



PUBLIC

SAP Information Steward

Document Version: 4.3 Support Package 2 (14.3.2.0) – 2024-02-09

Administrator Guide

Content

- 1 Accessing Information Steward for administrative tasks. 9**
- 2 Architecture. 10**
 - 2.1 Information Steward architecture overview. 10
 - Servers and services. 12
 - Services. 12
 - Web application server. 16
 - SAP BusinessObjects BI platform. 18
 - SAP BI platform components usage. 19
 - SAP Data Services components usage. 22
 - Metadata integrators. 23
 - 2.2 Distributed architecture. 25
 - Component communication channels. 25
 - Port assignments. 26
 - DSN-less and TNS-less connections. 27
 - 2.3 SAP integration. 29
 - 2.4 Information workflows. 29
 - Adding a table to a Data Insight project. 30
 - Profiling data 31
 - Scheduling and running a Metadata Management integrator source. 32
 - Creating a custom Cleansing Package with Cleansing Package Builder. 32
 - Creating a Data Cleansing Solution in Data Cleansing Advisor. 34
- 3 Securing SAP Information Steward. 36**
 - 3.1 Security Overview. 36
 - 3.2 Enterprise security. 36
 - Data protection and privacy. 37
 - Securing user data for Information Steward. 38
 - Storage of sensitive information. 38
 - Configuring Information Steward for SSL. 39
 - Reverse proxy servers. 43
- 4 Information Steward configuration settings. 44**
 - 4.1 Configuring Information Steward settings. 44
 - 4.2 Information Steward settings option descriptions. 45
- 5 Users and Groups Management. 52**
 - 5.1 Users and Groups overview. 52

5.2	Information Steward pre-defined users and groups.	52
5.3	Managing users in Information Steward.	54
	Creating users for Information Steward.	54
	Adding users and user groups to Information Steward groups.	56
	Denying access.	56
	Deleting users in Information Steward.	57
	Setting user e-mail address for notifications.	57
5.4	User rights in Data Insight.	58
	Data Insight predefined user groups.	59
	Type-specific rights for Data Insight objects	60
	Type-specific rights for source or failed data connections.	67
	Customizing rights on Data Insight objects.	67
5.5	User rights in Metadata Management.	72
	Metadata Management pre-defined user groups.	73
	Type-specific rights for Metadata Management objects.	73
	Assigning users to specific Metadata Management objects.	77
5.6	User rights in Metapedia.	79
	Metapedia pre-defined user groups.	79
	Type-specific rights for Metapedia objects	80
	Assigning users to specific Metapedia objects	82
5.7	User rights in Cleansing Package Builder.	83
	Group rights for cleansing packages.	84
5.8	User rights for Information Steward administrative tasks	84
	Viewing repository information.	85
5.9	User rights in Match Review.	86
	Match Review pre-defined user groups.	86
	Type-specific rights for Match Review.	87
6	Information Steward servers and services.	90
6.1	Verifying Information Steward servers are running.	91
6.2	Verifying Information Steward services	91
6.3	Server license measurement report.	93
	Generating a license measurement report.	94
6.4	Configuring Metadata Browsing Service and View Data Service	95
	Metadata Browsing Service configuration parameters.	96
	View Data Services configuration parameters.	98
6.5	E-mail notification configuration.	99
	Setting up e-mail notification configuration.	100
6.6	Job server group.	101
	Data Insight can use Data Services Job Servers.	102
	Configuring a Data Services Job Server for Data Insight.	103
	Displaying job servers for Information Steward.	104

	Removing a Job Server from the group.	104
6.7	Avoid exceptions in the Metadata Search service.	105
	Setting runtime parameters for the Metadata Search service.	106
7	Repository Management.	108
7.1	Viewing repository information.	108
7.2	Viewing Data Cleansing Advisor repository information.	108
7.3	Enabling Windows authentication and SSL/TLS for Microsoft SQL Server repositories.	109
7.4	Configuring SSL/TLS for SAP HANA repository.	111
7.5	Backing up your repository.	112
7.6	Repository utility overview.	112
	Creating a new repository.	113
	Recovering the repository.	114
	Upgrading the repository.	116
	Completing the repository recovery.	117
	Repository utility parameters.	118
8	Data Insight Administration.	123
8.1	Administration overview for Data Insight.	123
8.2	Data Insight performance settings.	124
8.3	Data Insight Connections.	125
	Defining a Data Insight connection to a database.	125
	Defining a Data Insight connection to an application.	149
	Execute options for ABAP program generation for SAP tables.	155
	Defining a Data Insight connection to a file.	156
	Displaying and editing Data Insight connection parameters.	157
	Deleting a Data Insight connection.	158
	Setting connection type and permissions to access failed data.	159
8.4	Data Insight projects.	160
	Creating a project.	160
	Editing a project description.	161
	Enterprise project.	162
	Deleting a project.	162
8.5	Data Insight tasks.	163
	Scheduling a task in the CMC.	164
	Export tasks for external scheduler.	166
	Recurrence options.	177
	Notification for task completion.	178
	Configuring notification when scheduling a task.	179
	Rule threshold notification.	181
	Capturing skipped records from rule filter.	182
	Monitoring a task.	183

	Pausing and resuming a schedule.	185
	Deleting failed data for a failed rule task.	186
	Common runtime parameters for Information Steward.	187
9	Metadata Management Administration.	189
9.1	Administration overview for Metadata Management	189
9.2	Configure metadata integrator sources.	189
	Configuring SAP BusinessObjects Enterprise integrator sources.	191
	Configuring SAP Business Warehouse integrator sources.	194
	Configuring SAP Data Services metadata integrator sources.	195
	Configuring SAP ERP Central Component integrator sources.	197
	Configuring Apache Hadoop integrator sources.	199
	Configuring SAP HANA integrator sources.	204
	Configuring SAP PowerDesigner integrator sources.	206
	Configuring Common Warehouse Model integrator sources.	210
	Configuring SAP Data Federator integrator sources	212
	Configuring sources for Meta Integration Metadata Bridge Metadata Integrator	212
	Configuring relational database system integrator sources.	213
	Configuring Excel Content integrator sources.	219
	Configuring Amazon Redshift integrator sources.	220
9.3	Manage integrator sources and instances.	222
	Viewing and editing an integrator source	223
	Deleting an integrator source	224
	Changing limit for storing integrator runs	225
	Integrator run required.	226
9.4	Running a Metadata Integrator	227
	Running a Metadata Integrator immediately	227
	Defining a schedule to run a Metadata Integrator	228
	Running SAP BusinessObjects Enterprise Metadata Integrator with Windows AD authentication.	229
9.5	Runtime parameters for integrator sources.	229
	Common runtime parameters for metadata integrators.	230
	Runtime parameters for specific metadata integrators.	230
	Changing runtime parameters for integrator sources.	247
9.6	Viewing integrator run progress, history, and log files.	248
9.7	Other options in the Integrator History page.	249
9.8	Troubleshooting	250
	Crystal Report message	250
	Desktop Intelligence document error	251
	Out of memory error.	251
	Parsing failure	251
	Parsing failure for derived table	251

	Unable to parse SQL	252
	Unable to retrieve SQL to parse	252
	Connection with Data Federator Designer	253
9.9	Grouping Metadata Integrator sources	253
	Creating source groups	253
	Modifying source groups	253
	Deleting source groups	254
9.10	Configuring BI Launch Pad for Information Steward.	254
9.11	Displaying user-defined reports in Information Steward.	255
10	Cleansing Package Builder Administration.	257
10.1	Changing ownership of a cleansing package.	257
10.2	Deleting a cleansing package.	257
10.3	Changing the description of a cleansing package.	258
10.4	Unlocking a cleansing package.	258
10.5	Content upgrade for person and firm cleansing packages.	259
	Content upgrade states.	260
	Content upgrade log file.	261
10.6	Cleansing package states and statuses.	261
	Cleansing package states.	262
	Cleansing package statuses.	263
11	Match Review Administration.	265
11.1	Match Review.	265
	Match Review workflow.	265
	Match result table.	266
	Match Review job status table.	267
11.2	Data Review configuration file.	268
11.3	Match Review connections.	269
	Defining a Match Review connection.	270
	IBM DB2 connection options for Match Review.	271
	Microsoft SQL Server connection options for Match Review.	272
	Oracle connection options for Match Review.	273
	SAP HANA connection options for Match Review.	273
	SAP ASE connection options for Match Review.	274
12	Information Steward Utilities.	276
12.1	Utilities overview.	276
	Compute and store lineage information for reporting.	279
	Recreate search indexes on Metadata Management.	280
12.2	Scheduling a utility.	280
12.3	Rescheduling a utility.	281

12.4	Running a utility on demand.	282
12.5	Monitoring utility executions.	282
12.6	Utility service notification alerts.	283
	Service monitor e-mail alert schedule.	284
	Activating e-mail alerts for utility service monitor.	285
12.7	Modifying utility configurations.	286
12.8	Additional utility configurations.	286
	Creating an additional utility configuration.	287
12.9	Deleting default schedule for Calculate Scorecard and Purge utilities.	288
13	Performance and Scalability Considerations.	289
13.1	Resource intensive functions.	289
13.2	Architecture.	290
13.3	Factors that influence performance and sizing.	292
	Data Insight.	292
	Metadata Management.	294
	Cleansing Package Builder.	295
13.4	Scalability and performance considerations.	296
	Scalability levels in deployment.	297
	Distributed processing.	297
	Scheduling tasks	300
	Queuing tasks.	301
	Degree of parallelism.	301
	Grid computing.	304
	Using SAP applications as a source.	306
	Multi-threaded file read.	307
	Data Insight result set optimization.	308
	Performance settings for input data.	308
	Settings to control repository size.	309
	Settings for Cleansing Package Builder.	312
	Settings for Metadata Management.	313
	Settings for Metapedia.	314
13.5	Best practices for performance and scalability.	315
	General best practices.	315
	Data Insight best practices.	316
	Metadata Management best practices.	319
	Cleansing Package Builder best practices.	320
	Match Review best practices.	320
13.6	SAP tables and ABAP-supported functions.	320
14	Migration.	322
14.1	Migration overview.	322

	Develop and test environments.	322
	Production environment.	323
	Lifecycle phases of a cleansing package.	324
14.2	Migrating between system landscapes.	326
	Objects supported for migration.	328
	Migrating CMS content using Promotion Management.	329
	Migrating the repository to a different database type.	332
14.3	SAP tables and ABAP program migration considerations.	333
	ABAP program naming conventions.	334
14.4	Object portability.	335
	Data Insight object portability.	335
	Metadata Management object portability.	336
	Cleansing Package Builder object portability.	337
	Match Review object portability.	337
	Information Steward general object portability.	338
14.5	Exporting and importing objects.	339
	Import and export custom attribute definitions.	339
	Exporting and importing Match Review configurations (task).	339
15	Supportability.	340
15.1	Information Steward logs.	340
	Metadata Management logs.	340
	Data Insight logs.	342
	Metadata Browsing Service and View Data Service logs.	343
	Viewing Information Steward logs.	344
15.2	Monitoring with CA Wily Introscope.	345
	CA Wily Introscope Prerequisites.	345
	Workflows.	346
16	Metadata Management export utility.	347
16.1	Metadata Management export utility.	347
16.2	Exporting objects to XML.	348

1 Accessing Information Steward for administrative tasks

Perform administrative tasks in the SAP BusinessObjects BI platform Central Management Console (CMC).

Ensure that you have administrator permission for Data Insight, Metadata Management, Data Review. This administrator permission is required to perform tasks in the CMC.

Access the CMC based on your system setup:

- If you installed the CMC on your computer, use the *Start* menu.
- If you installed the CMC on a different computer, enter the URL for the CMC into a browser window. Ensure that the URL contains the name of the computer and port number.

❁ Example

`http://<webserver>:8080/BOE/CMC` where `<webserver>` is the name of the web server machine.

If you changed the default virtual directory on the web server, type the URL accordingly. If necessary, change the default port number to the number you provided when you installed the SAP BusinessObjects BI platform.

1. Log in to the CMC by entering your user name and password.
2. Access various areas of the CMC to administer Information Steward.
 - To open the Information Steward home page, select *Information Steward* from the *CMC Home* dropdown list.
 - To perform tasks such as configure Information Steward settings or manage utilities, stay on the *CMC Home* page and click *Applications* under the *Manage* list. Then right-click *Information Steward Application* and select a task from the context menu.
3. To close the CMC, click *Log Off* in the upper right corner of the *Central Management Console* window.

Related Information

[Information Steward pre-defined users and groups \[page 52\]](#)

2 Architecture

2.1 Information Steward architecture overview

SAP Information Steward uses SAP BusinessObjects Business Intelligence (BI) platform for managing user security, scheduling integrator sources as tasks and utilities, and managing sources and on-demand services.

Before we get into the specifics of administering the various areas of Information Steward, it is important that you understand the overall architecture, system, and service components of the SAP Information Steward platform. Use the architecture information to form a plan for system deployment, management, and maintenance.

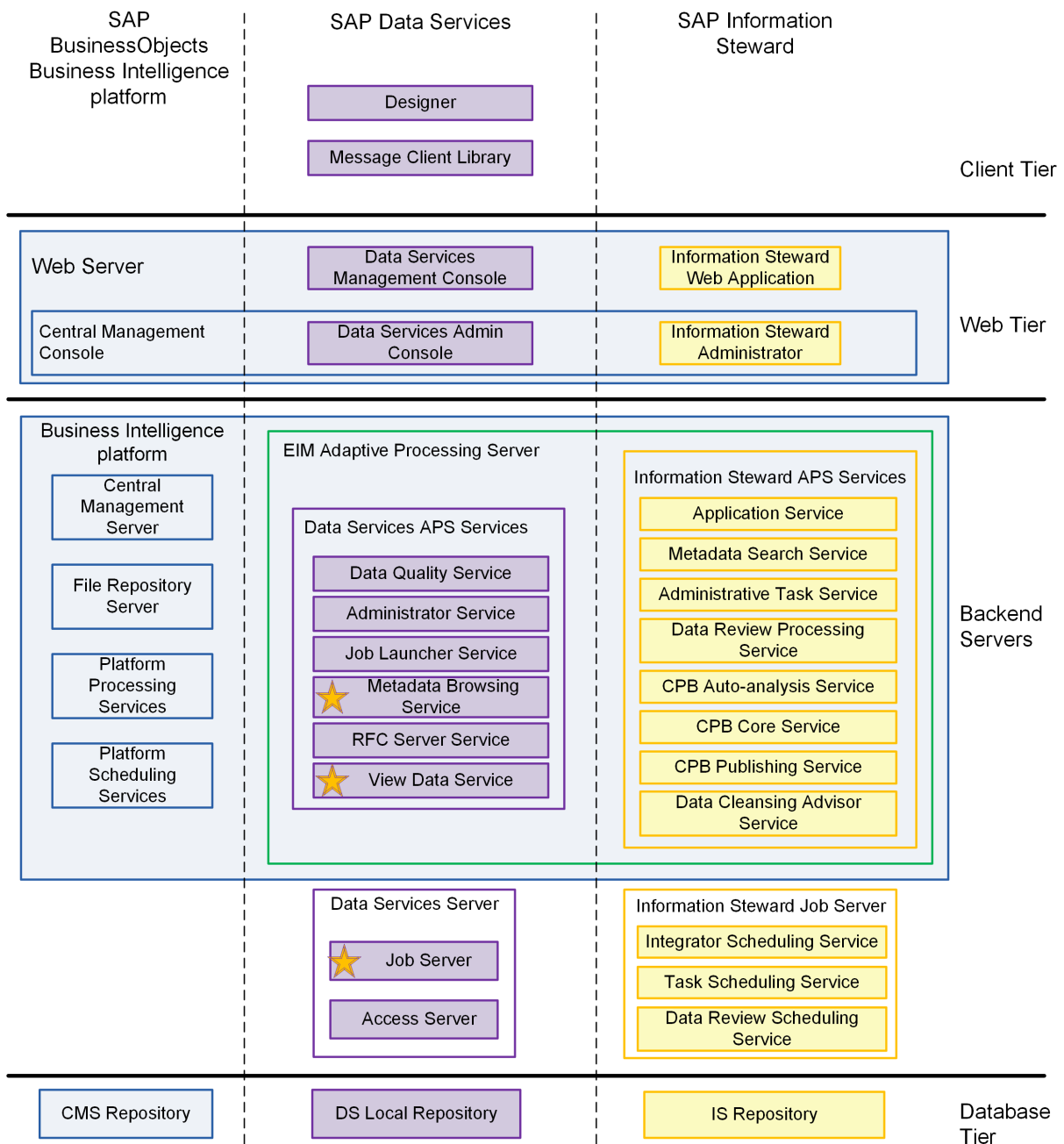
Information Steward uses SAP Data Services for profiling, rule tasks, browsing metadata, and viewing data.

This architecture allows deployment of server components from the following applications:

- SAP BusinessObjects BI platform
- SAP Data Services
- SAP Information Steward

The server components are located on different machines for flexibility, reliability, scalability, and better performance.

The following diagram shows the servers and services for SAP BusinessObjects BI platform, SAP Data Services, and SAP Information Steward. The services marked with a yellow star are used by the Data Insight module of Information Steward.



[Servers and services \[page 12\]](#)

We use the terms *server* and *service* throughout the SAP Information Steward documentation so it is important to understand what these terms mean.

[Services \[page 12\]](#)

SAP Information Steward uses specific services for the special processes it performs.

[Web application server \[page 16\]](#)

SAP Information Steward is deployed on a Web server using the Central Management Console (CMC) and a Web application server.

[SAP BusinessObjects BI platform \[page 18\]](#)

SAP Information Steward relies on SAP BusinessObjects Business Intelligence (BI) platform, or the SAP Business Objects Information Platform Services (IPS) for key functionality.

[SAP BI platform components usage \[page 19\]](#)

SAP Information Steward uses the various components of the BI platform or IPS, such as scheduling and Web Intelligence.

[SAP Data Services components usage \[page 22\]](#)

SAP Information Steward uses the services and servers from Data Services such as system management, data management, and data viewing.

[Metadata integrators \[page 23\]](#)

SAP Information Steward provides metadata integrators that access metadata from various source systems.

2.1.1 Servers and services

We use the terms *server* and *service* throughout the SAP Information Steward documentation so it is important to understand what these terms mean.

Server is used to describe an operating system level process (also known as a daemon) hosting one or more services. For example, the SAP Enterprise Information Management (EIM) Adaptive Processing server and SAP Information Steward Job Server host multiple services. A server runs under a specific operating system account and has its own process ID (PID).

A **service** is a server subsystem that performs a specific function. The service runs within the memory space of its server under the PID of the parent container (server). For example, the SAP Data Services Metadata Browsing service is a subsystem that runs within the SAP EIM Adaptive Processing Server.

Parent topic: [Information Steward architecture overview \[page 10\]](#)

Related Information

[Services \[page 12\]](#)

[Web application server \[page 16\]](#)

[SAP BusinessObjects BI platform \[page 18\]](#)

[SAP BI platform components usage \[page 19\]](#)

[SAP Data Services components usage \[page 22\]](#)

[Metadata integrators \[page 23\]](#)

2.1.2 Services

SAP Information Steward uses specific services for the special processes it performs.

The following table describes each of the services that Information Steward uses.

