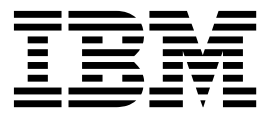


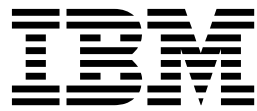
IBM StoredIQ Platform
Version 7.6.0.13

Overview Guide



IBM StoredIQ Platform
Version 7.6.0.13

Overview Guide



Note

Before using this information and the product it supports, read the information in Notices.

This edition applies to Version 7.6.0.13 of product number 5724M86 and to all subsequent releases and modifications until otherwise indicated in new editions.

© **Copyright IBM Corporation 2001, 2017.**

US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Table of contents

IBM StoredIQ Platform product library	v	IBM StoredIQ Administrator	2
Contacting IBM StoredIQ customer support	vii	IBM StoredIQ Data Workbench	4
Contacting IBM	vii	IBM StoredIQ eDiscovery	5
How to send your comments	ix	IBM StoredIQ Policy Manager	5
Overview of the IBM StoredIQ components	1	Notices	7
Components of the IBM StoredIQ solution	1	Privacy policy considerations	9
IBM StoredIQ interface components	1	Terms and conditions for product documentation.	10
		Trademarks	10
		Index	13

IBM StoredIQ Platform product library

The following documents are available in the IBM® StoredIQ Platform product library.

- *IBM StoredIQ Platform Deployment and Configuration Guide, SC27-6386*
- *IBM StoredIQ Platform Overview Guide, GC27-6398*
- *IBM StoredIQ Platform Data Server Administration Guide, SC27-5692*
- *IBM StoredIQ Administrator Administration Guide, SC27-5688*
- *IBM StoredIQ Data Workbench User Guide, SC27-5691*
- *IBM StoredIQ eDiscovery User Guide, SC27-5693*
- *IBM StoredIQ Policy Manager User Guide, SC27-5694*

Contacting IBM StoredIQ customer support

For IBM StoredIQ technical support or to learn about available service options, contact IBM StoredIQ customer support at this phone number:

Support and assistance

- 1-866-227-2068

To e-mail IBM StoredIQ customer support, use this email address:

- storediqsupport@us.ibm.com

For information about how to contact IBM, see the Contact IBM web site at <http://www.ibm.com/contact/us/>

IBM Knowledge Center

The IBM StoredIQ publications can be found from IBM Knowledge Center.

PDF publications

The IBM Publication Center site offers customized search functions to help you find all the IBM publications you need.

Contacting IBM

For general inquiries, call 800-IBM-4YOU (800-426-4968). To contact IBM customer service in the United States or Canada, call 1-800-IBM-SERV (1-800-426-7378).

For more information about how to contact IBM, including TTY service, see the Contact IBM website at <http://www.ibm.com/contact/us/>.

How to send your comments

Your feedback is important in helping to provide the most accurate and highest quality information.

Send your comments by using the online reader comment form at https://www14.software.ibm.com/webapp/iwm/web/signup.do?lang=en_US&source=swg-rcf.

Overview of the IBM StoredIQ components

The following sections describe three components of the IBM StoredIQ solution and the interfaces that IBM StoredIQ products provide to use the services of the IBM StoredIQ Platform.

Components of the IBM StoredIQ solution

The three components of the IBM StoredIQ solution are the gateway, data servers, and application stack (AppStack).

The 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

The IBM StoredIQ Data Server helps understand the data landscape of the enterprise. It 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 servers push 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. In addition to an administrator user interface, administrators can deploy the IBM StoredIQ Desktop Data Collector and index desktops from the data server.

Application stack

The application stack provides the user interface for the IBM StoredIQ Administrator, IBM StoredIQ Data Workbench, IBM StoredIQ eDiscovery, and the IBM StoredIQ Policy Manager products.

IBM StoredIQ interface components

The IBM StoredIQ interface components include IBM StoredIQ Platform Data Server, IBM StoredIQ Administrator, IBM StoredIQ Data Workbench, IBM StoredIQ eDiscovery, IBM StoredIQ Policy Manager, and IBM StoredIQ Desktop Data Collector.

IBM StoredIQ Platform Data Server

IBM StoredIQ Platform Data Server user interface provides access to data server functionality. It allows administrators to view the dashboard and see the status of the jobs and system details. Administrators can manage information about servers and conduct various configurations on the system and application settings.

IBM StoredIQ Administrator

IBM StoredIQ Administrator monitors and manages the distributed infrastructure at a client site. IBM StoredIQ Administrator sits between the IBM StoredIQ Platform interface and the applications and facilitates the transfer and communication of information. IBM StoredIQ Administrator understands and manages IBM StoredIQ Platform concepts such as volumes, indexes, harvests, and configurations. At the same time, it manages the application concerns such as infoSet lifecycle and creation,

volume configuration, and action and target-set management. To this end, it is divided into two sections, that is, platform and application, so that the administrators know where to accomplish a task.

