Welcome to IBM Security Guardium Analyzer

To help you get started with IBM Security Guardium Analyzer, please refer to these frequently asked questions

OVERVIEW

What is IBM Security Guardium Analyzer?
Guardium Analyzer is a software-as-a-service offering that helps users efficiently find regulated data (such as PII and personal and sensitive personal data), understand data and database exposures, assess risk, and act to address issues and minimize risk.

How does IBM Security Guardium Analyzer work?
The service applies next-generation data classification, as well as vulnerability scanning, to uncover risks associated with such data in cloud-based and on-premises databases. It then applies risk scoring to the classification and scanning results to identify and prioritize the databases that may be most likely to fail an audit, so you can act to minimize your risk.

What core features are part of Guardium Analyzer?
Guardium Analyzer is targeted for use by compliance managers, IT managers, and DBAs. The key features delivered as part of Guardium Analyzer include:

- Connectivity to cloud and on-premises databases
- Next-generation data classification capabilities
- Vulnerability scanning
- Risk scoring
- Prioritized remediation recommendations
- Progress dashboard
- Reports for auditors
- Multi-regulation support

For more details about these capabilities, please visit the Details tab on the IBM Security Guardium Analyzer Marketplace page.

Is there a free trial for Guardium Analyzer?
Yes – today, we offer a "freemium" version of Guardium Analyzer that supports unlimited scans for up to three databases for as long as you want. Please go to the IBM Security Guardium Analyzer Marketplace page to register and get access. Please note to successfully set up the trial, you need to have database access privileges to scan the desired database(s) AND access to a Windows machine, to download and run the IBM Security Guardium Data Connector.

If I am already a Guardium Data Protection customer, what added value can I get from Guardium Analyzer?

The added value an existing Guardium client can obtain from leveraging Guardium Analyzer is three-fold:

1. Guardium Analyzer can provide an easy-to-consume picture of data risk when layered on top of Guardium Data Protection, allowing clients to quickly and continuously assess risks related to the structured data that is being tested, monitored, and protected by Guardium Data Protection and Guardium Vulnerability Assessment.
2. Guardium Analyzer's next generation classification capabilities improve classification accuracy
3. For clients who have purchased Guardium Data Protection but not yet deployed, Guardium Analyzer can serve as an efficient way to prioritize deployment of protective measures by highlighting the databases that may contain the highest levels of risk. Analyzer results may be feed into Guardium Data Protection, to fine-tune data protection policies.

How is Guardium Analyzer priced and sold?

For pricing and packaging information, please visit the IBM Security Guardium Analyzer Marketplace page and click on the Purchase tab. Clients have the option of purchasing Guardium Analyzer directly from the Marketplace page.

GETTING STARTED

What data sources does Guardium Analyzer scan?
Today, Guardium Analyzer supports the following data sources: IBM Db2, Oracle, AWS RDS Oracle, Informix, Microsoft SQL Server, MySQL, IBM Db2 on IBM Db2 for zLinux, and IBM Integrated Analytics System (IIAS) (also known as Netezza Sailfish).

In addition to supporting on-prem data sources, Guardium Analyzer can also scan those data sources that are installed on an Infrastructure-as-a-Service (IaaS) solution (a cloud vm, for example). Guardium Analyzer *may* work with other databases hosted by a cloud provider, but they are not officially supported at this time. More platforms will be added over time. For a full listing and supported versions, refer to:
https://www.ibm.com/support/knowledgecenter/SS2RDF/GuardiumAnalyzer/sys_req.html

What do I need to have in place to use Guardium Analyzer (either the Trial or the paid offering)?
You need to have access to a Windows-based server to run the downloaded IBM Security Guardium Data Connector, which will connect to your databases and to Guardium Analyzer. You also must have user-level access to the databases you wish to connect. You can download the connector on MacOS but you need to run the executable on a Windows-based machine or virtual machine.

What is the ideal set up recommendation for the best Guardium Analyzer trial experience?
We recommend using Google Chrome for accessing Guardium Analyzer. As mentioned above, users also must have access to a Windows machine to download and run the Data Connector, as well as user entitlement to connect to and scan their databases.