Services for Information Steward

Service	Server on which service runs	Service description	Deployment comments
Cleansing Package Builder Auto-analysis Service	Enterprise Information Management (EIM) Adaptive Processing Server	<p>Performs data analysis when you create custom cleansing packages in Cleansing Package Builder.</p> <ul style="list-style-type: none"> Data analysis uses information provided in the custom cleansing package wizard. Creates abstract version of records using data analysis and statistical analysis information. Groups abstracted records and processes with data inference algorithms. Creates suggestions shown in Design mode of Cleansing Package Builder. 	<p>Must already have installed on this computer:</p> <ul style="list-style-type: none"> SAP Business Intelligence (BI) platform Platform Processing Services
Cleansing Package Builder Core Service	EIM Adaptive Processing Server	<p>Performs main functionality of Cleansing Package Builder, such as:</p> <ul style="list-style-type: none"> Create and open cleansing packages Design mode Advanced mode 	<p>Must already have installed on this computer:</p> <ul style="list-style-type: none"> SAP BI platform Platform Processing Services
Cleansing Package Builder Publishing Service	EIM Adaptive Processing Server	<p>Converts published cleansing packages to reference data format used by Data Services.</p> <div data-bbox="805 1467 1093 1635" style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p>Note</p> <p>Can take a significant period of time for large cleansing packages.</p> </div>	<p>Must already have installed on this computer:</p> <ul style="list-style-type: none"> SAP BI platform Platform Processing Services
Information Steward Administrator Task Service	EIM Adaptive Processing Server	<p>Performs tasks on Data Services such as:</p> <ul style="list-style-type: none"> test rules deletes objects (connection, datastore, workflow) from Data Services exports rules to Data Services 	<p>Must already have installed on this computer:</p> <ul style="list-style-type: none"> SAP BI platform Platform Processing Services

Service	Server on which service runs	Service description	Deployment comments
Data Cleansing Advisor Service	EIM Adaptive Processing Server	Performs functions for setting up data cleansing solutions in Data Cleansing Advisor (in Data Insight).	Must already have installed on this computer: <ul style="list-style-type: none"> SAP BI platform Platform Processing Services
Data Service Metadata Browsing Service	EIM Adaptive Processing Server	Provides Data Insight users the capability to browse and import metadata from different data sources: <ul style="list-style-type: none"> Relational database systems such as <ul style="list-style-type: none"> Oracle Microsoft SQL Server Applications such as: <ul style="list-style-type: none"> SAP ECC Oracle applications 	Must already be installed on this computer: <ul style="list-style-type: none"> SAP Data Services SAP BI platform Platform Processing Services
Data Services View Data Service	EIM Adaptive Processing Server	Provides the capability to view the external data in Data Insight connections in Information Steward.	Must already be installed on this computer: <ul style="list-style-type: none"> SAP Data Services SAP BI platform Platform Processing Services
Information Steward Data Review Service	EIM Adaptive Processing Server	<ul style="list-style-type: none"> Checks the input table for new match groups that are ready for review. If match results are ready for review, creates a match review task. 	Must already be installed on this computer: <ul style="list-style-type: none"> SAP BI platform Platform Processing Services
Information Steward Data Review Task Scheduling Service	Information Steward Job Server	Processes scheduled Data Review tasks in the Central Management Console (CMC).	Must already be installed on this computer: <ul style="list-style-type: none"> SAP BI platform Platform Scheduling Services

Service	Server on which service runs	Service description	Deployment comments
Information Steward Application Service	EIM Adaptive Processing Server	<ul style="list-style-type: none"> Enables Information Steward web application access (read and write) to the Information Steward repository Provides metadata object relationships (such as data lineage and change impact analysis) for Metadata Management. 	<p>Must already be installed on this computer:</p> <ul style="list-style-type: none"> SAP BI platform Platform Processing Services <p>Recommended:</p> <ul style="list-style-type: none"> Install Application Service on different computer than web application server
Information Steward Metadata Search Service	EIM Adaptive Processing Server	<p>Finds an object that exists in any integrator source while viewing metadata on Information Steward Metadata Management using the Lucene Search Engine.</p> <ul style="list-style-type: none"> Constructs the search index during the execution of the Metadata Integrators File repository server stores compressed search index Updates search index with changes to Metadata: <ul style="list-style-type: none"> terms categories policy sets policy statements custom attributes annotations 	<p>If Metadata Search Service is not available during construction and update processes, search might return incorrect results.</p> <p>For these situations, Information Steward provides a utility to reconstruct the search index.</p>
Information Steward Task Scheduling Service	Information Steward Job Server	Processes scheduled Data Insight profile and rule tasks in the Central Management Console (CMC).	<p>Must have installed on this computer:</p> <ul style="list-style-type: none"> SAP BI platform Platform Scheduling Services

Note

You can deploy multiple Application Services for load balancing and availability.

Note

You can deploy multiple search services for load balancing and availability.

Service	Server on which service runs	Service description	Deployment comments
Information Steward Integrator Scheduling Service	Information Steward Job Server	Processes scheduled Metadata Management integrator sources in the CMC.	Must already be installed on this computer: <ul style="list-style-type: none"> SAP BI platform Platform Scheduling Services
Information Steward Integrator Service	EIM Adaptive Processing Server	Helper service used for testing connections to integrator sources.	Installed with the integrators.

Parent topic: [Information Steward architecture overview \[page 10\]](#)

Related Information

[Servers and services \[page 12\]](#)

[Web application server \[page 16\]](#)

[SAP BusinessObjects BI platform \[page 18\]](#)

[SAP BI platform components usage \[page 19\]](#)

[SAP Data Services components usage \[page 22\]](#)

[Metadata integrators \[page 23\]](#)

2.1.3 Web application server

SAP Information Steward is deployed on a Web server using the Central Management Console (CMC) and a Web application server.

Administrators perform administrative tasks on the CMC to set up the various Information Steward modules.

Administrative task	Description
Manage user security	Set up and maintain user security permissions for the following Information Steward modules: <ul style="list-style-type: none"> Data Insight Metadata Management Cleansing Package Builder Data Review
Manage connections, projects, and tasks	<ul style="list-style-type: none"> Manage Data Insight connections and projects. Manage Data Review connections and tasks.

Administrative task	Description
Configure and run metadata integrators	<ul style="list-style-type: none"> • Create and maintain metadata integrators and manage run schedules. • Schedule Data Insight profile and rule tasks. • Run or schedule Information Steward utilities.
Define source groups	Define source groups to subset the metadata when viewing relationships such as same as, impact, and lineage.
Define projects	Define Cleansing Package Builder projects.
Control behavior and performance	Configure application settings that affect the behavior and performance of Data Insight profile and rule tasks.

With a Web application server, data stewards analyze data and associated metadata in Information Steward modules. The following table describes the tasks for each module.

Module	Tasks to analyze data and associated metadata
Data Insight	<ul style="list-style-type: none"> • Profile data in tables or files and analyze the resulting profile attributes. • Define rules and set up data quality scorecards. • View data quality score trends and view sample data that failed each rule. • View data quality impact on dependent data sources. • Create data cleansing solutions based on your data content-type identification results and SAP best practices for your specific data.
Metadata Management	<ul style="list-style-type: none"> • View metadata from different sources and search for objects without the need to know the source or application in which it exists. • Add annotations to an object and define custom attributes and values for an object. • Run pre-defined reports that answer typical business questions such as which universe objects are excluded from reports, or which reports use a particular table or source. • View impact and lineage of objects within the same source or within different sources.
Metapedia	Define Metapedia terms related to business data and organize terms into categories. You can also define and implement policy sets and statements that are derived from business policies.
Cleansing Package Builder	<ul style="list-style-type: none"> • Define cleansing packages to parse and standardize data. • Publish cleansing packages. • Export cleansing packages so they can be imported to SAP Data Services to generate base Data Cleanse transforms for jobs that cleanse the data.

Module	Tasks to analyze data and associated metadata
Worklist for Data Reviewer	<ul style="list-style-type: none"> Review results of automated matching on a regular basis and make necessary corrections. Reassign the master record in a match group. Review Metapedia terms and policy sets and statements to approve or reject them. Review Data Insight rules for accuracy and approve or reject them.

Note

Ensure that you install the Information Steward web application on the same web application server as the SAP BusinessObjects BI platform.

For specific version compatibility, refer to the [PAM](#).

Parent topic: [Information Steward architecture overview \[page 10\]](#)

Related Information

[Servers and services \[page 12\]](#)

[Services \[page 12\]](#)

[SAP BusinessObjects BI platform \[page 18\]](#)

[SAP BI platform components usage \[page 19\]](#)

[SAP Data Services components usage \[page 22\]](#)

[Metadata integrators \[page 23\]](#)

2.1.4 SAP BusinessObjects BI platform

SAP Information Steward relies on SAP BusinessObjects Business Intelligence (BI) platform, or the SAP Business Objects Information Platform Services (IPS) for key functionality.

The BI platform or IPS provides the following functionality to Information Steward:

- Manage user and group security
- Schedule and run on-demand services for Metadata Management integrator sources and utilities
- Schedule and run on-demand services for Data Insight profiling and rule tasks and utilities
- Perform administrative tasks for Information Steward with the Central Management Console (CMC)
- Provide scalability by load balancing and high availability

Note

The decision to use either the BI Platform or IPS is made at the time you install Information Steward. For information about the differences between BI Platform and IPS, see the *Master Guide*.

Parent topic: [Information Steward architecture overview \[page 10\]](#)

Related Information

[Servers and services \[page 12\]](#)

[Services \[page 12\]](#)

[Web application server \[page 16\]](#)

[SAP BI platform components usage \[page 19\]](#)

[SAP Data Services components usage \[page 22\]](#)

[Metadata integrators \[page 23\]](#)

2.1.5 SAP BI platform components usage

SAP Information Steward uses the various components of the BI platform or IPS, such as scheduling and Web Intelligence.

Information Steward relies on the components in SAP BusinessObjects Business Intelligence (BI) platform or SAP BusinessObjects Information Platform Services (IPS) for important functionalities as described in the following table.

BI platform or IPS component	Usage for Information Steward
Web Tier	Deploys Information Steward on the BI platform Central Management Console (CMC) through which administrative tasks for Information Steward are performed.
Central Management Console (CMC)	Manage the following using the CMC: <ul style="list-style-type: none">• SAP EIM Adaptive Processing Server and services• Information Steward Job Server and services• Metadata Management module• Data Insight module• Cleansing Package Builder module• Match Review module connections and tasks• Information Steward utilities• User security (authentication and authorization)• Repository and application settings

BI platform or IPS component	Usage for Information Steward
Central Management Server (CMS)	<p>Maintains a database of information about your BI platform system including the following:</p> <ul style="list-style-type: none"> • Information about users and groups • Security levels • Schedule information • BI platform content and servers <p>Stores the following objects in the Metadata Management module of Information Steward:</p> <ul style="list-style-type: none"> • Integrator source configurations • Source groups • Utilities configurations • Data Insight connections • Projects • Tasks
	<div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 10px;"> <p>Note</p> <p>Because integrator source configurations and source group definitions are stored in the CMS, you can use the upgrade management tool to move them from one version of the CMS to another. The schedules and rights information are considered dependencies of these configurations. For details, see the <i>SAP Information Steward Upgrade Guide</i>.</p> </div>
	<p>For more information about the CMS, see the <i>Business Intelligence Platform Administrator Guide</i>. If you installed Information Platform Service (IPS), see the <i>Information platform services Administrator Guide</i>.</p>
Information Platform Scheduling Services	<p>Executes profiling tasks and integrator tasks for the Information Steward Job Server, which may host the following services for Information Steward:</p> <ul style="list-style-type: none"> • Task scheduling service • Integrator scheduling service • Data Review scheduling service

BI platform or IPS component	Usage for Information Steward
Information Platform Processing Services	<p data-bbox="804 356 1385 450">Creates the SAP Enterprise Information Management (EIM) Adaptive Processing Server, which is required during Information Steward installation.</p> <p data-bbox="804 477 1385 539">The SAP EIM Adaptive Processing Server uses the platform processing services to host the following services:</p> <ul data-bbox="815 562 1321 1003" style="list-style-type: none"> <li data-bbox="815 562 1091 586">• Metadata Search service <li data-bbox="815 602 1123 627">• Metadata Integrator service <li data-bbox="815 642 1254 667">• Data Services Metadata Browsing service <li data-bbox="815 683 1161 707">• Data Services View Data service <li data-bbox="815 723 1315 748">• Information Steward Administrator Task service <li data-bbox="815 763 1235 788">• Cleansing Package Builder Core service <li data-bbox="815 804 1321 828">• Cleansing Package Builder auto-analysis service <li data-bbox="815 844 1294 869">• Cleansing Package Builder Publishing service <li data-bbox="815 884 1174 909">• Match Review Processing service <li data-bbox="815 925 1150 949">• Data Cleansing Advisor service <li data-bbox="815 965 1038 990">• Application service
Input File Repository Server and Output File Repository Server	<p data-bbox="804 1032 1206 1057">Stores files associated with the following:</p> <ul data-bbox="815 1077 1394 1238" style="list-style-type: none"> <li data-bbox="815 1077 1394 1133">• A published cleansing package, which Data Services accesses <li data-bbox="815 1144 1394 1200">• History logs for Data Insight task and metadata integrator execution <li data-bbox="815 1211 1289 1238">• Search index for metadata integrator objects

Parent topic: [Information Steward architecture overview \[page 10\]](#)

Related Information

[Servers and services \[page 12\]](#)

[Services \[page 12\]](#)

[Web application server \[page 16\]](#)

[SAP BusinessObjects BI platform \[page 18\]](#)

[SAP Data Services components usage \[page 22\]](#)

[Metadata integrators \[page 23\]](#)

2.1.6 SAP Data Services components usage

SAP Information Steward uses the services and servers from Data Services such as system management, data management, and data viewing.

Information Steward requires the following Data Services servers and services:

- Job Server
- Metadata Browsing Service
- View Data Service

Data Services Job Server

The Data Services Job Server provides the engine processes that perform the Data Insight data profiling and validation rule tasks. The engine processes use parallel execution and in-memory processing to deliver high data throughput and scalability. The Information Steward installer installs the Job Server on the primary computer.

The Data Services Job Server provides the system management tools that Information Steward requires for initial installation.

System management tool	Description
Repository Manager	<p>Creates the required Data Insight objects in the Information Steward repository.</p> <p>The Information Steward installer invokes the Repository Manager automatically when it creates the repository during the initial installation.</p>
Server Manager	<p>Creates the Information Steward Job Server group and job servers, and associates them to the Information Steward Repository.</p> <p>To add job servers to the Information Steward Job Server group, manually invoke the server manager.</p>

Data Services Metadata Browsing Service

The Metadata Browsing Service provides the capabilities to browse and import the metadata from Data Insight connections.

Data Services View Data Service

The View Data Service provides the capabilities to view the source data from Data Insight connections.

Note

When you install Data Services, ensure that you select the following components:

- Job Server under the Server category.
- SAP Solutions for EIM Adaptive Processing Server, which contains the Metadata Browsing Service and the View Data Service.

Parent topic: [Information Steward architecture overview \[page 10\]](#)

Related Information

[Servers and services \[page 12\]](#)

[Services \[page 12\]](#)

[Web application server \[page 16\]](#)

[SAP BusinessObjects BI platform \[page 18\]](#)

[SAP BI platform components usage \[page 19\]](#)

[Metadata integrators \[page 23\]](#)

2.1.7 Metadata integrators

SAP Information Steward provides metadata integrators that access metadata from various source systems.

Metadata integrators are programs that perform the following processes:

- Collect metadata from source systems and store the collected metadata into the Information Steward repository.
- Run at regular intervals based on a schedule.
- Update existing metadata.
- Run on one job server or multiple job servers for load balancing and high availability.
- Run as a separate process.

The following table describes the metadata integrators included with Information Steward, and the type of metadata that each integrator collects.

Information Steward metadata integrators

Metadata integrator name	Metadata collected
SAP BusinessObjects Enterprise Metadata Integrator	Collects metadata about objects such as universes, Crystal Reports, Web Intelligence documents, and Desktop Intelligence documents.
SAP NetWeaver Business Warehouse Metadata Integrator	Collects metadata about objects such as Queries, InfoProviders, InfoObjects, Transformations, and DataSources from an SAP NetWeaver Business Warehouse system.

Metadata integrator name	Metadata collected
Common Warehouse Model (CWM) Metadata Integrator	Collects metadata about objects such as catalogs, schemas, and tables from the relational package of CWM.
SAP Data Federator Metadata Integrator	Collects metadata about objects such as projects, catalogs, data sources, and mapping rules from a Data Federator repository.
SAP Data Services Metadata Integrator	Collects metadata about objects such as source tables, target tables, and column mappings from a Data Services repository.
Meta Integration Model Bridge (MIMB) Metadata Integrator (also known as MITI Integrator)	<p>Collects metadata from other third-party sources such as the following:</p> <ul style="list-style-type: none"> • Data Modeling metadata such as SAP Sybase PowerDesigner (Version 16.0 and older), Embarcadero ER/Studio, and Oracle Designer • Extract, Transform, and Load (ETL) metadata such as Oracle Warehouse Builder and Microsoft SQL Server Integration Services (SSIS) • OLAP and BI metadata such as IBM DB2 Cube Views, Oracle OLAP, and Cognos BI Reporting
Relational databases Metadata Integrator	Collects metadata from relational database management systems (RDBMS) which can be DB2, MySQL, Oracle, SQL Server, Java Database Connectivity (JDBC), or an SAP Universe connection. Collected metadata includes the definition of objects such as tables, view, synonyms, and aliases.
SAP HANA Metadata Integrator	Collects metadata about objects such as tables, views, and procedures from an SAP HANA database.
SAP Sybase PowerDesigner Metadata Integrator	Collects physical data model metadata for SAP Metadata Management, and collects glossary information for SAP Metapedia from the SAP Sybase PowerDesigner repository using PowerDesigner Version 16.1 and newer.
<div style="background-color: #f0f0f0; padding: 10px; border: 1px solid #ccc;"> <p>Note</p> <p>The SAP Sybase PowerDesigner application must be installed on the same machine as SAP Information Steward Metadata Management.</p> </div>	
Excel Metadata Integrator	Collects metadata from database models that aren't supported by other metadata integrators. For example, collect metadata from dBase, Microsoft Access, Microsoft FoxPro, Hadoop, SQLite, and so on. In addition, use the Excel Metadata Integrator to create your own metadata without using a physical database source. For example, use Unified Modeling Language to create metadata for a relational database design.

For information about each metadata integrator, see the *User Guide*.

Parent topic: [Information Steward architecture overview \[page 10\]](#)

Related Information

[Servers and services \[page 12\]](#)

[Services \[page 12\]](#)

[Web application server \[page 16\]](#)

[SAP BusinessObjects BI platform \[page 18\]](#)

[SAP BI platform components usage \[page 19\]](#)

[SAP Data Services components usage \[page 22\]](#)

2.2 Distributed architecture

With SAP Information Steward, use the distributed architecture to improve performance.

Use distributed architecture to distribute software components across multiple computers. Consider your environment for the best distribution to support the traffic and connectivity requirements of your network.

- Create a system that is minimally distributed for a development and testing environment
- Create a system that is highly distributed for scaling with the demands of a production environment

[Component communication channels \[page 25\]](#)

Use specific communication channels to plan your network configuration and firewall settings.

[Port assignments \[page 26\]](#)

On each host system, verify that all ports to be used by SAP Information Steward components are available and not in use by other programs.

[DSN-less and TNS-less connections \[page 27\]](#)

DSN-less and TNS-less (Oracle) connections are also known as server name connections.

2.2.1 Component communication channels

Use specific communication channels to plan your network configuration and firewall settings.

The following table describes each communication channel type and components.