IBM StoredIQ Data Workbench

With IBM StoredIQ Data Workbench, you can visualize the indexed data and identify potential red-flag issues to know how much and what types of data you have on different types of servers. It alerts people about potentially interesting or useful data. It helps ensure that the data of an enterprise is an asset, not a liability.

IBM StoredIQ eDiscovery

IBM StoredIQ eDiscovery is integral to the eDiscovery process. It helps discover, identify, refine, preserve, and collect data that is relevant to legal matters. With IBM StoredIQ eDiscovery, you can conduct discovery work more efficiently and ensure that you capture the appropriate data at the same time.

IBM StoredIQ Policy Manager

IBM StoredIQ Policy Manager acts on data in an automatic fashion at scale, running policies that affect data objects without requiring review.

IBM StoredIQ Desktop Data Collector

IBM StoredIQ Desktop Data Collector deploys from the IBM StoredIQ Platform Data Server Administrator interface. It indexes desktops as volumes. The volumes appear in the Data Server Administrator interface and IBM StoredIQ Data Workbench, where the data can be analyzed and acted upon, and the IBM StoredIQ Administrator application.

IBM StoredIQ Platform uses a non-invasive deployment, specifically:

- Lightweight client deployment with no browser plug-ins and no client-side UI installations. All major browsers, such as Internet Explorer and Firefox, are supported natively.
- One (common) index that is shared across all use cases. This index supports a wide variety of data sources and is used by application dashboards for many use cases.
- No agents are placed on data sources. The native API/protocol is used whenever it is possible. Third-party or custom connectors are used only when necessary.

Note: Agents are used for desktop collection.

IBM StoredIQ Administrator

IBM StoredIQ Administrator helps you to manage global assets common to the distributed infrastructure behind IBM StoredIQ Platform applications.

IBM StoredIQ Administrator provides at-a-glance understanding of the different issues that can crop up in the IBM StoredIQ Platform environment. These views are unique to the IBM StoredIQ Administrator application as they provide an overview of how the system is running. They allow access to various pieces of information that are being shared across applications or allow for the management of resources in a centralized manner.

The administrator is the person responsible for managing the IBM StoredIQ Platform. This individual has strong understanding of data sources, indexes, data servers, jobs, infosets, and actions. This list provides an overview as to how IBM StoredIQ Administrator works:

- **Viewing data servers and volumes:** Using IBM StoredIQ Administrator, the Administrator can identify what data servers are deployed, their location, what data is being managed, and the status of each data server in the system. Volume management is a central component of IBM StoredIQ Platform. IBM StoredIQ Administrator also allows the Administrator to see what volumes are currently under management, which data server is responsible for that volume, the state of the volume after indexing, and the amount and size of information that is contained by each volume. Administrators can also add volumes to and delete volumes from data servers through this interface.
- **Scheduling harvests:** Harvesting, which can also be referred to as indexing, is the process or task by which IBM StoredIQ Platform examines and classifies data in your network. Using IBM StoredIQ Administrator, harvests can be scheduled, edited, and deleted.
- **Creating system infosets:** System infosets that use only specific indexed volumes can be created and managed within IBM StoredIQ Administrator. Although infosets are a core component of IBM StoredIQ Data Workbench, system infosets are created as a shortcut for users in IBM StoredIQ Administrator.
- **Managing users:** The user management area allows administrators to create users and manage users' access to the various IBM StoredIQ applications.
- **Configuring and managing actions:** An action is any process that is taken upon the data that is represented by the indexes. Actions are run by data servers on indexed data objects. Any errors or warnings that are generated as a result of an action are recorded as exceptions in IBM StoredIQ Data Workbench.

Note: Actions can be created within IBM StoredIQ Administrator and then made available to other IBM StoredIQ applications such as IBM StoredIQ Data Workbench.

- **Managing target sets:** Provides an interface that allows the user to set the wanted targets for specific actions that require a destination volume for their actions.
- **Reports:** IBM StoredIQ Administrator provides a number of built-in reports, such as summaries of data objects in the system, storage use, and the number of identical documents in the system. You can create custom reports, including Query Analysis Reports for e-discovery purposes, and automatically email report notifications to administrators and other interested parties.
- **Auto-classification:** Automated document categorization, what IBM StoredIQ refers to as auto-classification models, integrates the IBM® Content Classification's classification model into the IBM StoredIQ Platform infoset-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 infoset in IBM StoredIQ Data Workbench to generate new metadata for the objects in the infoset. Metadata can be used in rule-based filters to create new infosets.
- **Cartridges:** Cartridges are compressed files that contain analysis logic. By adding 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. Step-up Analytics runs a selected cartridge on an infoset. IBM StoredIQ examines all documents in the infoset, applies the analytics that is contained in the cartridge to the document, and then stores the analysis results in the IBM StoredIQ index.