Are there any deployment configuration recommendations, recommended not-to-exceed database sizes, or other specifics that can help me maximize performance?
Allow enough time for the scan windows you are setting up, so that the connector can scan all connected databases. We recommend you determine the average scan time for your databases first, before deciding how many databases to connect, and then factor in an appropriate scan window; it is best to establish a baseline by scanning a sample database, and then using that scan time to derive what expected time would be for larger databases. Note that poorly optimized database tables will contribute to longer scan times.

How do you handle non-supported DBs? Can custom connections be developed for non-supported DBs?

There is no option to add drivers or connections for unsupported DBs - we will be adding support for more DBs as we go. Customers may try to connect to unsupported databases or databases that may no longer be supported. However, if the connection to such a database is successful, customers will be using those databases at their own risk, the scan results likely will not be accurate because the database may not be getting the latest security patches and updates, and IBM cannot provide support or product updates specific to those databases. For a full list of supported platforms, refer to: the IBM Security Guardium Analyzer Knowledge Center.

THE GUARDIUM ANALYZER DATA CONNECTOR

What does the IBM Security Guardium Data Connector do?
The Data Connector is provided with the Guardium Analyzer service, and it helps clients efficiently connect to their cloud and on-premises databases to uncover regulated data and vulnerabilities related to that data. Clients can connect to multiple databases simultaneously using one Data Connector, and nothing needs to be installed on database servers to do so. As part of the connection and scanning process, encryption techniques are applied to protect the data, and no personal data is uploaded to the cloud. For more details about the Data Connector, watch this short video.

How do I install the Data Connector, and where does it sit?
The connector should be installed on a Windows-based server within your local data center. After it is installed it will be able to connect to your on-premises and cloud databases and to Guardium Analyzer. DO NOT install the Data Connector on the database server(s). Think of it as a secure gateway to the service.

How long will it take me to download the Data Connector?
It can take a few minutes to download the Data Connector. In some regions of the world, it may take up to 10 minutes or more to download the Data Connector (depending on internet connectivity and speeds).

What is the ideal configuration for the Data Connector? What kind of performance should I expect?

- Do not install the Data Connector on database servers.
- We recommend a windows server with at least 8GB of RAM and 4 cores to run the Guardium Data Connector.
- On a single windows server, we recommend only having 1 connector installed. You may have multiple data connectors installed throughout your environment, but each should be on its own windows server.
- When you set your scan window, we recommend scheduling scans for off-peak hours. While the scans are running, Analyzer will use most of the available CPUs.
- Scan times will depend upon the number of tables in the database(s) that you connect (and number of columns within those tables), as well as the latency and network speed between the connector and the database(s). Please allow for a large enough scan window to allow database scanning to complete. As an example, in our lab environment, it took 7 hours to scan a database with 500,000 tables. For a database with a few hundred columns to scan across all tables and on a local LAN, the scan time is typically less than 10 minutes.

What type of access should my Windows server have to successfully install the Data Connector?

The Windows server where the Data Connector is installed should have public network access, be able to access https://www.datarisk.dsoc.ibm.com, and log in to that link using the IBM ID used to sign up for the trial. To verify this, launch https://www.datarisk.dsoc.ibm.com on Chrome/Firefox and ensure you can navigate to the URL and log in using that IBM ID. Finally, please note that when installing the Data Connector, you must do so as an admin.

How many DBs can be physically connected to one data connector? At what point do you recommend installing another connector?

We recommend connecting no more than 100 DBs to a single Data Connector. You should install additional connectors after that point. If you know that the databases you are connecting to the Data Connector are filled with very dense tables, we would recommend connecting fewer than 100 databases per Data Connector, for performance purposes.

GUARDIUM ANALYZER AND MY DATA

Does Guardium Analyzer move any of my sensitive data to the cloud?
No. Guardium Analyzer does not move any sensitive data to the cloud. Guardium Analyzer scans for regulated data and vulnerabilities in your databases, but only metdata is sent to the cloud-based Guardium Analyzer dashboard.

