

IBM StoredIQ Glossary

Gateway

The gateway communicates between the data servers and the application stack. The application stack polls the gateway for information about the data on the data servers. The data servers push the information to the gateway.

Data servers

A data server obtains the data from supported data sources and indexes it. By indexing this data, you gain information about unstructured data such as file size, file data types, file owners. The data server pushes the information about volumes and indexes to the gateway so it can be communicated to the application stack. Multiple data servers feed into a single gateway.

DataSeter - classic

Data servers can be categorized in two types: DataSeter - Classic and DataSeter - Distributed. DataSeter - Classic refers to the regular data servers. It uses either the current PostgreSQL or Lucene index as an index.

DataSeter - distributed

The distributed data server uses an Elasticsearch cluster instead of an embedded Postgres database. It increases the scalability and flexibility of the IBM StoredIQ deployment in a way that it can manage much larger amounts of data. Without adding more data servers, data that is managed by the IBM StoredIQ deployment can be increased by adding new nodes to the Elasticsearch cluster. Search queries perform better on DataSeter - Distributed.

The application stack provides the user interface for the IBM StoredIQ Administrator, IBM StoredIQ Data Workbench, IBM StoredIQ Insights, and the IBM StoredIQ Policy Manager products. The synchronization feature for integration with a governance catalog is also part of the application stack

Auto-classification

Automated document categorization, what IBM StoredIQ refers to as autotclassification models, integrates the IBM® Content Classification's classification model into the IBM StoredIQ infoet-generation process. Data Experts can use IBM Content Classification to train a classification model, which is then registered with IBM StoredIQ Administrator. The registered classification model can be applied to an existing infoet in IBM StoredIQ Data Workbench to generate new metadata for the objects in the infoet. Metadata can be used in rule-based filters to create new infoets.

Cartridges

Cartridges are compressed files that contain analysis logic. When you add a cartridge to IBM StoredIQ AppStack, it can detect new data in documents during indexing and make these new insights searchable. For example, a sensitive pattern cartridge can enable IBM StoredIQ to detect passport numbers, phone numbers, and other IDs.

To apply the analysis logic contained in the cartridge, you must run a Step-up Analytics action that uses the cartridge on an infoSet. IBM StoredIQ examines all documents in the infoSet, applies the analytics, and then stores the analysis results in the IBM StoredIQ index.

Connector API SDK

A connector is a software component of IBM StoredIQ that is used to connect to a data source such as a network file system and access its data. Using IBM StoredIQ Connector API SDK, developers of other companies can develop connectors to new data sources outside the IBM StoredIQ development environment. These connectors can be integrated with a live IBM StoredIQ application to index, search, manage, and analyze data on the data source.

IBM StoredIQ Insights

IBM StoredIQ Insights provides dynamic and interactive filtering for your data with easy access to all metadata and instant plain-text preview of document content for full-text indexed volumes.