- **Managing concepts:** Provides the ability to relate business concepts to indexed data.
- **Managing Mule scripts:** Helps you to create Mule scripts and upload script packages. These Mule scripts are used by IBM StoredIQ Policy Manager to create policies using the automation workflow.
- **DataServer - Distributed:** Data servers can be categorized in two types: DataServer - Distributed and DataServer - Classic. 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 DataServer - Distributed.
- **OneDrive for Business:** IBM StoredIQ supports index, copying from, and exporting from OneDrive for Business to CIFS and NFS shares.
- **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 Data Workbench

The following section provides a high-level explanation of what IBM StoredIQ Data Workbench is and its potential uses.

Big data is a pervasive problem, not a one-time occurrence. It is easy for most companies to realize that big data is problematic, but it is hard to identify what problems they have. Big data is all about the unknown, but the unknown cannot be off limits. IBM StoredIQ Data Workbench can help you learn about your data, make educated decisions with your most valuable asset, and turn your company's most dangerous risk into its most valuable asset.

IBM StoredIQ Data Workbench is a data visualization and management tool that helps you to actively manage your company's data. It helps you to determine how much data you have, where it is, who owns it, and when it was last used. When you have a clear understanding of your company's data landscape, IBM StoredIQ Data Workbench helps you take control of data. You can make informed decisions about your data and act on that knowledge by copying, copying to retention, or conducting a discovery export.

Here are just some of the examples that you can use IBM StoredIQ Data Workbench.

- You need to find all company email that is sent from or received by Eileen Sideways (esideways@thecompany.com). You can use IBM StoredIQ Data Workbench to find all email and then copy that data to a predefined repository. You can also use IBM StoredIQ Data Workbench to find all of the esideways@thecompany.com email that occurred between specific dates and then make that email available for review.
- As an administrator, you want to rid your networks and storage of unused data. You can use IBM StoredIQ Data Workbench to find all files that were not modified in more than five years.

- You want to find all image files that are created in 2007. Not only can IBM StoredIQ Data Workbench find all image files that were created in 2007. It also shows how much space they occupy on your network.
- A user needs to understand how data about Windows is being retained. Using IBM StoredIQ Data Workbench, you can provide that user with a visual overview of the number of objects that are retained and a breakdown of files per data source. Additionally, you can apply overlays to show the user if those files contain forbidden information such as credit-card numbers or Social Security numbers.

IBM StoredIQ eDiscovery

The following section provides a high-level explanation of what IBM StoredIQ eDiscovery is, by whom it is used, prerequisites, and its potential uses.

IBM StoredIQ eDiscovery is a user application that helps legal users during the initial phases of the eDiscovery process. By helping you to capture needed electronic data, you can use IBM StoredIQ eDiscovery to communicate easily with other users about the status and review process of work that is being done. IBM StoredIQ eDiscovery does not drive the eDiscovery process, but instead helps legal users to control and communicate those processes more effectively.

Here are just some of the examples you can use IBM StoredIQ eDiscovery.

- You need to find all electronic information about an upcoming personal-injury lawsuit. You can use IBM StoredIQ eDiscovery to create a matter for the suit, and then create boxes to contain email and reports about the case.
- As a legal user, you want to see the status of all currently active matters. You can use the Matter Dashboard to visualize the different matters' states.
- As a data expert, you know that you need to respond to users in a timely fashion. Using IBM StoredIQ eDiscovery, requests are displayed on the IBM StoredIQ Data Workbench navigation.

Proactive eDiscovery consists of identifying and cataloging data or types of data that might eventually be responsive to legal matters. It collects and retains the data and then produces that data in a way that matches accepted or required legal practices.

Within IBM StoredIQ eDiscovery, the user is a non-technical user who needs to obtain certain pieces of data to complete a legal process. The IBM StoredIQ eDiscovery user knows the information or parameters for what kinds of data they want, but does not always know how to obtain that data.

Before you use IBM StoredIQ eDiscovery, ensure that prerequisites are met:

- The IBM StoredIQ Platform must be deployed, configured, and ready for use.
- IBM StoredIQ eDiscovery is dependent upon other IBM StoredIQ applications such as IBM StoredIQ Data Workbench and IBM StoredIQ Administrator.