Communication channel	Components
CORBA	<p>All SAP BusinessObjects Business Intelligence (BI) solutions communicate with the Central Management Server (CMS) and each other on CORBA, Common Object Request Broker Architecture. CORBA is designed to manage communication of systems on diverse platforms.</p> <p>Any communication with SAP BusinessObjects BI solutions occurs through CORBA.</p> <p>CORBA communication can happen with or without SSL.</p>
HTTP/HTTPS	<p>SAP Information Steward communicates with the Central Management Console (CMC) and SAP Data Services Management Console using HTTP or HTTPS.</p> <p>SAP Data Federator metadata integrator uses HTTP/HTTPS to communicate with the Data Federator server.</p>
TCP/IP	<p>Information Steward is a web-based application. Information Steward and the subcomponents deployed on SAP BusinessObjects BI platform and Data Services job server or Data Services engine uses TCP/IP with or without SSL.</p> <p>Remote SAP BusinessObjects Enterprise Metadata Integrator, deployed on SAP BusinessObjects Enterprise XI 3.1 system, communicates with SAP BusinessObjects Enterprise Metadata Integrator, deployed on SAP BusinessObjects BI platform 4.1, using TCP/IP with or without SSL.</p>

Parent topic: [Distributed architecture \[page 25\]](#)

Related Information

[Port assignments \[page 26\]](#)

[DSN-less and TNS-less connections \[page 27\]](#)

2.2.2 Port assignments

On each host system, verify that all ports to be used by SAP Information Steward components are available and not in use by other programs.

For a development system, you can install many components on the same host. Installing on a single host simplifies many connections between components (the host name is always the same), but you must still define connections based on the TCP/IP protocol.

This table details the default ports used by Information Steward components:

Default Port	Description
5005	Remote Job Server port that receives commands to run an integrator on a previous version.
8080	Default web application server port that listens to HTTP requests from web browsers.
6576	Default port used by the Data Cleansing Advisor repository. For information about editing this port number, see "Viewing and editing Data Cleansing Advisor repository information" in the <i>Information Steward Administrator's Guide</i> .

Connections to databases also require access to a port that is defined when the database is set up. For details, see the *SAP BusinessObjects Business Intelligence platform Administrator Guide*.

Parent topic: [Distributed architecture \[page 25\]](#)

Related Information

[Component communication channels \[page 25\]](#)

[DSN-less and TNS-less connections \[page 27\]](#)

2.2.3 DSN-less and TNS-less connections

DSN-less and TNS-less (Oracle) connections are also known as server name connections.

SAP Information Steward provides server name connections to databases that you use as an Information Steward repository, profiling source, or storage for data that failed validation rules. Server name connections eliminate the need to configure the same DSN (TNS for Oracle) entries on every machine in a distributed environment.

The Information Steward repository supports server name connections for the databases described in the following table.

Note

For supported versions and the most updated list of supported databases for the repository, see the PAM.

Repository database required information

Database	Information required
Microsoft SQL Server	<ul style="list-style-type: none"> • CMS registration name • Database name • Server name • Port number (default is 1433)

Database	Information required
SAP HANA	<ul style="list-style-type: none"> • CMS registration name • Server name • Port number (default is 30015)
SAP ASE	<ul style="list-style-type: none"> • CMS registration name • Database name • Server name • Port number
IBM for DB2	<ul style="list-style-type: none"> • CMS registration name • Database name • Server name: DB2 database alias • Port number
Oracle	<ul style="list-style-type: none"> • CMS registration name • Server: TNSNAMES connect identifier • Port number • SID (instead of TNS name)
SAP SQL Anywhere	<p>Database server that is bundled with the BI or IPS installation.</p> <ul style="list-style-type: none"> • CMS registration name • Database name • Server name • Port number

For Data Insight profiling sources and failed data storage, the following database types are supported for server name connections:

- DB2 UDB
- Informix
- MySQL
- Netezza
- Oracle
- SAP HANA
- SQL Anywhere
- Sybase IQ
- Teradata

Note

For the most current list of supported databases for server name connections, see the *Release Notes*.

Parent topic: [Distributed architecture \[page 25\]](#)

Related Information

[Component communication channels \[page 25\]](#)

[Port assignments \[page 26\]](#)

2.3 SAP integration

SAP Information Steward integrates with your existing SAP infrastructure with the following SAP tools:

- **SAP System Landscape Directory (SLD)**
The system landscape directory of SAP NetWeaver is the central source of system landscape information relevant for the management of your software life-cycle. By providing a directory comprising information about all installable software available from SAP and automatically updated data about systems already installed in a landscape, you get the foundation for tool support to plan software life-cycle tasks in your system landscape.
- **Monitoring with CA Wily Introscope**
CA Wily Introscope is a web application management product that delivers the ability to monitor and diagnose performance problems that may occur within Java-based SAP modules in production, including visibility into custom Java applications and connections to back-end systems. It allows you to isolate performance bottlenecks in NetWeaver modules including individual Servlets, JSPs, EJBs, JCOs, Classes, Methods and more. It offers real-time, low-overhead monitoring, end-to-end transaction visibility, historical data for analysis or capacity planning, customizable dashboards, automated threshold alarms, and an open architecture to extend monitoring beyond NetWeaver environments.

Related Information

[CA Wily Introscope Prerequisites \[page 345\]](#)

2.4 Information workflows

When tasks are performed in SAP Information Steward, such as adding a table to a Data Insight project, running a Metadata Management integrator source, or creating a cleansing package, information flows through SAP information platform services, SAP Data Services, and SAP Information Steward. The servers and services within each of these software products communicate with each other to accomplish a task.

The following section describes some of the process flows as they would happen in SAP information platform services, SAP Data Services, and SAP Information Steward.

[Adding a table to a Data Insight project \[page 30\]](#)

After you've created a Data Insight project, and created connections in the Central Management Console, add tables to Data Insight.

[Profiling data \[page 31\]](#)

Perform profiling on data in Data Insight.

[Scheduling and running a Metadata Management integrator source \[page 32\]](#)

Schedule and run an integrator source in Metadata Management to collect metadata.

[Creating a custom Cleansing Package with Cleansing Package Builder \[page 32\]](#)

Create and publish a custom cleansing package using the SAP Information Steward Cleansing Package Builder.

[Creating a Data Cleansing Solution in Data Cleansing Advisor \[page 34\]](#)

SAP Information Steward helps you create a complete data cleansing solution that you can ultimately use to increase the quality of your data in SAP Data Services.

Related Information

[Information Steward architecture overview \[page 10\]](#)

[Services \[page 12\]](#)

2.4.1 Adding a table to a Data Insight project

After you've created a Data Insight project, and created connections in the Central Management Console, add tables to Data Insight.

The following workflow describes the process of adding a table to a Data Insight project.

1. Open the applicable project in Data Insight.
2. Select **► Add ► Tables ►** to open the *Browse Metadata* window.
3. The Web application server passes the request to the Central Management Server (CMS) and returns a list of connections that you have permission to view.
4. The CMS sends the request to the SAP Data Services Metadata Browsing Service.
5. The Metadata Browsing Service obtains the metadata from the connection and sends the metadata to the Web application server.
6. The Web application server displays the metadata in the Data Insight *Browse Metadata* window.
7. When you choose a table and select *Add to Project*, the Web application asks the Application Service to store the metadata in the Information Steward repository.

Task overview: [Information workflows \[page 29\]](#)

Related Information

[Profiling data \[page 31\]](#)

[Scheduling and running a Metadata Management integrator source \[page 32\]](#)

[Creating a custom Cleansing Package with Cleansing Package Builder \[page 32\]](#)

[Creating a Data Cleansing Solution in Data Cleansing Advisor \[page 34\]](#)

2.4.2 Profiling data

Perform profiling on data in Data Insight.

The following workflow describes the process of running a profile task in Data Insight.

Note

Running validation rules are similar to the following steps.

1. Select a table name or file in Workspace Home and select *Profile*.
2. Choose the table or tables to profile.
3. Save the task and schedules to run it.
4. The Web application server passes the request to the Central Management Server (CMS).
5. The Information Steward Web application determines from the CMS system if you have the right to run profile tasks on the connection that contains the table or file.
6. The Administrator determines if you have the right to create a profile task for the connection, and have the right to schedule the task. If so, the task is scheduled in the CMS system.
7. When the scheduled time arrives, the CMS sends the task information to the Information Steward Task Scheduling Service.
8. The Information Steward Task Scheduling Service sends the profile task to the SAP Data Services Job Server.
9. The Data Services Job Server partitions the profile task based on the performance application settings.
10. The Data Services Job Server runs the profile task and stores the results in the Information Steward repository.
11. The web application server displays the profile results in Data Insight Workspace Home.

Task overview: [Information workflows \[page 29\]](#)

Related Information

[Adding a table to a Data Insight project \[page 30\]](#)

[Scheduling and running a Metadata Management integrator source \[page 32\]](#)

[Creating a custom Cleansing Package with Cleansing Package Builder \[page 32\]](#)

[Creating a Data Cleansing Solution in Data Cleansing Advisor \[page 34\]](#)

2.4.3 Scheduling and running a Metadata Management integrator source

Schedule and run an integrator source in Metadata Management to collect metadata.

The following workflow describes the process:

1. You schedule an integrator source in the Central Management Console (CMC). The CMC sends the request to the Central Management Server (CMS) system.
2. If you have the appropriate rights to schedule the object, the CMS commits the scheduled integrator request to the CMS system.
3. When the scheduled time arrives, the CMS finds a suitable SAP Information Steward Job Server based on the associated integrator Job Server group and passes the job.
4. The integrator process collects metadata and stores the metadata in the Information Steward repository.
5. The integrator process generates the Metadata Management search index files and loads them to the EIM Adaptive Processing Server Input File Repository Server.
6. After uploading the search index files, the integrator source notifies the Metadata Management search service.
7. The Metadata Management search service downloads the generated index files and consolidates them into a master index file.
8. The Information Steward Integrator Scheduling Service updates the CMS with the job status.

Task overview: [Information workflows \[page 29\]](#)

Related Information

[Adding a table to a Data Insight project \[page 30\]](#)

[Profiling data \[page 31\]](#)

[Creating a custom Cleansing Package with Cleansing Package Builder \[page 32\]](#)

[Creating a Data Cleansing Solution in Data Cleansing Advisor \[page 34\]](#)

2.4.4 Creating a custom Cleansing Package with Cleansing Package Builder

Create and publish a custom cleansing package using the SAP Information Steward Cleansing Package Builder.

The following workflow describes the process of creating and publishing a custom cleansing package in Cleansing Package Builder.

1. Open the *Cleansing Package Builder* tab.
2. The Cleansing Package Builder application sends your login information to the Cleansing Package Builder Web Service.

3. The Cleansing Package Builder Web Service sends the information to the EIM Adaptive Processing server. The EIM Adaptive Processing Server runs on the SAP BusinessObjects Business Intelligence platform.
4. The EIM Adaptive Processing Server determines which rights the user has in Cleansing Package Builder.
5. The EIM Adaptive Processing Server sends the information back through the Cleansing Package Builder Web Service to the Cleansing Package Builder application. You can view only the cleansing packages to which you have the rights.
6. To start creating a cleansing package, open the *Cleansing Packages Tasks* window and select ► *New Cleansing Package* ► *Custom Cleansing Package*. ►
You provide the required information and sample data to create the cleansing package.
7. The Cleansing Package Builder sends the information through the Cleansing Package Builder Web service to the Cleansing Package Builder Core Service, using the SAP Enterprise SDK mechanism. The Cleansing Package Builder Core Service handles the main functions of Cleansing Package Builder. The Cleansing Package Builder Core Service runs on the EIM Adaptive Processing Server.
8. The Cleansing Package Builder Core Service sends the response back through the Cleansing Package Builder Web Service to the Cleansing Package Builder application. The new cleansing package is created in Cleansing Package Builder.
9. Information Steward communicates with the Cleansing Package Builder Auto-Analysis Service through the Cleansing Package Builder Web Service. The Cleansing Package Builder Auto-Analysis Service analyzes the data to create suggestions of standard forms and variations. The Cleansing Package Builder Auto-Analysis Service runs on the EIM Adaptive Processing Server.
10. When you've finished refining the cleansing package, select *Publish* in the *Cleansing Packages Tasks* window.
11. The Cleansing Package Builder application communicates with the Cleansing Package Builder Publishing Service through the Cleansing Package Builder Web Service. The Cleansing Package Builder Publishing Service assists in the cleansing package's conversion to the reference data format used by SAP Data Services. The Cleansing Package Builder Publishing Service runs on the EIM Adaptive Processing Server.
12. Information Steward sends the published cleansing package to the SAP BusinessObjects BI platform input file repository so that it can be accessed by Data Services.
13. Data Services communicates directly with the SAP BusinessObjects BI platform to sync with the published cleansing packages.

Task overview: [Information workflows \[page 29\]](#)

Related Information

[Adding a table to a Data Insight project \[page 30\]](#)

[Profiling data \[page 31\]](#)

[Scheduling and running a Metadata Management integrator source \[page 32\]](#)

[Creating a Data Cleansing Solution in Data Cleansing Advisor \[page 34\]](#)

2.4.5 Creating a Data Cleansing Solution in Data Cleansing Advisor

SAP Information Steward helps you create a complete data cleansing solution that you can ultimately use to increase the quality of your data in SAP Data Services.

The following workflow describes the process of creating and publishing a data cleansing solution with the Data Cleansing Advisor:

1. You select an input source in the Data Insight Workspace Home, and then choose ► *Profile* ► *Content Type* ▾.
2. After content type profiling is complete, Data Insight places the *Launch Data Cleansing Advisor* icon next to the name of the input source, if applicable.

The icon appears only when Data Insight is able to assign content types for one or more columns in the source.

3. Select the *Launch Data Cleansing Advisor* icon to start the Data Cleansing Advisor wizard.
4. The Data Cleansing Advisor wizard prompts you to provide necessary information for creating a data cleansing solution.
5. For each wizard step, the Web application server communicates with the Data Cleansing Advisor service to retrieve optimal results based on your current selection.
6. When you select *Finish* in the Data Cleansing Advisor wizard, the Web application server passes the request to the Data Cleansing Advisor service.
7. Data Insight starts the creation process for this data cleansing solution on the EIM Adaptive Processing Server.
8. Data Insight redirects you to the Data Cleansing Solutions list in Data Cleansing Advisor. Data Insight lists the newly created cleansing solution with a status of "Creating."
9. The Data Cleansing Advisor service stores all settings in the Central Management Server so that it can retrieve the settings for processing in later steps.
10. The Data Cleansing Advisor service copies data from the input source to the Information Steward local repository. The service then starts a new SAP Data Services job on the Data Services Job Server.
11. The Data Cleansing Advisor service cleanses the data using content-type profiling results and options selected in the Data Cleansing Advisor wizard.

The Data Cleansing Advisor services starts two Data Services jobs on the Data Services Job Server for this step.

12. Data Cleansing Advisor performs matching by starting a new Data Services job on the Data Services Job Server.
13. Data Insight lists the data cleansing solution with a "Ready" status. You can view the results and edit cleanse and match settings in Data Cleansing Advisor.
14. When you satisfied with the data cleansing solution, you can do one of two things: or you can
 - Export the solution as an ATL file, which you can later import into the Data Services Designer.
 - Send the solution directly to the Data Services repository, which you can open as a data flow in the Data Services workspace for completion.

ⓘ Note

For the job to be viable, you must add a source and target object to the data flow and complete the remaining mapping.

Task overview: [Information workflows \[page 29\]](#)

Related Information

[Adding a table to a Data Insight project \[page 30\]](#)

[Profiling data \[page 31\]](#)

[Scheduling and running a Metadata Management integrator source \[page 32\]](#)

[Creating a custom Cleansing Package with Cleansing Package Builder \[page 32\]](#)

3 Securing SAP Information Steward

3.1 Security Overview

SAP Information Steward uses the security framework that SAP BusinessObjects Business Intelligence platform (BI platform) provides.

The SAP BusinessObjects BI platform architecture addresses the many security concerns that affect today's businesses and organizations. It supports features such as distributed security, single sign-on, resource access security, granular object rights, and third-party authentication in order to protect against unauthorized access.

For details about how BI platform addresses enterprise security concerns, see the *SAP Business Intelligence Platform Administrator's Guide*. For more information about securing the platform, see the topic "Securing the BI platform" in the same guide (*SAP Business Intelligence Platform Administrator's Guide*).

Running the Information Steward application

As a security requirement, the latest Adobe Flash player must be installed in the web browser used for Information Steward. JavaScript is also used, but you do not need to install it.

Related Information

[Securing user data for Information Steward \[page 38\]](#)

[Storage of sensitive information \[page 38\]](#)

[Configuring Information Steward for SSL \[page 39\]](#)

[Reverse proxy servers \[page 43\]](#)

[Information Steward pre-defined users and groups \[page 52\]](#)

3.2 Enterprise security

SAP Information Steward is a web-based application that uses enterprise security provided by SAP Business Intelligence platform. Information Steward takes advantage of the following SAP Enterprise security features:

- Secure access to user data
- Storage of sensitive information
- Secure connections

- Reverse proxy servers

Note

You should have a virus scanner running on the Business Intelligence platform server to help avoid any downtime on server systems due to corrupt files. The virus scanner can quarantine any malicious files uploaded to FRS, for example, Cleansing Package Builder sample data or Metadata Management Integrator configuration files.

Related Information

[Securing user data for Information Steward \[page 38\]](#)

[Storage of sensitive information \[page 38\]](#)

[Configuring Information Steward for SSL \[page 39\]](#)

[Reverse proxy servers \[page 43\]](#)

3.2.1 Data protection and privacy

Data protection is associated with numerous legal requirements and privacy concerns. In addition to compliance with applicable data privacy regulations, it is necessary to consider compliance with industry-specific legislation in different countries. SAP provides specific features and functions to support compliance with regards to relevant legal requirements, including data protection. SAP does not give any advice on whether these features and functions are the best method to support company, industry, regional, or country-specific requirements. Furthermore, this information does not give any advice or recommendation in regards to additional features that would be required in particular IT environments; decisions related to data protection must be made on a case-by-case basis, under consideration of the given system landscape and the applicable legal requirements.

Note

In the majority of cases, compliance with applicable data protection and privacy laws will not be covered by a product feature. SAP software supports data protection compliance by providing security features and specific data protection-relevant functions, such as simplified blocking and deletion of personal data. SAP does not provide legal advice in any form. Definitions and other terms used in this document are not taken from any given legal source.

Audit data and log files are handled by the SAP BusinessObjects Business Intelligence Platform (BIP). For more information about managing the data and log files, read about access logging and deleting personal data in the *Business Intelligence Platform Administrator Guide* version 4.2 SP5 or higher.

3.2.2 Securing user data for Information Steward

SAP Information Steward has access to the following data which might contain sensitive information:

- Source data in Data Insight connections on which users run profile and rule tasks
- Sample data from profiling results that Data Insight stores in the Information Steward repository
- Sample data that failed validation rules that Data Insight stores in the Information Steward repository
- All data that failed validation rules that a user chooses to store in a database accessed through a Data Insight connection.
- Results data in Match Review is saved in the staging repository. The user must secure the staging repository.

The Database Administrator (DBA) secures the data in these databases by managing user permissions on them:

- Source systems used by Data Insight
- Repository used by Information Steward
- Failed data repository used by Data Insight
- Staging repository used by Match Review

The following information is stored in the Information Steward repository during these processes:

- Data Insight information and profile sample data is stored when profiling
- Data Insight sample failed data is stored when rule tasks are calculated
- Match Review source identification is stored when a review task is running

In addition, the Data Insight Administrator or Administrator control access to the data by using the Central Management System (CMS) to manage the following rights on the Data Insight connections:

- View Data
- Profile/Rule permission
- View Sample Data
- Export Data

Related Information

[User rights in Data Insight \[page 58\]](#)

3.2.3 Storage of sensitive information

Information Steward uses the SAP Enterprise cryptography which is designed to protect sensitive data stored in the CMS repository. Sensitive data includes user credentials, data source connectivity data, and any other info objects that store passwords. This data is encrypted to ensure privacy, keep it free from corruption, and maintain access control. Sample data stored in the Information Steward repository is protected by the user access permissions. For more information, see the "Overview of SAP BusinessObjects Enterprise data security" section of the *SAP BusinessObjects Enterprise Administrator's Guide*.

