12/18/2012 15:08	Pending

COLUMN NAME	DB TYPE	LENGTH	DISCOVERY ACCURACY	DISCOVERY RULE NAME	DISCOVERY METHOD	SAMPLES
work_city	VARCHAR2	100		Town - Column Name	Name Based Search	

SCAN	DATA TYPE	DB	SCHEMA	TABLE	TABLE TYPE	TABLE STATUS	SITE	SERVER GROUP	SERVICE	DB	TABLE GROUP	TG STATUS	SENSITIVE	DATE	ACTION
Oracle Classification	Payment Card	xe	HR	CONT...	Table	New	ESXi Lab	Oracle WIN2...	Oracle 11gR2		ESXi Lab - Oracle	Existing	✓	12/18/2012 15:08	Accepted Automatically

COLUMN NAME	DB TYPE	LENGTH	DISCOVERY ACCURACY	DISCOVERY RULE NAME	DISCOVERY METHOD	SAMPLES
ccnum	VARCHAR2	16	100%	Credit Card No. - Content	Content Based Search	

Figure 1: A database classification scan is easy to configure using customer or predefined data types.

In addition, Imperva Data Classification analyzes sensitive data relationships by using heuristics and statistical analysis. By automating the identification of data relationships, Imperva Data Classification significantly reduces the manual effort and enables a more efficient sensitive data analysis process, allowing you to detect and audit changes to the sensitive data landscape over time.

Classification for risk mitigation

The Imperva data classification tool complements other Imperva data discovery tools and can be easily configured to scan the entire network and identify servers and services to which data monitoring, data analytics and data protection can then be applied. This ensures that your most critical data is safeguarded from unauthorized access and targeted attacks.

After the initial scope has been defined, continuous monitoring of changes to servers and other enterprise assets will identify any new instances of servers containing sensitive data. Scans can be scheduled periodically to check that sensitive data is not being stored in new locations on existing servers.

DATA TYPE	REGULAR EXPRESSION
US Social Security number	^([0-6]\d{2}|7[0-6]\d|77[0-2])([\-]?(\d{2})\2(\d{4})$
Credit Card numbers for AMEX, VISA, Mastercard	^((4\d{3})|(5[1-5]\d{2}))(-?|\040?)(\d{4}(-?|\040?)){3}|^(3[4,7]\d{2})(-?|\040?)\d{6}(-?|\040?)\d{5}
Email address	^[A-Z0-9._%+-]+@[A-Z0-9.-]+\.[A-Z]{2,4}$

Figure 2: Sample regular expressions for sensitive data

How data classification works

Data classification can be performed based on content, context, or user selections:

- Content-based classification – involves reviewing the content of database tables, and classifying it.
- Context-based classification – involves classifying database columns based on metadata such as the column and table name.

Imperva Data Classification uses regular expression rules, pattern matching, and predefined name-based or content-based data classification types as it scans database contents to tag matches that it finds. Customized, customer-specific data classification types can be added and scans can be scheduled and repeated to ensure on-going awareness of types of data within an organization's databases.

Creating your data classification policy

A data classification policy defines who is responsible for data classification – typically by defining Program Area Designees (PAD) who are responsible for classifying data for different programs or organizational units.

The data classification policy should consider the following questions:

- Which person, organization or program created and/or owns the information?
- Which organizational unit has the most information about the content and context of the information?
- Who is responsible for the integrity and accuracy of the data?
- Where is the information stored?
- Is the information subject to any regulations or compliance standards, and what are the penalties associated with non-compliance?

Data classification can be the responsibility of the information creators, subject matter experts, or those responsible for the correctness of the data.

The policy also determines the data classification process: how often data classification should take place, for which data, which type of data classification is suitable for different types of data, and what technical means should be used to classify data. The data classification policy is part of the overall information security policy, which specifies how to protect sensitive data.

Data classification and compliance

Data Classification supports compliance as it categorizes and tracks your most business critical data. Imperva provides out-of-the-box regular expressions that automatically identify many of the data types covered by industry-specific regulatory mandates such as SOX, HIPAA, PCI DSS, CCPA and GDPR which may require classification of different data attributes.

Data classification in the cloud or hybrid environment

Imperva Cloud Data Security provides visibility, oversight and security of your data in all the places it may lie, on-premises, or in an AWS or Azure public cloud, enabling you to comprehensively categorize your data enterprise-wide by sensitivity and determine the risks associated with it.

Classification benefits an entire data governance strategy

When properly executed, data classification makes all of your compliance, security and operational data governance practices better. You will save time and reduce the expense of compliance management and audits because you have prioritized the relevant data and won't waste effort on what's not regulated.

Implementing a robust data classification solution allows organizations to make better-informed decisions on how to protect their sensitive data. Your security analytics can help you understand what users are doing with the most sensitive data and whether any of the users exhibit dangerous behaviors. Security policies can use more granular controls for sensitive data, and utilize rules that prevent abusive or malicious actions such as downloading an excessive number of sensitive data records.

Imperva is an analyst-recognized, **cybersecurity leader** championing the fight to **secure data and applications** wherever they reside.

Copyright © 2020 Imperva. All rights reserved