IBM StoredIQ Policy Manager

IBM StoredIQ Policy Manager enables policy execution at scale.

IBM StoredIQ Policy Manager allows users to run mature policies and processes at scale across a wider range of data. The users can define and run systemwide policies, focusing on the execution of the process rather than understanding or reviewing affected data objects. Additionally, with IBM StoredIQ Policy Manager's

reports, you can record what actions were conducted, when they were conducted, and what data was affected by the policy's execution.

Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US*

For license inquiries regarding double-byte character set (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

*Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan*

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those

websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

*IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US*

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

The performance data discussed herein is presented as derived under specific operating conditions. Actual results may vary.

The client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

All IBM prices shown are IBM's suggested retail prices, are current and are subject to change without notice. Dealer prices may vary.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the

names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Each copy or any portion of these sample programs or any derivative work, must include a copyright notice as follows: © Copyright IBM Corp. 2004, 2015. All rights reserved.

Privacy policy considerations

IBM Software products, including software as a service solutions, ("Software Offerings") may use cookies or other technologies to collect product usage information, to help improve the end user experience, to tailor interactions with the end user or for other purposes. In many cases no personally identifiable information is collected by the Software Offerings. Some of our Software Offerings can help enable you to collect personally identifiable information. If this Software Offering uses cookies to collect personally identifiable information, specific information about this offering's use of cookies is set forth below.

[Depending upon the configurations deployed, this Software Offering may use [session] [and] [persistent] cookies that collect each user's [name,] [user name,] [password,] [profile name,] [position,] or [other personally identifiable information] for purposes of [session management,] [authentication,] [enhanced user usability,] [single sign-on configuration] [or other usage tracking or functional purposes.] These cookies [cannot be disabled] [can be disabled, but disabling them will also [likely] eliminate the functionality they enable].

If the configurations deployed for this Software Offering provide you as customer the ability to collect personally identifiable information from end users via cookies and other technologies, you should seek your own legal advice about any laws applicable to such data collection, including any requirements for notice and consent.

For more information about the use of various technologies, including cookies, for these purposes, See IBM's Privacy Policy at <http://www.ibm.com/privacy> and IBM's Online Privacy Statement at <http://www.ibm.com/privacy/details> the section entitled "Cookies, Web Beacons and Other Technologies" and the "IBM Software Products and Software-as-a-Service Privacy Statement" at <http://www.ibm.com/software/info/product-privacy>.

Terms and conditions for product documentation

Permissions for the use of these publications are granted subject to the following terms and conditions.

Applicability

These terms and conditions are in addition to any terms of use for the IBM website.

Personal use

You may reproduce these publications for your personal, noncommercial use provided that all proprietary notices are preserved. You may not distribute, display or make derivative work of these publications, or any portion thereof, without the express consent of IBM.

Commercial use

You may reproduce, distribute and display these publications solely within your enterprise provided that all proprietary notices are preserved. You may not make derivative works of these publications, or reproduce, distribute or display these publications or any portion thereof outside your enterprise, without the express consent of IBM.

Rights

Except as expressly granted in this permission, no other permissions, licenses or rights are granted, either express or implied, to the publications or any information, data, software or other intellectual property contained therein.

IBM reserves the right to withdraw the permissions granted herein whenever, in its discretion, the use of the publications is detrimental to its interest or, as determined by IBM, the above instructions are not being properly followed.

You may not download, export or re-export this information except in full compliance with all applicable laws and regulations, including all United States export laws and regulations.

IBM MAKES NO GUARANTEE ABOUT THE CONTENT OF THESE PUBLICATIONS. THE PUBLICATIONS ARE PROVIDED "AS-IS" AND WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT, AND FITNESS FOR A PARTICULAR PURPOSE.

Trademarks

IBM, the IBM logo, and [ibm.com](http://www.ibm.com) are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at <http://www.ibm.com/legal/copytrade.shtml>

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Java™ and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Other company, product, and service names may be trademarks or service marks of others.

Index

A

action 3
appstack 1

C

components of the IBM StoredIQ
solution 1

D

data server 1
Data Workbench
about 4
potential uses of 4
Desktop Agent 1

E

eDiscovery
about 5
potential uses of 5
prerequisites for using 5
exceptions 3

G

gateway 1

I

IBM StoredIQ 1
components 1
IBM StoredIQ Administrator 1
IBM StoredIQ Data Workbench 1
IBM StoredIQ eDiscovery 1
IBM StoredIQ Platform Data
Server 1
IBM StoredIQ Policy Manager 1
IBM StoredIQ Policy Manager
about 5

L

legal
notices 7
trademarks 10

N

notices
legal 7

T

trademarks 10



Printed in USA

GC27-6398-15