Encryption of sensitive information, such as passwords, is done in the following Information Steward areas:

- Information Steward repositories
- Metadata Integrator sources
- Connections to data sources

Storing cookies

A cookie is a small text file that stores session state on the client side: the user's web browser caches the cookie for later use. The following cookies are used by Information Steward: MM_Logon, user, cms, authentication, sap id, sap client number, token, vintela sso, metadata login.

3.2.4 Configuring Information Steward for SSL

You can configure the Secure Sockets Layer (SSL) protocol for SAP Information Steward by:

- configuring the web application (see the Business Intelligence platform *Web Application Deployment Guide*)

Note

When configuring the web application, be certain that the HTTP access to the Business Intelligence platform web application is maintained after enabling SSL. Otherwise, the main Information Steward web application is not able to initialize.

- configuring the Business Intelligence platform (see the Business Intelligence platform *Administrator Guide*)
- configuring the Remote Job Server (below)

Configure the Remote Job server

The SAP Enterprise Metadata Integrator in Information Steward can collect metadata from an SAP BusinessObjects Enterprise XI 3.x system by using the Remote Job Server. When you install Information Steward, you install the Remote Job Server component on the computer where the Enterprise XI 3.x system resides. Then you use the Information Steward Service Configuration to configure the Remote Job Server. For information about installing the Remote Job Server on the SAP BusinessObjects Enterprise XI 3.x system, see "Remote Job Server Installation" in the *Installation Guide*.

If you are using the Secure Sockets Layer (SSL) protocol for all network communication between clients and servers in your SAP BusinessObjects Enterprise XI 3.x and SAP Business intelligence platform 4.0 deployments, you can use SSL for the network communication between the Remote Job Server and the Metadata Integrator on Information Steward. In this environment:

- The server is the Remote Job Server on the SAP BusinessObjects Enterprise XI 3.x. To enable SSL, the server must have both keystore and truststore files defined.

- The client is the Metadata Integrator on Information Steward. To enable SSL, the client must use the same truststore and password as the server.

To set up SSL between the Remote Job Server and the Metadata Integrator, you need to perform the following tasks:

- Create keystore and truststore files for the Remote Job Server and copy the truststore file to the Business intelligence platform 4.0 system.
- Configure the location of SAP BusinessObjects Enterprise XI 3.x SSL certificates and key file names (from Server Intelligence Agent (SIA)).

Related Information

[Creating the keystore and truststore files for the Remote Job Server \[page 40\]](#)

[Configuring the SSL protocol for the Remote Job Server \[page 41\]](#)

[SSL and HTTPS setup for Information Steward \[page 42\]](#)

3.2.4.1 Creating the keystore and truststore files for the Remote Job Server

To set up SSL protocol for communication to the Remote Job Server, use the `keytool` command to:

- Create a certificate and store it in a keystore file on the computer where SAP BusinessObjects Enterprise XI 3.x resides
 - Create a trust certificate and store it in a truststore file on the computer where SAP BusinessObjects Enterprise XI 3.x resides
 - Copy the trust certificate into a truststore file on the computer where SAP BusinessObjects Business Intelligence Platform 4.0. resides
1. On the computer where you installed SAP BusinessObjects Enterprise XI 3.x, generate a keystore file and export it:
 - a. Open a cmd window and go to the directory where Metadata Management configuration files are stored for Information Steward.

For example, type the following command:

```
cd C:\Program Files (x86)\SAP BusinessObjects\InformationSteward\MM\config
```

- b. Generate the keystore file using the `keytool` command.
For example, type the following command to generate a keystore file with the name `is.keystore` and password `mypwstore`:

```
%JAVA_HOME%\bin\keytool -genkey -keystore is.keystore
-keyalg RSA -dname "CN=localhost,OU=ICC,O=ICC,L=PA,ST=CA,C=US"
-keypass mypwkey -storepass mypwstore
```

The value for `keypass` protects the generated key.

This command creates the key and certificate in the keystore file in the `C:\Program Files (x86)\SAP BusinessObjects\InformationSteward\MM\config` directory.

- c. Export the certificate from the keystore file using the `keytool` command.
For example, type the following command to export to the file `isClient.cer`:

```
%JAVA_HOME%\bin\keytool -export -keystore is.keystore  
-file isClient.cer -keypass mypwkey -storepass mypwstore
```

- d. Import the certificate from the export file to create the truststore file.
For example, type the following command to import the certificate into a truststore file named `is.truststore.keystore`:

```
%JAVA_HOME%\bin\keytool -import -v -trustcacerts  
-file isClient.cer -keystore is.truststore.keystore  
-keypass mypwkey -storepass mypwstore
```

This command stores the certificate in the truststore file in the `InformationSteward\MM\Config` directory.

2. Copy the truststore file to the computer where SAP Business Intelligence Platform 4.0. resides.
 - a. Ensure that the `is.truststore.keystore` file is in a directory that is available to both the computer where SAP BusinessObjects Enterprise XI 3.x is installed and the computer where SAP Business Intelligence Platform 4.0. is installed.
 - b. Copy the `is.truststore.keystore` file to the directory where Metadata Management configuration files are stored for Information Steward.

For example:

```
C:\Program Files (x86)\SAP BusinessObjects\InformationSteward\MM\config
```

3.2.4.2 Configuring the SSL protocol for the Remote Job Server

After you create a key and certificate on the Remote Job Server computer and store them in a secure location, you need to provide the Information Steward Service Configuration on SAP BusinessObjects Enterprise XI 3.x with the secure location.

To configure the SSL protocol in the Central Configuration Manager (CCM):

1. Obtain the SSL certificate and key file names from the SIA Properties file.
 - a. Click **Start** > *All Programs* > *SAP BI platform 4* > *SAP BI platform* > *Central Configuration Manager* to open the CCM.
 - b. Right-click *Server Intelligence Agent* and choose *Properties*.
 - c. Obtain the values for the following options so that you can enter them in a later step:
 - *SSL Certificates Folder*
 - *Server SSL Certificate File*
 - *SSL Trusted Certificate File*
 - *SSL Private Key File*
 - *SSL Private Key File Passphrase File*

For more information about these options, see the *SAP Business Intelligence Platform Administrator's Guide*.

- d. Click *Cancel* to close the *Server Intelligence Agent Properties* window.

2. In Windows, select **Start** > **Programs** > **<Information_Steward_version>** > **Information Steward Service Configuration**.
3. Open the **Protocol** tab.
4. Make sure **Enable SSL** is selected.
5. Enter the values for the SIA options that you obtained from the CCM in step 1c above.
6. Enter the values for **Keystore File** and **Keystore Password** with the values that you specified in the `keystore` and `storepass` options of the `keytool -genkey` command.
7. Enter the values for **Truststore File** and **Truststore Password** with the values that you specified in the `keystore` and `storepass` options of the `keytool -import` command.
8. Click **Apply**.
9. Restart the Remote Job Server.

For more information about SSL servers, see *SAP Business Intelligence Platform Administrator's Guide: Configuring servers for SSL*.

Note

To run an integrator source with SSL enabled on the Remote Job Server, set the following runtime JVM parameter.

```
-Dbusinessobjects.migration=on
```

Related Information

[Creating the keystore and truststore files for the Remote Job Server \[page 40\]](#)

[Setting the runtime parameter for SAP Enterprise 3.x integrator with SSL \[page 238\]](#)

3.2.4.3 SSL and HTTPS setup for Information Steward

These steps will help you avoid potential problems running Cleansing Package Builder if you have configured the Secure Socket Layer (SSL) protocol to work with an HTTPS setup to access a web server for SAP Information Steward.

These steps vary based on the web server you use. Information Steward's default web server is Apache Tomcat.

Before you start these steps you should read about how to deploy and undeploy web services using the WDeploy tool in the *Web Application Deployment Guide* located on the customer portal. For instructions specific to Information Steward, read about how to use WDeploy in the *Installation Guide*.

1. Create a copy of the file named `axis2.xml` located in `<Install_Dir>\SAP BusinessObjects\SAP BusinessObjects Enterprise XI 4.0\warfiles\webapps\CPBWebService\WEB-INF\conf`, and save it to a different directory as a backup.
2. Open the original `axis2.xml` file with a text editor.

3. Add the following text to the file, updating the port number and the host as indicated in the bulleted list after the code:

```
<transportReceiver name="https"  
class="org.apache.axis2.transport.http.SimpleHTTPServer">  
  <parameter name="port">443</parameter>  
  <parameter name="hostname">https://[host]</parameter>  
</transportReceiver>
```

- The port number is the SSL port number that you use with your BusinessObjects Enterprise (BOE) and Information Steward web service (we show “443” as an example).
 - Host is the name of the server that hosts the Information Steward application. For example, if you use a web application named “myserver” to access Information Steward, set [host] to `myserver` (omitting the brackets).
4. Undeploy the CPBWebService using WDeploy. (See the topic “To undeploy one web component” in the *Installation Guide*).
 5. Perform the predeployment steps explicitly for CPBWebService. For more information see the topic “WDeploy prerequisites” in the *Installation Guide*.
 6. Redeploy the CPBWebService. For details see the topic “Deploying web applications with WDeploy” in the *Web Application Deployment Guide*.

3.2.5 Reverse proxy servers

SAP Information Steward can be deployed in an environment with one or more reverse proxy servers. A reverse proxy server is typically deployed in front of the web application servers in order to hide them behind a single IP address. This configuration routes all Internet traffic that is addressed to private web application servers through the reverse proxy server, hiding private IP addresses.

Because the reverse proxy server translates the public URLs to internal URLs, it must be configured with the URLs of the Information Steward web applications that are deployed on the internal network.

For information about supported reverse proxy servers and how to configure them, see “Information platform services and reverse proxy servers” and “Configuring reverse proxy servers for Information platform” in the *SAP information platform services Administrator's Guide*.

4 Information Steward configuration settings

Use configuration settings to establish directory locations, to control software behavior when performing tasks, and to control performance and scalability.

The *Information Steward Settings* window in the Central Management Console (CMC) contains many settings that control how the software behaves when performing tasks in Data Insight, Metadata Management, Metapedia, Cleansing Package Builder, and Match Review.

4.1 Configuring Information Steward settings

Administrators or users with the required permissions can set or change the Information Steward configuration settings.

1. Log in to the Central Management Console (CMC) and select *Applications*.
2. Double-click *Information Steward application*.

The *Information Steward Settings* window opens.

3. Select *Configure Application* in the navigation pane at left.

The configuration settings appear at right.

4. For each option, accept the current setting or adjust the settings as applicable. For option descriptions, see [Information Steward settings option descriptions \[page 45\]](#).
5. Click *Save* to apply the settings. Changes that affect the user interface are visible to users when they log out and log back in.

To reset the settings to their default values, click *Reset*. To cancel the changes that you made to the settings and before you have saved them, click *Cancel*.

4.2 Information Steward settings option descriptions

The *Information Steward Settings* window in the *Applications* area of the Central Management Console (CMC) contains options that control software behavior and job performance in Information Steward modes.

Group	Parameters	Default value	Description
<i>Profiling</i> (Data Insight)	<i>Max input size</i>	-1	Maximum number of rows to profile. -1 processes all records Set a positive maximum number of rows to input for profiling to reduce memory consumption, especially for profiling large tables.
	<i>Input sampling rate</i>	1	Rate at which rows are read. 1 reads every row, 2 reads every second row, and so on.
	<i>Max sample data size</i>	25	Maximum number of records to save for each profile attribute.
	<i>Number of distinct values</i>	25	Number of distinct values to store for the data distribution result.
	<i>Number of patterns</i>	25	Number of patterns to store for pattern distribution result.
	<i>Number of words</i>	25	Number of words to store for word distribution result.
	<i>Results retention period</i>	90	Duration in days before the profiling results are deleted from the Information Steward repository by the Purge utility.
	<i>Optimization period</i>	24	Duration in hours before the software replaces profiling results with results from a scheduled task. Set to 0 hours to refresh results with every task run.

Group	Parameters	Default value	Description
	<i>Address Reference Data directory</i>	<install_dir>\Data Services\DataQuality	<p>Directory location for reference data used for address profiling and Data Cleansing Advisor processing.</p> <p>If no value is entered, Data Cleansing Advisor uses address reference data stored in the Data Quality subdirectory of the Data Services installation directory.</p>
<i>Rules (Data Insight)</i>	<i>Max input size</i>	-1	<p>Maximum number of rows to validate for rules.</p> <p>-1 validates rules for all rows in a source.</p> <p>Set a positive maximum number of rows for rule validation to reduce memory consumption, especially for large tables.</p>
	<i>Input sampling rate</i>	1	<p>Rate at which rows are read.</p> <p>1 reads every row, 2 reads every second row, and so on.</p>
	<i>Max sample data size</i>	100	<p>Maximum number of failed records to save for each rule.</p> <p>You can set up to a maximum of 500 rows.</p>
	<i>Score retention period</i>	1095	<p>Number of days before the scores are purged.</p> <p>The default, 1095 days, is approximately 3 years.</p>
	<i>Optimization period</i>	1	<p>Number of hours before a new score is calculated for a table by a scheduled task.</p> <p>Set to 0 to always refresh immediately.</p>

Group	Parameters	Default value	Description
	<i>Default low threshold</i>	5	<p>Default value to indicate a low threshold when binding a rule and defining a scorecard.</p> <p>The range is 1 to 10.</p> <p>You can change this value, or override the value for an individual rule or scorecard.</p>
	<i>Default high threshold</i>	8	<p>Default value to indicate a high threshold when binding a rule and defining a scorecard.</p> <p>The range is 1 to 10.</p> <p>You can change this value, or override the value for individual rule or scorecard.</p>
	<i>Allow author to approve</i>	Yes	<p>Allows an author in the approval group to approve any new rules that they create. If you choose <i>No</i>, the author will not be able to approve the rules they create, even if they are part of the approval group.</p>
<i>Performance</i> (Data Insight)	<i>Degree of parallelism</i>	4	<p>Number of parallel processing threads for a task.</p>
	<i>Average concurrent tasks</i>	5	<p>Average number of profile or rule tasks that can execute in parallel in the Data Services job server group of Information Steward.</p> <p>Information Steward multiplies this value by the number of job servers available in the job server group. For example, if <i>Average concurrent tasks</i> is set to 5 and the job server group contains 2 job servers, then 10 tables or subtables can be processed in parallel.</p>

Group	Parameters	Default value	Description
	<i>Number of processing threads</i>	2	Maximum number of threads for file processing.
	<i>Pageable cache directory</i>	N/A	<p>Directory location to use for caching if profiling and rule tasks require more memory.</p> <p>Ensure that you specify a pageable cache directory that contains enough disk space for the amount of data you plan to profile.</p>
	<i>Task distribution level</i>	Table	<p>Level to distribute the execution of a profile or rule task in an SAP Data Services job server group. The term "table" in this context means table, file, or view.</p> <p>Table: The entire table executes on one job server.</p> <p>Subtable: The table is partitioned to execute in parallel on multiple job servers.</p> <p>If you specify Subtable, ensure that:</p> <ul style="list-style-type: none"> You set up a pageable cache directory on a shared directory on which all job servers in the server group have access. The network is efficient between the job servers in the group. The requirements for a Data Services job server group are met.
<i>Debug Options</i> (Data Insight)	<i>Data Services engine options</i>	N/A	Values passed as command line parameters for debugging.

Group	Parameters	Default value	Description
<i>Permissions</i> (Data Insight)	<i>Connection on which to base failed data</i>	Source Connection	<p>Connection type to base view and export permission for failed data.</p> <p>Source connection: Users need permission to the source connection to view and export failed data. Administrators grant permission to the source connection and assign the user required permissions. If user does not have these permissions, user cannot view or export failed data.</p> <p>Failed data connection: Users need permission to the failed data connection to view and export failed data. Administrators grant permission to the failed data connection and assign the required permissions. If user does not have these permissions, user cannot view or export failed data.</p> <p>Required permissions:</p> <ul style="list-style-type: none"> • View data • View sample data • Export data
	<i>Data Review and Worklist settings</i>	<i>Data Review tasks retention period</i>	-1
	<i>Worklist URL</i>	localhost	<p>URL to which match review tasks are assigned. Reviewer receives an e-mail containing the URL.</p>

Group	Parameters	Default value	Description
<i>Data Cleansing Advisor Settings</i>	<i>Data Cleansing Advisor retention period</i>	120	<p>Number of days before the Data Cleansing Advisor staged data is deleted. The longer you keep the data, the larger the repository size.</p> <p>If you set the value to -1, the staged data is never purged.</p>
<i>Search preferences</i>	<i>Update time (seconds)</i>	30	<p>Delay time before changes in Information Steward affect the search results.</p>
<i>Metadata Management Options</i>	<i>Custom reports directory</i>	N/A	<p>Full path to the directory for custom reports, such as Crystal reports, and associated properties file. Applicable for Crystal Reports version 12 or later.</p>
<i>Metapedia options</i>	<i>Allow duplicated term names</i>	No	<p>Allows Metapedia terms to share the same name.</p> <p>The Technical Name field in Metapedia must be unique. Therefore, even though the term shares the same name with another term, the technical name is unique.</p>
	<i>Allow author to approve</i>	Yes	<p>Allows an author in the approval group to approve any new rules that they create. If you choose <i>No</i>, the author will not be able to approve the rules they create, even if they are part of the approval group.</p>

Group	Parameters	Default value	Description
<i>Email Settings</i>	<i>To Email Address</i>	No	<p>E-mail address for the user or user group who receive utility service monitor alerts. Utility service monitor alerts inform the designated recipient or group that a scheduled utility has failed to run or is still running.</p> <p>If you do not enter an e-mail address for this option (<i>To Email Address</i>), Information Steward uses the e-mail address in the <i>To</i> option in the EIM Adaptive Job Server <i>Destination</i> settings in the CMC.</p>

Related Information

[Data Insight performance settings \[page 124\]](#)

[Resource intensive functions \[page 289\]](#)

[Setting up e-mail notification configuration \[page 100\]](#)

5 Users and Groups Management

5.1 Users and Groups overview

The Central Management Console (CMC) manages security information, such as user accounts, group memberships, and object rights that define user and group privileges. When a user attempts an action on an Information Steward object, the CMC authorizes the action only after it verifies that the user's account or group membership has sufficient privileges.

Use SAP BusinessObjects Business Intelligence platform security to create users and authorize user access to the objects and actions within the Information Steward modules:

- Data Insight
- Metadata Management
- Metapedia (under Metadata Management in the Central Management Console)
- Cleansing Package Builder
- Match Review

Information Steward provides pre-defined user groups that have specific rights on objects unique to each module. These user groups enable you to grant rights to multiple users by adding the users to a group instead of modifying the rights for each user account individually. You also have the ability to create your own user groups.