What data is being sent to the cloud dashboard from the connector? What parameters are sent? Are they sent over a secure channel?
All communication is over HTTPS/TLS only. The data returned is only metadata: table name, column name, name of pattern found (e.g. tb_employees.emp_name: "First Name"), and any VA test that failed (ID only). While HTTPS/TLS is the industry standard for most internet communication and is highly secure, even if a hacker is able to decrypt this information, the transmitted data does not provide any value as the hacker cannot use this information to pinpoint the location of the client Database. Also, the VA ID is only understood by the Guardium Analyzer service and cannot be correlated back to the actual CVE ID.

It is possible to preview the scan results metadata generated by the Guardium Analyzer Data Connector before the metadata is shared with the Guardium Analyzer cloud dashboard?
Yes, it is possible to preview the results in a preliminary format before they are shared with the Guardium Analyzer cloud dashboard and receive their risk score and prioritization. To do this, you would set the Data Connector to ‘local’ mode. Once you are satisfied with the raw results, you can deselect ‘local’ mode and share the results with the dashboard.

How is the data about databases (metadata) being stored locally on the connector? Is it encrypted if sensitive information is gathered through the scan?
The connector stores only the connection data (db type, port/ip user/password) in a local encrypted database. No scan data is saved.

THE GUARDIUM ANALYZER CLASSIFICATION ENGINE

How is Guardium Analyzer's data classification engine different?
The new classification engine inside Guardium Analyzer is based on System T, which is part of the IBM Watson offerings. Please note that System T does not involve or include machine learning or artificial intelligence.

Here's a breakdown of different classification methods available today:
Catalog-based search (no customization):

- Based on searching table column names only
- Column names often do not match content, so results aren't accurate
- Example: Sensitive at-risk data identified via metadata search as 'SSN' actually had no SSN info in table. If column names were changed to 'A' and 'B', if SSN info DID appear inside the table, no results would be found through metadata search

Regular Expression:

- Catalog-based search, or "data sampling' search, with simple regular expression
- Better accuracy than metadata search
- Greater richness in the number of rules and expressions considered
- Can look for tokens, but does not allow for dictionary lookup

AQL and System T (next generation data classification) – inside Guardium Analyzer:

- Extracts data from a table, crawls it, applies taxonomy, and supports dictionary lookup
- Data classification rules and dictionaries can be more expressive
- Performs rules-based matching then applies a "checksum" algorithm to validate the match, which creates higher accuracy

How can I add or delete or import my own custom data classification patterns?
To change, update, or import new data classification patterns, you must either be using the Guardium Analyzer Trial (for use with up to 3 data sources) or have access to the Guardium Analyzer Professional Plan. Classification pattern customization is not supported for Standard Plan subscribers.

For those with the Trial or the Professional Plan, you can add either a regex or a dictionary based custom pattern to your classification by going to settings -> manage patterns -> Add pattern.

To delete a custom data classification, go to settings -> manage patterns. Then you can highlight the pattern you want to remove and click on the 'garbage can' icon on the bottom right side of the page.

Are the classification and vulnerability scans different scans, or can these be combined as one scan? If one, can data and vulnerability scans be segregated?
The goal of a scan is to create an overall risk assessment that is based on the risks associated with the amount of regulated and/or sensitive data in the environment combined with the vulnerabilities associated with each database within which the regulated and/or sensitive data resides. Because the scan results are used to derive the risk score, the classification results and the vulnerability assessment results are tied together and may not be separated.

SUPPORTED LANGUAGUES

What languages does the Guardium Analyzer classification engine support?
Today, the Guardium Analyzer classification dictionary supports English, French, German, Spanish, Italian, Japanese (Kanji and Hiragana), Dutch, Danish, Australian PII, Brazilian Portuguese, and more. Additional languages will be added over time.

What languages is Guardium Analyzer globalized and translated into (the product itself, the UIs, etc)?
Guardium Analyzer went live in English, and it has been globalized and translated for: French, Spanish, German, Japanese, Italian, Polish, and Korean. Additional languages may be added over time.