Related Information

[Information Steward pre-defined users and groups \[page 52\]](#)

[Managing users in Information Steward \[page 54\]](#)

[User rights in Data Insight \[page 58\]](#)

[User rights in Metadata Management \[page 72\]](#)

[User rights in Metapedia \[page 79\]](#)

[User rights in Cleansing Package Builder \[page 83\]](#)

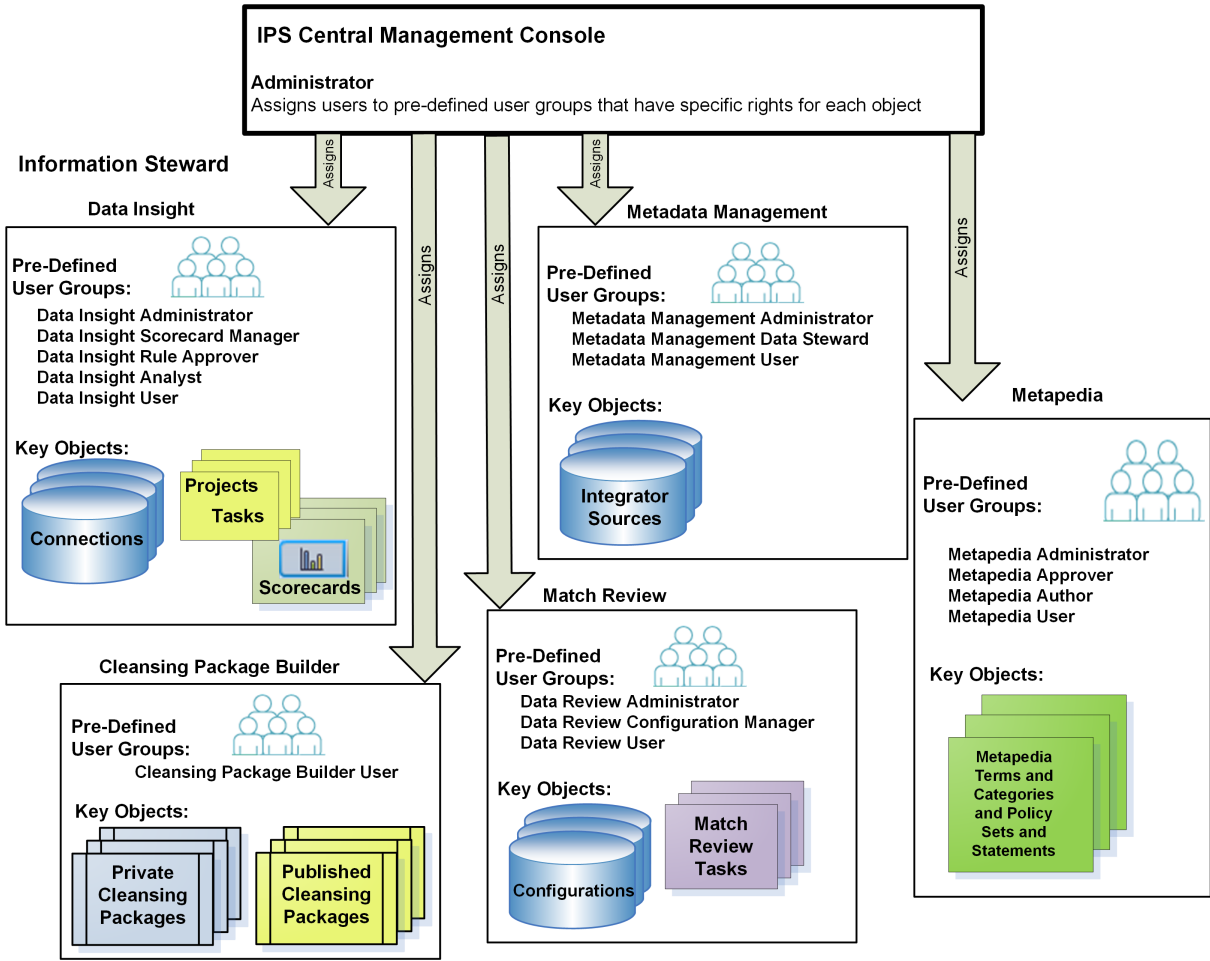
[User rights in Match Review \[page 86\]](#)

5.2 Information Steward pre-defined users and groups

Information Steward provides pre-defined user groups for each module to facilitate managing security on the objects within each module.

Each pre-defined user group has specific rights to objects in Information Steward modules.

The following diagram provides an overview of the pre-defined Information Steward user groups and their relation to the Administrator group in the Information Platform Services (IPS).



- Members of the Administrator group are responsible for the following tasks in the IPS:
 - Create users and custom groups for all Information Steward modules
 - Perform all tasks within all Information Steward modules
 - Grant access to cleansing packages

The table below describes the role for administrators for each Information Steward module.

Administrator roles in Information Steward

Module	Administrator role
Data Insight	<ul style="list-style-type: none"> Grants users and groups access to connections and projects. (By default, all Data Insight pre-defined groups are granted access to all connections and projects.) Accesses all Data Insight actions.
Metadata Management	<ul style="list-style-type: none"> Grants user access to integrator sources. Accesses all Metadata Management actions.

Module	Administrator role
Metapedia	<ul style="list-style-type: none"> Grants user access to term approval roles. Grants user access to policy set and statement approval roles. Accesses all Metapedia actions.
Data Review	<ul style="list-style-type: none"> Creates connections for match review and schedules for match review configurations. Accesses all Match Review actions.

Within each Information Steward module, additional user groups have specific rights for the objects within that module. For example, the Data Insight Analyst group can create profile tasks and rules, but only the Data Insight Rule Approver can approve rules.

Related Information

[Data Insight predefined user groups \[page 59\]](#)

[Type-specific rights for Data Insight objects \[page 60\]](#)

[Metadata Management pre-defined user groups \[page 73\]](#)

[Type-specific rights for Metadata Management objects \[page 73\]](#)

[Group rights for cleansing packages \[page 84\]](#)

[User rights in Cleansing Package Builder \[page 83\]](#)

[Match Review pre-defined user groups \[page 86\]](#)

[Type-specific rights for Match Review \[page 87\]](#)

5.3 Managing users in Information Steward

This section contains the steps to create users and add them to Information Steward groups.

5.3.1 Creating users for Information Steward

Create user accounts and assign them to groups or assign rights to control their access to objects in SAP Information Steward.

To create users:

1. Log in to the Central Management Console (CMC) with a user name that belongs to the Administrator group.
2. At the CMC home page, click *Users and Groups*.
3. Click **Manage** > **New** > **New User**.

4. To create a user:

- a. Select your authentication type from the *Authentication Type* list.

Information Steward can use the following user authentication:

- Enterprise (default)
- LDAP (can be used with single sign-on)
- Windows Active Directory (AD) (can be used with single sign-on)
- SAP ERP and Business Warehouse (BW)

For more information, see the “Single Sign-On Setup” topic in the *SAP BusinessObjects Business Intelligence platform Administrator Guide*.

- b. Type the account name, full name, email, and description information.

→ Tip

Use the description area to include extra information about the user or account.

- c. Specify the password information and settings.

5. To create a user that will logon using a different authentication type, select the appropriate option from the *Authentication Type* list, and type the account name.

6. Specify how to designate the user account according to options stipulated by your SAP Business Intelligence platform license agreement.

If your license agreement is based on user roles, select one of the following options:

- *BI Viewer*: Access to Business Intelligence platform applications for all accounts under the BI Viewer role is defined in the license agreement. Users are restricted to access application workflows that are defined for the BI Viewer role. Access rights are generally limited to viewing business intelligence documents. This role is typically suitable for users who consume content through Business Intelligence platform applications.
- *BI Analyst*: Access to SAP Enterprise applications for all accounts under the BI Analyst role is defined in the license agreement. Users can access all applications workflows that are defined for the BI Analyst role. Access rights include viewing and modifying business intelligence documents. This role is typically suitable for users who create and modify content for Business Intelligence platform applications.

If your license agreement is not based on user roles, specify a connection type for the user account.

- Choose *Concurrent User* if this user belongs to a license agreement that states the number of users allowed to be connected at one time.
- Choose *Named User* if this user belongs to a license agreement that associates a specific user with a license. Named user licenses are useful for people who require access to SAP BusinessObjects Enterprise regardless of the number of other people who are currently connected.

7. Click *Create & Close*.

The user is added to the system and is automatically added to the Everyone group. An inbox is automatically created for the user, as is an Enterprise alias. You can now add the user to a group or specify rights for the user. For more information, see the “Managing users and groups” topic in the *SAP BusinessObjects Business Intelligence platform Administrator Guide*.

5.3.2 Adding users and user groups to Information Steward groups

Groups are collections of users who share the same rights to different objects. SAP Information Steward provides groups for Data Insight and Metadata Management, such as Data Insight Analyst group and Metadata Management User group.

To add a user to a group:

1. Log in to the Central Management Console (CMC) with a user name that belongs to the Administrator group.
2. At the CMC home page, click *Users and Groups*.
3. Select the *User List* or *Group List* node in the navigation tree.
4. Select the name of the user or user group in the right panel.
5. Click **► Actions ► Join Group ▾**.
6. On the *Join Group* dialog box, select the *Group List* node in the navigation tree.
7. Select one or more names from the *Available Groups* list, and click > to place them in the *Destination Group(s)* list.
8. Click *OK*.

Related Information

[Data Insight predefined user groups \[page 59\]](#)

[Group rights for connections \[page 63\]](#)

[Group rights for projects \[page 64\]](#)

[Group rights for tasks \[page 66\]](#)

[Metadata Management pre-defined user groups \[page 73\]](#)

5.3.3 Denying access

To deny access to Information Steward for an individual user or group:

1. Log in to the Central Management Console (CMC) with a user name that is a member of either the Administrator group or the Data Insight Administrator group.
2. At the CMC home page, click *Information Steward*.
3. Click **► Manage > Security > User Security ▾**.
4. On the *User Security* page, click *Add Principals*.
5. In the list of *Available users/groups*, select the name of each user or group that you want to deny access, and click the > button to move the names to the *Selected users/groups* list.
6. Click *Add and Assign Security*.
7. Click the *Advanced* tab.

8. Click the [Add/Remove Rights](#) link.
9. On the General rights list, deny the [View objects](#) right.
10. Click [Apply](#) or [OK](#).

5.3.4 Deleting users in Information Steward

Deleting a user removes them from all user groups to which they belong and removes their right to access and control objects in SAP Information Steward.

To delete a user:

1. Log in to the Central Management Console (CMC) with a user name that belongs to the Administrator group.
2. At the CMC home page, click [Users and Groups](#).
3. Select the [User List](#) node in the navigation tree.
4. Select the name of the user in the right panel.
5. Click [Manage > Delete](#).
6. Click [OK](#).

Related Information

[Creating users for Information Steward \[page 54\]](#)

[Adding users and user groups to Information Steward groups \[page 56\]](#)

5.3.5 Setting user e-mail address for notifications

Administrators add users e-mail information in the Central Management Console (CMC) so the user can receive e-mail notifications for specific events.

Before SAP Information Steward sends e-mail notifications to users, you must configure the EIM Adaptive Job Server Destination E-mail settings for SMTP domain, host, port, and other settings.

1. Log in to the CMC and select [Users and Groups](#) from the [CMC Home](#) dropdown menu.
2. Click [User List](#) or [Group List](#) as applicable from the navigation tree at left.
A list of all users appears at right.
3. Select the applicable user name and click [Manage > Properties](#).
4. Complete the [Email](#) text box with the user e-mail address.
5. Click [Save & Close](#).

Related Information

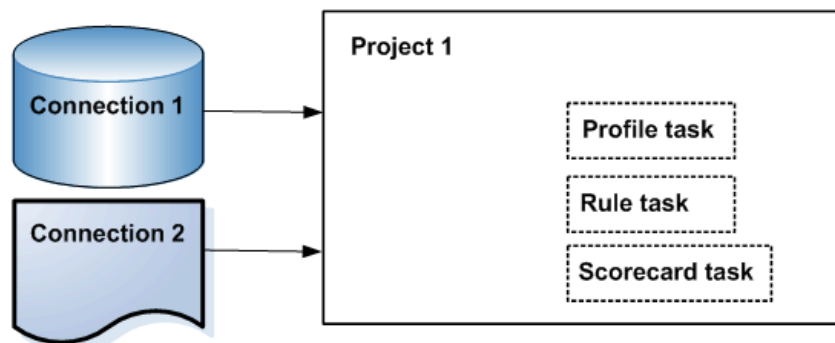
[Setting up e-mail notification configuration \[page 100\]](#)

5.4 User rights in Data Insight

The Data Insight module of SAP Information Steward contains the following objects that have specific rights that allow various actions on them.

- Connections through which users view data sources and import tables and files to profile the data. In addition to rights to the connection, a user must also be granted permission on the source data: For database connections, the Database Administrator must grant privileges on the tables to the user. For file connections, the users that run the following services must have permissions on the directory where the file resides:
 - Information Steward Web Application Server (for example, Tomcat)
 - Data Services service
 - Server Intelligence Agent that runs EIMAdaptiveProcessingServer and ISJobServer
- Views that can join tables and files from multiple connections.
- Projects that contain profile tasks, rule tasks, and scorecards in specific business areas, such as HR or Sales.
- Profile tasks that collect profile attributes to help you determine the quality and structure of the data.
- Rule tasks that validate the data according to your business and quality rules.

The following diagram shows users and groups who are granted rights to access Connections, Projects, and Tasks for Data Insight.



Data Insight User Groups	Rights on Connections	Rights on Projects	Rights on Tasks
Data Insight Administrator	<ul style="list-style-type: none"> Create connection Edit connection Delete connection 	<ul style="list-style-type: none"> Create Project Delete Project 	
Scorecard Administrator		<ul style="list-style-type: none"> Create scorecard 	
Rule Approver		<ul style="list-style-type: none"> Approve or reject rule 	
Profiler	<ul style="list-style-type: none"> Create profile task Create rule task 	<ul style="list-style-type: none"> Create profile task Create or edit rule Bind rule 	<ul style="list-style-type: none"> Edit task Schedule task Delete Schedule View Schedule or task
Data Insight Business User	<ul style="list-style-type: none"> View connection View data View sample data 	<ul style="list-style-type: none"> View Project View Profile results View rules View Scorecard results 	

Related Information

[Data Insight predefined user groups \[page 59\]](#)

[Type-specific rights for Data Insight objects \[page 60\]](#)

[Customizing rights on Data Insight objects \[page 67\]](#)

[Managing users in Information Steward \[page 54\]](#)

5.4.1 Data Insight predefined user groups

Information Steward provides the following predefined Data Insight user groups to facilitate the assignment of rights on connections, projects, and tasks. These groups enable you to change the rights for multiple users in one place (a group) instead of modifying the rights for each user account individually.

→ Tip

Users working on the same project should have access to the same set of connections so that they can collaborate on the same data set within a project.

The following table describes the predefined user groups in ascending order of rights.

User group	Description
Data Insight User	Users can view the connections, projects, source data, profile results, sample profile data, rules, sample data that failed rules, and scorecard results.
Data Insight Analyst	Users have all of the rights of a Data Insight User, plus the following rights: <ul style="list-style-type: none"> • Add tables and files to a project • Remove tables and files from a project • Create, edit, and delete profile tasks and rule tasks • Create, edit, and delete rules, bind rules • Schedule profile tasks and rule tasks
Data Insight Rule Approver	Users have all of the rights of a Data Insight Analyst, plus the right to approve and reject rules.
Data Insight Scorecard Manager	Users have all of the rights of a Data Insight Analyst, plus the right to create and edit scorecards that consist of rules for specific Key Data Domains.
Data Insight Administrator	Users have all of the user group rights listed above plus the following rights: <ul style="list-style-type: none"> • Configure, edit, delete, run, schedule, view history of Information Steward utilities • Create, edit, and delete Data Insight connections and projects • Configure Information Steward application settings • Change Information Steward repository user and password

Related Information

[Group rights for Data Insight folders and objects in the CMC \[page 61\]](#)

[Group rights for connections \[page 63\]](#)

[User rights for views \[page 65\]](#)

[Group rights for projects \[page 64\]](#)

[Group rights for tasks \[page 66\]](#)

5.4.2 Type-specific rights for Data Insight objects

Rights are the base units for controlling user access to the objects, users, applications, servers, and other features in SAP BusinessObjects Enterprise.

Type-specific rights are rights that affect specific object types only, such as Data Insight connections, Data Insight projects, profile tasks, or rule tasks.

Rights are set on objects, such as a Data Insight connection or project, rather than on the "principals" (the users and groups) who access them. By default, the pre-defined Data Insight user groups are granted access to newly created connections and projects. If you want some users to access only certain connections and projects, then do not add them to a pre-defined group, but add their user names to the list of principals for each individual connection and project and assign the appropriate type -specific rights. For example, to give

a user access to a particular connection, you add the user to the list of principals who have access to the connection.

Type-specific rights consist of the following:

- General rights for the object type
These rights are identical to general global rights (for example, the right to add, delete, or edit an object), but you set them on specific object types to override the general global rights settings.
- Specific rights for the object type
These rights are available for specific object types only. For example, the right to view data appears for Data Insight connections but not for Data Insight projects.

For more information about user rights, see "How rights work in SAP Enterprise" in the *Enterprise Administrator's Guide*.

Related Information

[Group rights for connections \[page 63\]](#)

[Group rights for projects \[page 64\]](#)

[Group rights for tasks \[page 66\]](#)

5.4.2.1 Group rights for Data Insight folders and objects in the CMC

Each pre-defined Data Insight user group provides specific rights on the folders and objects in the CMC, as the following tables show.

Rights for Data Insight folders in the CMC

CMC Folder or Object	Right Name	Description	Pre-defined User Groups				
			Data Insight Administrator	Data Insight Scorecard Manager	Data Insight Rule Approver	Data Insight Analyst	Data Insight User
Data Insight folder	View objects	View Data Insight folder on the CMC	Yes	Yes	Yes	Yes	Yes
	Modify Rights	Modify the rights users have to Data Insight	Yes	No	No	No	No
Connections folder	View objects	View connections in Connections folder	Yes	Yes	Yes	Yes	Yes
	Add Objects to Folder	Create connection	Yes	No	No	No	No

CMC Folder or Object	Right Name	Description	Pre-defined User Groups				
			Data Insight Administrator	Data Insight Scorecard Manager	Data Insight Rule Approver	Data Insight Analyst	Data Insight User
	Modify Rights	Modify the rights users have to connections	Yes	No	No	No	No
Data Insight Connection	View objects	View connection properties	Yes	Yes	Yes	Yes	No
	Edit objects	Edit connection properties	Yes	No	No	No	No
	Delete objects	Delete connection	Yes	No	No	No	No
	Modify Rights	Modify the rights users have to Data Insight connections	Yes	No	No	No	No
Projects folder	View objects	View Projects folder	Yes	Yes	Yes	Yes	Yes
	Add objects	Create project	Yes	No	No	No	No
	Modify Rights	Modify the rights users have to projects	Yes	No	No	No	No
Projects	View objects	View project properties	Yes	Yes	Yes	Yes	No
	Edit objects	Edit project properties	Yes	No	No	No	No
	Delete objects	Delete project	Yes	No	No	No	No
	Modify Rights	Modify the rights users have to projects	Yes	No	No	No	No
Profile task or Rule task	View objects	View task properties	Yes	Yes	Yes	Yes	Yes
	View instance	View task history and logs	Yes	Yes	Yes	Yes	No
	Schedule	Schedule task	Yes	Yes	Yes	Yes	No
	Reschedule	Change the schedule of task	Yes	No	No	No	No
	Delete objects	Delete task	Yes	Yes	Yes	Yes	No
	Modify Rights	Modify the rights users have to tasks.	Yes	No	No	No	No

Related Information

[User rights for Information Steward administrative tasks \[page 84\]](#)

5.4.2.2 Group rights for connections

SAP Information Steward provides pre-defined Data Insight user groups that have specific rights to connections. You can add users to these groups to control their rights to connections.

For example, in *Data Insight*, if a user has the right to view a connection, the connection is visible in the *Browse Metadata* window. In addition, the tables and views in that connection are visible.

Rights for Data Insight Connections

Right	Description	Pre-defined User Groups				
		Data Insight Administrator	Data Insight Scorecard Manager	Data Insight Rule Approver	Data Insight Analyst	Data Insight User
View objects	View connections and tables, browse metadata in the connection	Yes	Yes	Yes	Yes	Yes
View Data	View external data in the connection	Yes	Yes	Yes	Yes	Yes
View Sample Data	View profile sample data and sample data that failed rules	Yes	Yes	Yes	Yes	Yes
Export Data	Export viewed data, profile sample data, and sample data that failed rules	Yes	Yes	Yes	Yes	Yes
Profile/Rule permission	Create profile tasks and rule tasks	Yes	Yes	Yes	Yes	No

Note

- Rights to a Data Insight connection are granted to users or groups when the Data Insight Administrator adds them to the *Principals* list for the connection.
- The Administrator and the Data Insight Administrator can create, edit, and delete connections in the CMC.
- For a database connection, the Database Administrator must grant the Data Insight user access to the tables.
- For a file connection, the users that run the following services must have permissions on the directory where the file resides:
 - Information Steward Web Application Server (for example, Tomcat)
 - Data Services service
 - Server Intelligence Agent that runs EIM Adaptive Processing Server and IS Job Server

Related Information

[Group rights for Data Insight folders and objects in the CMC \[page 61\]](#)

[Group rights for Data Insight folders and objects in the CMC \[page 61\]](#)

[User rights in Data Insight \[page 58\]](#)

[Data Insight predefined user groups \[page 59\]](#)

[Assigning users to specific Data Insight objects \[page 69\]](#)

5.4.2.3 Group rights for projects

Each pre-defined Data Insight user group provides specific rights on projects in Information Steward, as the following table shows.

Rights for Data Insight Projects on Information Steward

Right	Description	Pre-defined User Groups				
		Data Insight Administrator	Data Insight Scorecard Manager	Data Insight Rule Approver	Data Insight Analyst	Data Insight User
View objects	<ul style="list-style-type: none"> View project View profile results View rule results View rules View scorecard result View Preferences page 	Yes	Yes	Yes	Yes	Yes
Edit objects	<ul style="list-style-type: none"> Add and remove tables and files in project Add, edit, and remove views Update Preferences page 	Yes	Yes	Yes	Yes	No
Add objects	<ul style="list-style-type: none"> Create profile tasks Create rule tasks Copy views 	Yes	Yes	Yes	Yes	No
Manage rule	<ul style="list-style-type: none"> Create, edit, and remove rules Bind rules 	Yes	Yes	Yes	Yes	No
Import	<ul style="list-style-type: none"> Import rules Import views 	Yes	Yes	Yes	Yes	No
Manage scorecard	Add, edit, and remove Key Data Domains	Yes	Yes	No	No	No

Pre-defined User Groups						
Right	Description	Data Insight Administrator	Data Insight Scorecard Manager	Data Insight Rule Approver	Data Insight Analyst	Data Insight User
Approve rule	Approve or reject rules on Information Steward	Yes	No	Yes	No	No
Admin rule	Edit, review and re-assign rules that are assigned to others.	Yes	No	No	No	No

On the *Data Insight* tab of Information Steward:

- A project is visible if a user has the right to view a project.
- The *Add Tables* and *Remove* buttons are enabled in the Workspace if a user has the right to edit a project.
- The *Bind* and *Delete* buttons are enabled on the *Rule* tab if a user has the right to manage rules.
- The *Profile* and *Calculate Score* buttons are enabled in the Workspace if a user has the right to add objects to the project.

Note

Only the Administrator and Data Insight Administrator can create, edit, and delete projects in the CMC. For more information, see [Group rights for Data Insight folders and objects in the CMC \[page 61\]](#).

5.4.2.4 User rights for views

The rights each user has on a view is inherited from the rights the user has on the connections that comprise the view.

For example, suppose View1 is comprised of the following connections and tables:

- ConnectionA, Table1
- ConnectionB, Table2

Suppose User1 has the Edit right on ConnectionA but not on ConnectionB. Therefore, User1 cannot edit View1 because the denied Edit right is inherited from ConnectionB.

Similarly, if User2 has the Edit right on ConnectionA and ConnectionB, then User2 can edit View1.

This inheritance applies to all of the rights on views, as the following table shows.

Rights for Data Insight Views on Information Steward

Right	Description	Required inherited right
View objects	The view name and columns are visible in Workspace Home window.	View objects right on all source connections
View Data	Look at the source data in each table or file that comprise a view	View Data right on all source connections

Right	Description	Required inherited right
View Sample Data	View profile sample data and sample data that failed rules	View Sample Data right on all source connections
Profile/Rule permission	Create profile tasks and rule tasks	Profile/Rule right on all source connections
Export Data	Export viewed data, profile sample data, and sample data that failed rules	Export Data right on all source connections

Note

The following actions require rights on the project:

- To add, edit, or remove views, a user must have the Edit objects right on the project.
- To copy a view, a user must have the Add objects right on the project.

5.4.2.5 Group rights for tasks

Each pre-defined Data Insight user group provides specific rights on profile tasks and rule tasks in Information Steward, as the following table shows.

Rights for Data Insight Tasks on Information Steward

Right	Description	Pre-defined User Groups				
		Data Insight Administrator	Data Insight Scorecard Manager	Data Insight Rule Approver	Data Insight Analyst	Data Insight User
View objects	View the task in the Tasks tab of the Workspace	Yes	Yes	Yes	Yes	Yes
Edit objects	<ul style="list-style-type: none"> • Edit profile task • Edit rule task or score calculator task 	Yes	Yes	Yes	Yes	No
Delete objects	Delete profile or rule task	Yes	Yes	Yes	Yes	No
Schedule	<ul style="list-style-type: none"> • Run profile task • Run rule task 	Yes	Yes	Yes	Yes	No

Related Information

[Group rights for Data Insight folders and objects in the CMC \[page 61\]](#)

5.4.3 Type-specific rights for source or failed data connections

Administrators control user rights and the access connection type for viewing and exporting failed data.

Users can view failed data, view more failed data, and export failed data when they have permission to access the connection and the following specific permissions in the connection:

- View data
- View sample data
- Export data

Administrators control the access connection on which the users view or export failed data. The option is *Connection on which to base failed data* and it is in the Central Management Console in *Data Insight Settings*.

Connection on which to base failed data	Description
Source connection	Users need permission to the source connection to view and export failed data.
Failed data connection	Users need permission to the failed data connection to view and export failed data.

Related Information

[Assigning users to specific Data Insight objects \[page 69\]](#)

5.4.4 Customizing rights on Data Insight objects

To facilitate user management, assign users to a pre-defined Information Steward user group. By default, Information Steward assigns all pre-defined user groups to all Data Insight connections and projects. However, you might want to limit a user's access in the following ways:

- Only view and profile data in a subset of connections
- Only create rules on a subset of projects
- Only create scorecard and approve rules on a subset of projects
- Restrict project access
- Restrict access to Data Insight
- Only view and export failed data in the source or failed data connection

To limit access, take one of the following actions:

- Add the user or group to only a subset of connections. In addition, create projects for specific business areas, such as HR or Sales, and assign only certain users to access these projects to create profile tasks and rule tasks.

- Deny a right from an existing user group for a specific connection and specific project.

Related Information

[Assigning users to specific Data Insight objects \[page 69\]](#)

[Denying user rights to specific Data Insight objects \[page 68\]](#)

[Data Insight predefined user groups \[page 59\]](#)

5.4.4.1 Denying user rights to specific Data Insight objects

By default, the pre-defined Data Insight user groups are added to the access list of connections and projects when you create them. However, you might want to deny one or a small subset of users access to a specific Data Insight connection and project.

To allow a user or group to access all but one specific Data Insight connection and project:

1. Log in to the Central Management Console (CMC) with a user name that is a member of either the Administrator group or the Data Insight Administrator group.
2. Add the user name to a pre-defined Data Insight group because the user would still have access to most connections and projects. For details, see [Adding users and user groups to Information Steward groups \[page 56\]](#).
3. At the CMC home page, click *Information Steward*.
4. Select the object type.
 - To deny rights to a connection:
 - Select the *Connections* node in the Tree panel.
 - Select the connection name in the right panel.
 - To deny rights to a project:
 - Expand the *Data Insight* node, and expand the *Projects* node in the Tree panel.
 - Select the project name in the Tree panel.
 - To deny rights to a profile or rule task:
 - Expand the *Data Insight* node, and expand the *Projects* node in the Tree panel.
 - Select the project name in the Tree panel.
 - Select the task name in the right panel.
5. Click **Manage > Security > User Security**.
6. On the *User Security* page, click *Add Principals*.
7. In the list of *Available users/groups*, select the name of each user or group that you want to deny rights to this connection or project, and click the > button to move the names to the *Selected users/groups* list.

Note

To select multiple names, hold down the **Ctrl** key when you click each name.

8. Click *Add and Assign Security*.

9. Click the *Advanced* tab.
10. Click the *Add/Remove Rights* link.
11. To deny rights to this connection:
 - a. Expand the *Application* node and select *Data Insight Connection*.
 - b. Click the *Denied* column for each right that you want to deny this user or group.

For example, to deny the right to view and profile the data in this connection, click the *Deny* column for the following *Specific Rights for Data Insight Connection*:

 - *Profile/Rule permission*
 - *View Data*
 - *View Sample Data*
12. To deny project rights:
 - a. Expand the *Application* node and select *Data Insight Project*.
 - b. Click the *Denied* column for each right that you want to deny this user or group.

For example, to deny the right to create scorecard and approve rules in this project, click the *Denied* column for the following *Specific Rights for Data Insight Connection*:

 - *Approve Rule*
 - *Manage Rule*
13. To deny profile or rule task rights:
 - a. Expand the *Application* node and select *Information Steward Profiler Task*.
 - b. Click the *Override General Global* column and the *Denied* column for each right that you want to deny this user or group.

For example, to deny the right to schedule a profile task or rule task, click the *Override General Global* column and the *Denied* column for the following *General Rights for Data Insight Profiler Task*:

 - *Reschedule instances that the user owns*
 - *Schedule document to run*
 - *View document instances*
14. Click *OK* and verify that the list under *Right Name* does not display the rights you denied.
15. Click name of the principal you just added, click *View Security* and verify that the list of rights that you denied has the red icon in the *Status* column.
16. Click *OK* and close the *User Security* window.

Related Information

[Data Insight predefined user groups \[page 59\]](#)

5.4.4.2 Assigning users to specific Data Insight objects

Administrators can limit user or group access to only specific Data Insight connections, projects, and tasks. Add the user or group to the list of principals for each specific Data Insight object.

Note

You must assign the user to both the connection and the project to enable them to add tables or files, create profile tasks, and create rules.

To add a user or group to a specific connection and project:

1. Log in to the Central Management Console (CMC) with a user name that is a member of either the Administrator group or the Data Insight Administrator group.
2. At the CMC home page, click *Information Steward*.
3. Select the object type.
 - For a connection:
 1. Select the *Connections* node in the tree panel.
 2. Select the connection name in the right panel.
 - For a project:
 1. Expand the *Data Insight* node, and expand the *Projects* node in the tree panel.
 2. Select the project name in the tree panel.
 - For a profile task or rule task:
 1. Expand the *Data Insight* node, and expand *Projects* node in the tree panel.
 2. Select the task name in the right panel.
4. Click **Manage** > *User Security*.
5. On the *User Security* page, click *Add Principals*.
6. In the list of *Available Users or Groups*:
 - a. Select the name of each user or group that you want to authorize to this connection, project, or task.
 - b. Click the button to move the names to the *Selected Users or Groups* list.
7. Click *Add and Assign Security*.
8. Click the *Advanced* tab.
9. Click the *Add/Remove Rights* link.
10. To assign connection rights:
 - a. Expand the *Application* node and select *Information Steward Connection*.
 - b. Select the cell under the green check mark column to grant the specific right to the user or group.

For example, to grant the right to view and profile the data in the connection, click the cell in the green check mark column for the following *Specific Rights for Information Steward Connection*:

- *Profile/Rule permission*
- *View Data*
- *View Sample Data*

11. To assign project rights:
 - a. Expand the *Application* node and select *Information Steward Project*.
 - b. Click the cell under the green check mark column for each right that you want this user or group to have.

For example, to grant the right to create scorecards and approve rules in this project, click the cell under the green check mark icon column for the following *Specific Rights for Data Insight Connection*:

- *Approve Rule*
- *Manage Rule*

12. To assign profile or rule task rights:
 - a. Expand the *Application* node and select *Information Steward Profiler Task*.
 - b. Click the box in the *Override General Global* column and the click the cell under the green check mark icon column for each right that you want this user or group to have.

For example, to grant the right to schedule a profile task or rule task, click the *Override General Global* column and the green check mark column for the following *General Rights for Data Insight Profiler Task*:

- *Reschedule instances that the user owns*
- *Schedule document to run*
- *View document instances*

13. Click *OK* and verify that the list under the *Right Name* column displays the rights you just added.
14. Click *OK* and verify that the list of principals includes the name or names you just added.
15. Close the *User Security* window.

Related Information

[Data Insight predefined user groups \[page 59\]](#)

5.4.4.3 Restricting rights of a user group to one project

Whenever a project is created, all pre-defined Data Insight user groups are automatically added to its principal list. This feature facilitates user rights management when you want the same user or same group of users to have rights on all projects.

You might want to limit rights of a subset of users to only one project. For example, you might want to limit the Manage Scorecards right on the Human Resources project to only User A, and you want only User B to have the Manage Scorecards right on the Finance project.

To restrict the Manage Scorecards right to a specific user for each project:

1. Create User A and User B. For details, see [Creating users for Information Steward \[page 54\]](#).
2. Add User A and User B to the Data Insight Analyst user group. For more information, see [Adding users and user groups to Information Steward groups \[page 56\]](#).

This Data Insight Analyst user group has all of the rights of the Data Insight Scorecard Manager user group except the Manage Scorecard right, which you will grant to specific users within a project in subsequent steps.

3. Create the Human Resources project and the Finance project. For details, see [Creating a project \[page 160\]](#).
4. To grant User A the Manage Scorecard right on the Human Resources project:
 - a. Log in to the Central Management Console (CMC) with a user name that is a member of either the Administrator group or the Data Insight Administrator group.
 - b. At the CMC home page, click *Information Steward*, expand the *Data Insight* node, and expand the *Projects* node in the Tree panel.
 - c. Select the *Human Resources* project and click ► *Manage* ► *Security* ► *User Security* ►.

- d. On the *User Security* page, click *Add Principals*.
 - e. In the list of *Available users/groups*, select User A and click the > button to move the names to the *Selected users/groups* list.
 - f. Click *Add and Assign Security*.
 - g. Click the *Advanced* tab.
 - h. Click the *Add/Remove Rights* link.
 - i. Expand the *Application* node and click *Information Steward Project*.
 - j. Click the *Granted* column for the *Manage Scorecards* right under *Specific Rights for Data Insight Connection*.
 - k. Click *OK* and verify that the list under *Right Name* displays the *Manage Scorecards* right.
 - l. Click *OK* and close the *User Security* window.
5. Repeat steps 3a through 3j in the *Finance* project for User B.

Related Information

[Creating users for Information Steward \[page 54\]](#)

[Adding users and user groups to Information Steward groups \[page 56\]](#)

[Denying user rights to specific Data Insight objects \[page 68\]](#)

5.5 User rights in Metadata Management

The Metadata Management module of Information Steward contains the following objects that have object-specific rights that allow various actions on them.

- Metadata Management application through which users can view relationships (such as Same As, Impact, and Lineage) between integrator sources.
- Integrator Sources through which users collect metadata.
- Integrator Source Groups to subset the metadata when viewing relationships.

The CMS manages security information, such as user accounts, group memberships, and object rights that define user and group privileges. When a user attempts an action on a Metadata Management object, the CMS authorizes the action only after it verifies that the user's account or group membership has sufficient privileges.

Related Information

[Metadata Management pre-defined user groups \[page 73\]](#)

[Group rights for Metadata Management objects in Information Steward \[page 76\]](#)

[Group rights for Metadata Management folders and objects \[page 74\]](#)

[Assigning users to specific Metadata Management objects \[page 77\]](#)

5.5.1 Metadata Management pre-defined user groups

Information Steward provides the following Metadata Management user groups to enable you to change the rights for multiple users in one place (a group) instead of modifying the rights for each user account individually.

Metadata Management pre-defined user groups

Pre-defined User Group	Description
User	Users that can view metadata only in the Metadata Management tab of Information Steward.
Data steward	Users that have all the rights of a Metadata Management User, plus the following rights: <ul style="list-style-type: none">• Create and edit annotations• Define and approve Metapedia categories and terms
Administrator	Users that have all the rights of a Metadata Management Data Steward, plus the following rights: <ul style="list-style-type: none">• Create, edit, and delete Metadata Management integrator sources and source groups• Run and schedule Metadata Integrators• Create custom attributes and edit values of custom attributes• Configure, edit, delete, schedule, view history of Information Steward utilities• Manage Metapedia users and rights and re-assign terms to other groups

Related Information

[Information Steward pre-defined users and groups \[page 52\]](#)

5.5.2 Type-specific rights for Metadata Management objects

Type-specific rights affect only specific object types, such as integrator sources. The following topics describe type-specific rights for each Information Steward object in the CMC and Information Steward.

Related Information

[Group rights for Metadata Management folders and objects \[page 74\]](#)

[Group rights for Metadata Management objects in Information Steward \[page 76\]](#)

5.5.2.1 Group rights for Metadata Management folders and objects

Each pre-defined Metadata Management user group provides specific rights to the folders and objects in the Central Management Console (CMC).

Rights for Metadata Management folders in the CMC

CMC Folder or Object	Right Name	Description	Pre-defined User Groups		
			Metadata Management Administrator	Metadata Management Data Steward	Metadata Management User
Metadata Management folder	View objects	View Metadata Management folder	Yes	Yes	Yes
	Edit objects	Change limits for integrator source runs	Yes	No	No
	Modify Rights	Manage user security for Metadata Management	Yes	No	No
	Establish Relationship in Metadata Management	Establish relationships between objects using the object tray	Yes	Yes	Yes
Integrator Source folder	View objects	View integrator sources, their run history, their logs, and so on.	Yes	Yes	Yes
	Add objects to the folder	Configure new integrator sources	Yes	No	No
	Modify Rights	Modify the rights users have to integrator sources	Yes	No	No
Integrator Sources	View Instance	View history of metadata integrator runs	Yes	No	No
	Add objects to the folder	Create integrator source configurations	Yes	No	No
	Edit objects	Edit the properties of integrator sources	Yes	No	No
	Delete objects	<ul style="list-style-type: none"> Delete integrator sources Delete integrator source instances Delete the integrator source schedule 	Yes	No	No
	Schedule	Schedule and set runtime parameters for integrator source	Yes	No	No

CMC Folder or Object	Right Name	Description	Pre-defined User Groups		
			Metadata Management Administrator	Metadata Management Data Steward	Metadata Management User
	Reschedule	Change the schedule or run-time parameters for integrator source	Yes	No	No
	Modify Rights	Modify the rights users have to integrator sources.	Yes	No	No
Source Groups	View objects	View Source Groups and access properties	Yes	No	No
	Add objects to the folder	Create integrator source groups	Yes	No	No
	Edit objects	Edit integrator source groups	Yes	No	No
	Delete objects	Delete integrator source groups	Yes	No	No
	Modify Rights	Modify the rights users have to source groups	Yes	No	No
Metapedia Folder	View objects	View terms, categories, policy sets, and policy statements in Information Steward	Yes	Yes	Yes
	Edit	Modify terms, categories, policy sets, and policy statements	Yes	Yes	No
	Delete	Remove terms, categories, policy sets, and policy statements	Yes	Yes	No
	Add objects to the folder	Create terms, categories, policy sets, and policy statements	Yes	Yes	No
	Import/Export data	Import or Export the data in an Excel file	Yes	Yes	No
	Approve term	Approve or reject the term, policy set, or policy statements	Yes	Yes	No

CMC Folder or Object	Right Name	Description	Pre-defined User Groups		
			Metadata Management Administrator	Metadata Management Data Steward	Metadata Management User
	Manage term	Assign users to groups, assign rights to users, re-assign terms, policy sets, and policy statements to other groups	Yes	No	No

5.5.2.2 Group rights for Metadata Management objects in Information Steward

Each pre-defined Metadata Management user group provides specific rights on objects in Information Steward, as the following table shows.

Rights for Metadata Management objects in Information Steward

Right Name	Description	Pre-defined User Groups		
		Metadata Management Administrator	Metadata Management Data Steward	Metadata Management User
View objects	<ul style="list-style-type: none"> View all of the metadata objects and their relationships View the custom attributes and values View the Preferences page Search all metadata sources View Metadata Management lineage from View Lineage option on the Documents tab of BI Launch Pad. View Preferences page 	Yes	Yes	Yes
Edit objects	<ul style="list-style-type: none"> Update Preferences page Edit custom attributes and display order Edit custom attribute values Edit annotations Edit user-defined relationships between objects Update Preferences page 	Yes	Yes	No

Right Name	Description	Pre-defined User Groups		
		Metadata Management Administrator	Metadata Management Data Steward	Metadata Management User
Add objects	<ul style="list-style-type: none"> Create new custom attribute Associate custom attribute to an object type Update Preferences page Create custom attributes and display order Create custom attribute values Create annotations Create user-defined relationships between objects 	Yes	Yes	No
Delete objects	<ul style="list-style-type: none"> Delete custom attribute Create annotations Delete user-defined relationships between objects 	Yes	Yes	No

5.5.3 Assigning users to specific Metadata Management objects

By default, the pre-defined Metadata Management user groups are added to the access list of integrator sources and source groups when you create them.

You might want to allow only certain users to configure integrator sources, define source groups, or define Metapedia categories and terms or policy sets and statements. In these cases, you would add the user or group to the specific Metadata Management object's access list (instead of adding to a pre-defined Metadata Management user group).

- Log in to the Central Management Console (CMC) with a user name that is a member of either the Administrator group or the Metadata Management Administrator group.
- At the CMC home page, click [Information Steward](#).
- Expand the [Metadata Management](#) node in the Tree panel.
- Select the object type.
 - For an integrator source:
 - Select the [Integrator Sources](#) node in the Tree panel.
 - Select the integrator source name in the right panel.
 - For Metapedia categories and terms:
 - Select the [Metapedia](#) node in the Tree panel.
 - For a source group:
 - Expand the [Source Groups](#) node in the Tree panel.
 - Select the source group name in the right panel.
- Click [Manage](#) > [Security](#) > [User Security](#).

6. On the *User Security* page, click *Add Principals*.
7. In the list of *Available users/groups*:
 - a. Select the name of each user or group that you want to authorize to this integrator source or Metapedia or source group.

ⓘ Note

To select multiple names, hold down the Ctrl key when you click each name.

- b. Click the > button to move the names to the *Selected users/groups* list.
8. Click *Add and Assign Security*.
9. Click the *Advanced* tab.
10. Click the *Add/Remove Rights* link.
11. To assign integrator source rights:
 - a. Expand the *Application* node and select *Metadata Management Integrator configuration*.
 - b. Click the *Override General Global* column and the *Granted* (green check mark icon) column for each right that you want this user or group to have.

For example, to grant the right to schedule and view integrator instances, click the *Override General Global* column and the *Granted* column for the following *General Global Rights General* items:

- *Pause and resume document instances*
- *Schedule document to run*
- *View document instances*

12. To assign Metapedia rights:
 - a. Expand the *Content* node and select *Folder*.
 - b. Click the *Granted* column for each right that you want this user or group to have.

For example, to grant the right to create, edit, and delete Metapedia categories and terms, click the *Granted* column for the following *General Rights for Folder*:

- *Add objects to folder*
- *Delete objects*
- *Edit objects*
- *View objects*

13. To assign source group rights:
 - a. Expand the *Application* node and select *Metadata Management Source Group*.
 - b. Click the *Override General Global* column and the *Granted* (green check mark icon) column for each right that you want this user or group to have.

For example, to grant the right to schedule and view integrator instances, click the *Override General Global* column and the *Granted* column for the following *General Rights for Metadata Management Integrator configuration*:

- *Pause and resume document instances*
- *Schedule document to run*
- *View document instances*

14. Click *OK* and verify that the list under *Right Name* displays the rights you just added.
15. Click *OK* and verify that the list of principals includes the name or names you just added.
16. Close the *User Security* window.

5.6 User rights in Metapedia

In Metapedia, users create objects such as business terms and policy statements, that adhere to their business requirements and practices.

Users create categories for terms, and policy statements for policy sets. After authors have created a term or policy it is sent through an approval process.

The CMS manages security information, such as user accounts, group memberships, and object rights that define user and group privileges. When a user attempts an action on a Metapedia object, the CMS authorizes the action only after it verifies that the user's account or group membership has sufficient privileges.

Related Information

[Metapedia pre-defined user groups \[page 79\]](#)

[Type-specific rights for Metapedia objects \[page 80\]](#)

[Group rights for Metapedia folders \[page 80\]](#)

[Group rights for Metapedia terms in Information Steward \[page 81\]](#)

[Assigning users to specific Metapedia objects \[page 82\]](#)

5.6.1 Metapedia pre-defined user groups

SAP Information Steward provides the following Metapedia user groups to enable you to change the rights for multiple users in one place (a group) instead of modifying the rights for each user account individually.

Metapedia pre-defined user groups

Pre-defined user Group	Description
Metapedia User	Users that can view terms, categories, policy sets, and policy statements in the Metapedia tab of Information Steward.
Note A Metapedia author can optionally assign a Metapedia user as an observer when creating a term, policy set, or policy statement. An observer may receive email notifications when a term, policy set, or policy statement is approved or deleted.	
Metapedia Author	Users that have all the rights of a Metapedia User, plus the following rights: <ul style="list-style-type: none">• Create, edit and delete terms and categories.• Create, edit and delete policy sets and statements.• Import and export terms, categories, policy sets, and policy statements using Excel.• Receive email notification when a term, policy set, or policy statement is approved, rejected, or deleted.

Pre-defined user Group	Description
Metapedia Approver	<p>Users that have all the rights of a Metapedia Author, plus the rights to approve or reject terms, policy sets, and policy statements.</p> <p>Approvers receive all email notifications that a Metapedia Author receives, plus an email notification when a term, policy set, or policy statement is submitted for approval, is near or past a task due date, or when a policy set or statement has expired.</p>
Metapedia Administrator	<p>Users that have all the rights of a Metapedia Approver, plus the rights to manage terms, policy sets, and policy statements. For example, to re-assign terms, policy sets, and policy statements to other users or groups and to edit terms, policy sets, and policy statements that are not assigned to them as authors.</p>

5.6.2 Type-specific rights for Metapedia objects

Type-specific rights affect only certain object types, such as integrator sources or Metapedia objects. The following topics describe type-specific rights for each Information Steward object in the CMC and Information Steward.

Related Information

[Group rights for Metapedia folders \[page 80\]](#)

[Group rights for Metapedia terms in Information Steward \[page 81\]](#)

5.6.2.1 Group rights for Metapedia folders

Each pre-defined Metapedia user group provides specific rights on the folders and objects in the Central Management Console (CMC).

Rights for Metapedia folders in the CMC

CMC Folder or Object	Right Name	Description	Metapedia pre-defined User Groups			
			Administrator	Approver	Author	User
Metapedia Folder	View objects	View terms and categories in Information Steward	Yes	Yes	Yes	Yes
	Edit	Edit terms and categories	Yes	Yes	Yes	No
	Delete	Delete terms and categories	Yes	Yes	Yes	No

CMC Folder or Object	Right Name	Description	Metapedia pre-defined User Groups			
			Administrator	Approver	Author	User
	Add objects to the folder	Create terms and categories	Yes	Yes	Yes	No
	Import/Export data	Import or Export the data in an Excel file	Yes	Yes	Yes	No
	Approve term	Approve or reject the term	Yes	Yes	No	No
	Manage term	Re-assign terms to other users or groups and edit terms that are not assigned to them as authors	Yes	No	No	No

5.6.2.2 Group rights for Metapedia terms in Information Steward

Each pre-defined Metapedia user group provides specific rights on objects in Information Steward.

Rights for Metapedia terms in Information Steward

Right Name	Description	Pre-defined User Groups			
		Metapedia Administrator	Metapedia Approver	Metapedia Author	Metapedia User
View	<ul style="list-style-type: none"> View <i>Metapedia</i> tab View categories View terms View Preferences page Export to Excel 	Yes	Yes	Yes	Yes

Right Name	Description	Pre-defined User Groups			
		Metapedia Ad- ministrator	Metapedia Ap- prover	Metapedia Au- thor	Metapedia User
Add objects to the folder,	<ul style="list-style-type: none"> Create categories Create terms 	Yes	Yes	Yes	No
Edit objects,	<ul style="list-style-type: none"> Import to Excel 				
Delete ob- jects	<ul style="list-style-type: none"> Edit categories Edit terms Add terms to categories Relate terms Associate objects to a term Update Preferences page Delete related terms Delete associated objects Delete associated terms Delete term task Delete categories Delete terms 				
Approve terms	<ul style="list-style-type: none"> Approve terms Reject terms 	Yes	Yes	No	No

5.6.3 Assigning users to specific Metapedia objects

You might want to allow only certain users to define Metapedia categories and terms. In these cases, you would add the user or group to the specific Metapedia object's access list (instead of adding to a pre-defined Metapedia user group).

1. Log in to the Central Management Console (CMC) with a user name that is a member of either the Administrator group or the Metapedia Administrator group.
2. At the CMC home page, click *Information Steward*.
3. Expand the *Metadata Management* node in the Tree panel.
4. Select *Metapedia*.
5. Click **► Manage ► User Security ►**.
6. On the *User Security* window, click *Add Principals*.
7. In the list of *Available users/groups*, click *User List* or *Group List*.
 - a. Select the name of each user or group that you want to authorize to Metapedia group.

Note

To select multiple names, hold down the `Ctrl` key when you click each name.

- b. Click the **>** button to move the names to the *Selected users/groups* list.
8. Click *Add and Assign Security*.

9. Click the *Advanced* tab.
10. Click the *Add/Remove Rights* link.
11. To assign Metapedia rights:
 - a. Expand the *Content* node and select *Folder*.
 - b. Click the *Granted* column for each right that you want this user or group to have.

For example, to grant the right to create, edit, and delete Metapedia categories and terms, click the *Granted* column for the following *General Rights for Folder*:

 - *Add objects to folder*
 - *Delete objects*
 - *Edit objects*
 - *View objects*
12. To assign application rights:
 - a. Expand the *Application* node and select *Information Steward Metapedia*.
 - b. Click the *Granted* column for each right that you want this user or group to have.

For example, to grant the right to create, edit, and delete Metapedia categories and terms, click the *Granted* column for the following *Specific Rights for Information Steward Metapedia*:

 - *Approve Term*
 - *Export Data*
 - *Import data*
 - *Manage Term*
13. Click *OK* and verify that the list under *Right Name* displays the rights you just added.
14. Click *OK* and verify that the list of principals includes the name or names you just added.
15. Close the *User Security* window.

5.7 User rights in Cleansing Package Builder

The Cleansing Package Builder module of SAP Information Steward contains the following objects to which you control access.

- **Private cleansing packages:** Private cleansing packages are viewed or edited by the user who owns them and are listed under My Cleansing Packages. Private cleansing packages include those created by using the New Cleansing Package Wizard or by importing a published cleansing package.
- **Published cleansing packages:** Published cleansing packages are cleansing packages included with SAP Information Steward or cleansing packages which a data steward created and then published. Published cleansing packages are available to all users and can be used in an SAP Data Services Data Cleanse transform or imported and used as the basis for a new cleansing package.

Related Information

[Group rights for cleansing packages \[page 84\]](#)

[Managing users in Information Steward \[page 54\]](#)

5.7.1 Group rights for cleansing packages

The Administrator and the pre-defined Cleansing Package Builder User groups can perform specific actions on cleansing packages, as the following table shows.

Rights for Cleansing Package Builder

Rights	Description	Software in which this action is performed	Pre-defined User Groups	
			Administrator	User
Create	Create cleansing package.	Cleansing Package Builder	Yes	Yes
Edit	Change, delete, and rename your own private cleansing packages.	Cleansing Package Builder	Yes	Yes
Publish	Publish cleansing packages.	Cleansing Package Builder	Yes	Yes
View	Browse your own private cleansing packages and all published cleansing packages.	Cleansing Package Builder	Yes	Yes
Access the CMC	Log in to the CMC.	Central Management Console	Yes	No
Control Cleansing Package Builder server	Start and stop the Cleansing Package Builder Service.	Central Management Console	Yes	No
Set up users and groups	Create users and add users to groups.	Central Management Console	Yes	No
Reassign	Reassign ownership of private cleansing packages.	Central Management Console	Yes	No
Delete	Delete private and published cleansing packages.	Central Management Console	Yes	No
View all	See all private and published cleansing packages.	Central Management Console	Yes	No
Edit	Change description, owner, and state of cleansing packages.	Central Management Console	Yes	No

5.8 User rights for Information Steward administrative tasks

The following table describes the Information Steward actions in the [Applications](#) area of the CMC. To perform any of these Information Steward actions, a user must belong to one of the following groups:

- Administrator
- Data Insight Administrator
- Metadata Management Administrator

Actions for Information Steward in the CMC Applications area

Item	Action	Description
View Data Services Job Server	View	View list of job servers in the Information Steward job server group
Information Steward Utilities	View instance	View history of utility runs
	Create	Create utility instance
	Edit	Edit run-time parameters for utility
	Delete	Delete utility instance
	Schedule	Define schedule utility
	Reschedule	Change the schedule of utility
Information Steward Settings	Edit	Change configuration settings for <ul style="list-style-type: none">• Data Insight Profiling, Rules, or Performance• Data Review tasks• Search preferences
Information Steward repository	Edit	Update repository password

Related Information

[Job server group \[page 101\]](#)

[Utilities overview \[page 276\]](#)

[Information Steward configuration settings \[page 44\]](#)

[Viewing repository information \[page 85\]](#)

[Group rights for Data Insight folders and objects in the CMC \[page 61\]](#)

[Group rights for Metadata Management folders and objects \[page 74\]](#)

5.8.1 Viewing repository information

Access your SAP Information Steward repository to view or change settings, such as user name, password, and connection type.

You must have the appropriate credentials to change any repository settings. Additionally, most settings are “view only”.

1. Log into the Central Management Console (CMC) with a user name that is a member of the Administrator group.
2. Choose *Applications* from the navigation list at the top of the CMC Home page.
3. Select *Information Steward Application* from the *Application Name* list.
4. Select *Configure Repository* from the navigation pane at left.

Repository information appears at right. For most database types, the information is “view only”. Any changes that you can make require the database password.

If the database type is Oracle RAC, you can modify the connection string to add another server or tune parameters for failover.

ⓘ Note

If you have the appropriate credentials to access the repository database, you can change the user name and password. If you change the user name and password, you must restart the Web Application Server and Server Intelligence Agent (SIA).

5. After you're finished, select *Save*. To discard changes, select *Cancel*.

5.9 User rights in Match Review

The Match Review module of SAP Information Steward contains the following objects that have specific rights that allow various actions on them.

- Connections through which users view match results tables in a staging database to profile the data.
- Match review configurations
- Match review tasks through which users decide whether a record belongs in a match group or whether the master record should be reassigned.

ⓘ Note

In addition to rights to the connection, the Database Administrator must grant privileges on the tables to the user in the staging database.

5.9.1 Match Review pre-defined user groups

SAP Information Steward provides pre-defined Match Review user groups to facilitate the assignment of rights on connections, configurations, and tasks. These groups enable you to change the rights for multiple users in one place (a group) instead of modifying the rights for each user account individually.

The following table describes the pre-defined user groups in ascending order of rights.

Match Review User Groups

User group	Description
Data Review User	Users who are participants in match review as a reviewer or approver and have the following rights: Decides whether a record belongs in a match group. Decides whether the master record should be reassigned.

User group	Description
Data Review Configuration Manager	<p>Users that have all the rights of a Data Review User, plus the following rights:</p> <ul style="list-style-type: none"> Creates and manages match review configurations. Runs match review configurations immediately. Manages match review tasks as follows: <ul style="list-style-type: none"> Edits task settings. Cancel the task if the task is obsolete. Force completes the task if the remaining match review process can be bypassed.
Data Review Administrator	<p>Users that have all the above rights on Match Review objects, plus the following rights:</p> <ul style="list-style-type: none"> Creates connections for match review. Creates and manages match review configuration schedules.

5.9.2 Type-specific rights for Match Review

Type-specific rights affect only specific object types, such as Match Review configurations or tasks. The following topic describes type-specific rights for each Information Steward object in the CMC and SAP Information Steward.

Related Information

[Group rights for Data Review folder and Match Review objects in the CMC \[page 87\]](#)

5.9.2.1 Group rights for Data Review folder and Match Review objects in the CMC

Each pre-defined Data Review user group provides specific rights on the folders and objects in the Central Management Console (CMC), as the following table shows.

Rights for Data Review folders in the CMC

CMC Folder or Object	Right Name	Description	Pre-defined User Groups		
			Data Review Administrator	Data Review Data Steward	Data Review User
Connection folder	View objects	View connections in the Connections folder	Yes	Yes	Yes
	Add objects to folder	Create a connection	Yes	No	No

CMC Folder or Object	Right Name	Description	Pre-defined User Groups		
			Data Review Administrator	Data Review Data Steward	Data Review User
	Delete objects	Delete a connection	Yes	No	No
	Edit objects	Manage a connection	Yes	No	No
Data Review folder	View objects	View objects in the Data Review folder	Yes	Yes	Yes
	Add objects to folder	Create Match Review configurations	Yes	No	No
	Delete objects	Delete Match Review configurations	Yes	No	No
	Edit objects	Modify Match Review configurations	Yes	No	No
	Schedule	Create an instance of Match Review configuration	Yes	Yes	No
	Pause	Pause the instance of the Match Review configuration	Yes	No	No
	Reschedule	Reschedule the instance of the Match Review configuration	Yes	No	No
	View Instance	View the status of the Match Review configuration	Yes	Yes	No
	Delete instance	Delete an instance of the Match Review configuration	Yes	No	No
Match Review Configuration Folder	View objects	View objects in the Match Review Configuration folder	Yes	Yes	Yes
	Add objects to folder	Create Match Review configurations	Yes	Yes	No
	Delete objects	Delete Match Review configurations	Yes	Yes	No
	Edit objects	Modify Match Review configurations	Yes	Yes	No
	Schedule	Create an instance of the Match Review configuration	Yes	Yes	No
	Pause	Pause the instance of the Match Review configuration	Yes	No	No
	Reschedule	Reschedule the instance of the Match Review configuration	Yes	No	No
	View instance	View the status of a Match Review configuration	Yes	Yes	No
	Delete instance	Delete an instance of the Match Review configuration	Yes	No	No

CMC Folder or Object	Right Name	Description	Pre-defined User Groups		
			Data Review Administrator	Data Review Data Steward	Data Review User
	Manage task	Manage the Match Review tasks	Yes	Yes	No

6 Information Steward servers and services

Both the SAP Business Intelligence (BI) platform and the IPS (Information Platform Services) have a collection of services that run under a Server Intelligence Agent (SIA) on a host.

SAP Information Steward uses the EIM (Enterprise Information Management) Adaptive Processing Server, which has the following services:

- Cleansing Package Builder Auto-analysis Service
- Cleansing Package Builder Core Service
- Cleansing Package Builder Publishing Service
- Information Steward Administration Task Service
- Information Steward Data Review Service
- Data Cleansing Advisor Service
- Application Service
- Metadata Search Service
- Data Services Metadata Browsing Service
- Data Services View Data Service

Information Steward uses the Job Server that has the following services:

- Information Steward Task Scheduling Service
- Information Steward Integrator Scheduling Service
- Information Steward Data Review Task Scheduling Service

Information Steward uses the SAP Data Services Job Server, which executes Data Insight profiling tasks. You can create multiple Job Servers, each on a different computer, to use parallel execution for the profiling tasks.

[Verifying Information Steward servers are running \[page 91\]](#)

[Verifying Information Steward services \[page 91\]](#)

Verify that the features that you chose during SAP Information Steward installation have the corresponding services in the SAP Central Management Console (CMC).

[Server license measurement report \[page 93\]](#)

SAP Information Steward provides a utility for generating a report containing metrics about the license usage across all nodes where Information Steward is used.

[Configuring Metadata Browsing Service and View Data Service \[page 95\]](#)

SAP Information Steward uses the Metadata Browsing Service and the View Data Service to connect to and view data in profiling sources.

[E-mail notification configuration \[page 99\]](#)

For all SAP Information Steward e-mail notifications, configure the destination information in the EIM (Enterprise Information Management) Adaptive Job Server.

[Job server group \[page 101\]](#)

[Avoid exceptions in the Metadata Search service \[page 105\]](#)

To prevent the `TooManyClauses` or `OutOfMemory` exceptions when searching in Metadata Management or Metapedia, adjust settings in the EIM Adaptive Processing Server.

Related Information

[Servers and services \[page 12\]](#)

6.1 Verifying Information Steward servers are running

To verify that the Information Steward servers are running and enabled:

1. From the CMC Home page, select Servers from the *CMC Home* drop-down list.
2. Expand the *Service Categories* node and select *Enterprise Information Management Services*.
The list of servers in the right pane includes a *State* column that provides the status for each server in the list.
3. Verify that the following servers are “Running” and “Enabled”:
 - AdaptiveProcessingServer
 - JobServer
4. If a server is not running or enabled, do the following:
 - a. Select the server name from the list.
 - b. Click *Actions* from the toolbar and select *Start Server* or *Enable Server*.

For more information, see “Services” in the *Information Steward Administrator Guide*.

Task overview: [Information Steward servers and services \[page 90\]](#)

Related Information

[Verifying Information Steward services \[page 91\]](#)

[Server license measurement report \[page 93\]](#)

[Configuring Metadata Browsing Service and View Data Service \[page 95\]](#)

[E-mail notification configuration \[page 99\]](#)

[Job server group \[page 101\]](#)

[Avoid exceptions in the Metadata Search service \[page 105\]](#)

6.2 Verifying Information Steward services

Verify that the features that you chose during SAP Information Steward installation have the corresponding services in the SAP Central Management Console (CMC).

1. Open the CMC and select *Servers* from the *CMC Home* drop-down list.

2. Expand *Service Categories* in the left pane and select *Enterprise Information Management Services*. The right pane lists information about the following kinds of servers:
 - AdaptiveProcessingServer (EIM)
 - JobServer (IS)
3. Right-click the applicable server row and select *Stop Server* from the drop-down list. The *State* column changes to *Stopped*.
4. Right-click the applicable server row again and select *Select Services* from the drop-down list. The *Select Services: <computer_name>. <server_type>* window opens.
5. Verify that the list of services in the *Services of <computer_name>. <server_type>* column in the right pane is correct for each feature that was installed.

The table below lists each feature and the applicable service for each type of server.

Feature Installed	Services	
	EIM Adaptive Processing Server	Information Steward Adaptive Job Server (ISJobServer)
Information Steward Task Server	Information Steward Administrator Task Service	<ul style="list-style-type: none"> • Information Steward Task Scheduling Service • Information Steward Integrator Scheduling Service
Information Steward Data Review Server	N/A	Information Steward Data Review Task Scheduling Service
Metadata Search Service	Information Steward Metadata Search Service	N/A
Application Service	Information Steward Application Service	N/A
Cleansing Package Builder Service	<ul style="list-style-type: none"> • Cleansing Package Builder Auto-analysis Service • Cleansing Package Builder Core service • Cleansing Package Builder Publishing Service 	N/A
Information Steward Data Review Service	Information Steward Data Review Service	N/A
Data Cleansing Advisor Service	Data Cleansing Advisor Service	N/A
All metadata integrators	N/A	<ul style="list-style-type: none"> • Information Steward Integrator Scheduling Service • Information Steward Integrator Service

6. To add any missing service to the *Services of <computer_name>. <server_type>* list, select the service from the *Available services* list in the left pane and use the arrow button > to move it to the left-pane list.

Note

If you want to be able to use Information Steward Data Cleansing Advisor feature, you must move the Data Cleansing Advisor service from the list of available services to the list of services for *EIMAdaptiveProcessingServer*.

7. Click *OK* when you have finished.
8. Right-click the applicable server row and click *Start Server* or *Enable Server* from the drop-down list to start the server again.

Task overview: [Information Steward servers and services \[page 90\]](#)

Related Information

[Verifying Information Steward servers are running \[page 91\]](#)

[Server license measurement report \[page 93\]](#)

[Configuring Metadata Browsing Service and View Data Service \[page 95\]](#)

[E-mail notification configuration \[page 99\]](#)

[Job server group \[page 101\]](#)

[Avoid exceptions in the Metadata Search service \[page 105\]](#)

6.3 Server license measurement report

SAP Information Steward provides a utility for generating a report containing metrics about the license usage across all nodes where Information Steward is used.

The SAP Global License Auditing Service (GLAS) performs customer audits on a periodic basis. Generate the server license measurement report when the GLAS requests it. You can generate the report at any time for your own use.

The utility name is `LicenseMeasurementUtility.bat` (`.sh` for Linux). The utility collects information about the following servers:

- All Information Steward Job Servers registered in the Central Management Server (CMS)
- All the job servers hosted on the Data Services Server Manager for running Information Steward jobs

The generated report is an XML file that contains license information and CPU metrics for each job server. The report contains the following information about each job server:

- Server name
- IP address
- Port number
- License key code
- License type

- License expiration date
- CPU type
- Number of cores

Note

The license measurement utility is available on Windows and Linux.

[Generating a license measurement report \[page 94\]](#)

Generate the license measurement report upon request by the SAP Global License Auditing Services (GLAS).

Parent topic: [Information Steward servers and services \[page 90\]](#)

Related Information

[Verifying Information Steward servers are running \[page 91\]](#)

[Verifying Information Steward services \[page 91\]](#)

[Configuring Metadata Browsing Service and View Data Service \[page 95\]](#)

[E-mail notification configuration \[page 99\]](#)

[Job server group \[page 101\]](#)

[Avoid exceptions in the Metadata Search service \[page 105\]](#)

6.3.1 Generating a license measurement report

Generate the license measurement report upon request by the SAP Global License Auditing Services (GLAS).

Ensure that all applicable repository connections are active and that the SAP Information Steward Job Server is started.

Note

If the repository and Job Server are not connected, active, and started, the utility does not collect metrics about the repository or Job Server.

Administrators with access permission to the Information Steward Job Server can generate the license measurement report.

1. For Windows: From a command line, call the utility from `<INSTALL_DIR>\InformationSteward\MM\bin\LicenseMeasurementUtility.bat`. For Linux, call the utility `<INSTALL_DIR>/InformationSteward/MM/bin/LicenseMeasurementUtility.sh`.
2. Enter the following arguments:

License Measurement Utility arguments

Argument	Description
boeUser	Central Management Server (CMS) administrator name.
boePassword	Password related to the value in boeUser.
boeServer	Central Management Server (CMS) host name.
outputFilePath	Output file location and output file name.
password	Information Steward repository password.

Note

The software adds the extension `_metrics.xml` to the output file name.

Example

```
For Windows: C:\Program Files (x86)\SAP
BusinessObjects\InformationSteward\MM\bin>LicenseMeasurementUtility.bat
boeUser <user_name> boePassword <password> boeServer <cms_host_name>
outputFilePath <"outputlocation_filename"> password <repository_password>
```

The utility obtains all other Information Steward repository details from `%MM_CORE%\config\InstallUtilty.properties`.

3. Click . The utility generates and outputs the report to the named location and file.

6.4 Configuring Metadata Browsing Service and View Data Service

SAP Information Steward uses the Metadata Browsing Service and the View Data Service to connect to and view data in profiling sources.

When you install SAP Data Services, the installation configures the two services with default settings, and lists them under the server `EIMAdaptiveProcessingServer`. You can change the configuration settings to more effectively integrate Information Steward with your hardware, software, and network configurations.

To change the configuration settings for Metadata Browsing Service and View Data Service, perform the following steps in the Central Management Console (CMC):

1. Open the [Servers](#) management area of the CMC.
2. Expand [Service Categories](#) node the tree panel and select [Enterprise Information Management Services](#).
3. Double-click `<computername.> EIMAdaptiveProcessingServer` in the list in the right pane.

The [Properties](#) page opens.

4. Double-click the service to change.
5. Change the configuration as applicable and select *Save* or *Save & Close*.

Note

Not all changes occur immediately. If a setting can't change immediately, the *Properties* page displays both the current setting (in red text) and the updated setting. When you return to the *Servers* management area, the server is labeled as *Stale*. When you restart the server, it uses the updated settings from the *Properties* page and removes the *Stale* flag from the server.

[Metadata Browsing Service configuration parameters \[page 96\]](#)

You can change the following properties of the Metadata Browsing Service.

[View Data Services configuration parameters \[page 98\]](#)

You can change the following properties of the View Data Service.

Task overview: [Information Steward servers and services \[page 90\]](#)

Related Information

[Verifying Information Steward servers are running \[page 91\]](#)

[Verifying Information Steward services \[page 91\]](#)

[Server license measurement report \[page 93\]](#)

[E-mail notification configuration \[page 99\]](#)

[Job server group \[page 101\]](#)

[Avoid exceptions in the Metadata Search service \[page 105\]](#)

6.4.1 Metadata Browsing Service configuration parameters

You can change the following properties of the Metadata Browsing Service.

Server Configuration Parameter	Description	Possible Values
Service Name	Name of the service configuration.	Alphanumeric string with a maximum length of 64. The Service Name cannot contain any spaces. Default value: MetadataBrowsingService
Maximum Data Source Connections	Maximum number of data source connections that can be opened at any time under a service instance.	integer. Default value: 200

Server Configuration Parameter	Description	Possible Values
Retry attempts to launch Service Provider	Maximum number of attempts to launch a new service provider when there is contention to access a shared service provider.	Default value: 1
Stateful Connection Timeout (seconds)	Maximum duration in which a stateful connection is open. Stateful connections include SAP Applications and SAP BW Source.	Default value: 1200
Stateless Connection Timeout (seconds)	Maximum duration in which a stateless connection is open. Stateless connections include all relational database sources.	Default value: 1200
Recycle Threshold	Maximum number of requests that the service processes before the Data Services back end engine is recycled to free memory that was allocated for metadata browsing.	Default value: 50000
Log Level	<p>Level of logging of trace messages to the log file.</p> <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 10px; margin: 10px 0;"> <p>Note</p> <p>If there is more than one instance of Metadata Browsing Service configured in the CMS, the same level of information is collected from all instances. The log level defined for the first running service is the level used.</p> </div>	<p>Information Steward logs:</p> <ul style="list-style-type: none"> • None: Logging disabled. • Info: Logging disabled. (same as None) • Finer: All traces, requests, and responses. <p>Data Services logs:</p> <ul style="list-style-type: none"> • None: Logging disabled. • Info: All traces. • Finer: All traces, requests, and responses. <p>Default level is Info.</p>
Collect Connection Statistics	Enable or disable the collection of statistical information for each open connection.	Default is enabled.
Listener Port	<p>Port number used to communicate with the Data Services back end engine.</p> <p>If you change the port number, you must restart the EIMAdaptiveProcessingServer for the change to take effect.</p>	<p>Four-digit port number that is not currently in use.</p> <p>Default value: 4010</p>
JMX Connector Port	<p>Port number used for the JMX Connector.</p> <p>If you change the port number, you must restart the EIMAdaptiveProcessingServer for the change to take effect.</p>	<p>Four-digit port number that is not currently in use.</p> <p>Default value: 4011</p>

6.4.2 View Data Services configuration parameters

You can change the following properties of the View Data Service.

Server Configuration Parameter	Description	Possible Values
Service Name	Name of the service configuration.	Alphanumeric string with a maximum length of 64. The Service Name cannot contain any spaces. Default value: ViewData
Listener Port	Port number used to communicate with the Data Services backend engine. If you change the port number, you must restart the EIMAdaptiveProcessingServer for the change to take effect.	Four-digit integer. Default value: 4012
JMX Connector Port	Port number used for the JMX Connector. If you change the port number, you must restart the EIMAdaptiveProcessingServer for the change to take effect.	Four-digit integer. Default value: 4013
Batch Size (kilobytes)	Size of the data to be stored in a view data response.	Minimum value: 1000 Maximum value: 5000 Default value: 1000
Minimum Shared Service Providers	Minimum number of shared Data Services backend engines that need to be launched at the startup time of the service.	Default value: 1
Maximum Shared Service Providers	Maximum number of shared Data Services backend engines that can be launched during the time to service the view data requests.	Default value: 5
Maximum Dedicated Service Providers	Maximum number of dedicated Data Services backend engines that can be launched at any instant of time.	Default value: 10
Recycle Threshold	Maximum number of requests that will be processed by a service before the Data Services backend engine is recycled to free memory that was allocated for viewing data.	Any integer. Default value: 200
Number of attempts to launch service provider	Number of attempts to be made to try launching the Data Services backend engine instance.	Default value: 1
Maximum idle time for shared service provider (minutes)	Maximum number of minutes that a Data Services backend engine can remain without processing any requests. After this time is exceeded, the Data Services backend engine is shut down.	Default value: 120

Server Configuration Parameter	Description	Possible Values
Log Level	Level of logging of trace messages to the log file. <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 10px; margin-top: 10px;"> <p>Note</p> <p>If there is more than one instance of View Data Service configured in the CMS, the same level of information is collected from all instances. The log level defined for the first running service is the level used.</p> </div>	Information Steward logs: <ul style="list-style-type: none"> • None: Logging disabled. • Info: Logging disabled. (same as None) • Finer: All traces, requests, and responses. Data Services logs: <ul style="list-style-type: none"> • None: Logging disabled. • Info: All traces. • Finer: All traces, requests, and responses. Default level is Info.

6.5 E-mail notification configuration

For all SAP Information Steward e-mail notifications, configure the destination information in the EIM (Enterprise Information Management) Adaptive Job Server.

Set Information Steward to send automatic e-mail notifications for events such as when a specific rule threshold is met, when a rule score meets a specific sensitivity setting, or when a task completes. Before automatic e-mails function correctly in any module, you must configure the EIM Adaptive Processing Server with universal e-mail settings.

Note

You must have the required permissions to make e-mail settings in the EIM Adaptive Job Server.

Parent topic: [Information Steward servers and services \[page 90\]](#)

Related Information

[Verifying Information Steward servers are running \[page 91\]](#)

[Verifying Information Steward services \[page 91\]](#)

[Server license measurement report \[page 93\]](#)

[Configuring Metadata Browsing Service and View Data Service \[page 95\]](#)

[Job server group \[page 101\]](#)

[Avoid exceptions in the Metadata Search service \[page 105\]](#)

6.5.1 Setting up e-mail notification configuration

Configure e-mail notifications in the Enterprise Information Management (EIM) Adaptive Job Server in the Central Management Console (CMC).

Ensure that you have the required permissions, then perform the following steps to configure SAP Information Steward for e-mail notifications.

1. Log in to the CMC and select *Servers* from the *Organize* column in the *CMC Home* page.
2. Click the *Servers List* node at left.

A list of all existing servers appears at right.

3. Select the **Adaptive Job Server** from the list at right to highlight the row.
4. Select **Manage > Properties** from the tools menu at the top.

The *Properties* dialog box opens.

5. Select *Destination* from the navigation list at left.
6. Select *Email* from the *Destination* dropdown menu and click *Add*.

If an e-mail destination already exists, *Email* is not included in the *Destination* dropdown list. If necessary, edit the existing e-mail information by selecting *Email* from the existing destinations list.

A list of e-mail options appears.

7. Complete the e-mail options as described in the following table.

E-mail option descriptions

Option	Description
<i>Domain Name</i>	Required. Specifies the fully qualified domain of the SMTP server.
<i>Host</i>	Required. Specifies the e-mail server IP address.
<i>Port</i>	Required. Specifies the port number on which the SMTP server is listening. The default SMTP port is 25.
<i>Authentication</i>	Specifies the authentication that the SMTP server uses to authenticate the client to the server. <ul style="list-style-type: none"> • <i>Plain</i>: Uses the AUTH PLAIN mechanism to authenticate the client to the server. • <i>Login</i>: Uses the AUTH LOGIN mechanism to authenticate the client to the server. • <i>None</i>: Does not use an authentication method. Use when the SMTP server does not require the Plain or Login methods. <i>None</i> is the default setting.
<i>User Name</i>	Specifies the name of the user who has permission to send e-mail and attachments through the SMTP server.
<i>Password</i>	Specifies the password for the specified <i>User Name</i> .

Option	Description
<i>From</i>	Required. Specifies the e-mail address of the sender. Authorized users may override this setting later when they set e-mail notification settings for a specific event.
<i>To</i>	Specifies the e-mail address of the e-mail alert subscriber. Authorized users may override this setting later when they set e-mail notification settings for a specific event.
<i>CC</i>	Optional. Specifies an e-mail for recipients of a carbon copy of the e-mail alerts.
<i>Subject</i>	Specifies a subject heading as the default heading for all system alert e-mails. Either type a subject or select a placeholder from the <i>Add Placeholder</i> dropdown list.
<i>Message</i>	Specifies an e-mail message as the default message for all system alert e-mails. SAP recommends that you select an option from the <i>Add Placeholder</i> dropdown list.
<i>Deliver Document(s) as Attachment</i>	Specifies to send the alert as an attachment.

Note

Consider entering the `<%SI_EMAIL_ADDRESS%>` placeholder. If you specify a specific e-mail address or recipient now, Information Steward sends all system alerts to that address by default.

Note

Enter values for *Domain Name*, *Host* and *From*. Additionally, the *Port* entry must be numeric. If the required elements are not set, or you entered a non-numeric entry for Port, Information Steward presents a warning the next time you log in to Information Steward.

- Click *Save & Close*.

6.6 Job server group

The SAP Data Services Job Server performs the profile and rule tasks on data in Data Insight connections. The Information Steward Task Server sends Data Insight profile tasks to the Data Services Job Server which partitions the data and uses parallel processing to deliver high data throughput and scalability.

This section contains the tasks to manage Data Services Job Servers for Information Steward.

[Data Insight can use Data Services Job Servers \[page 102\]](#)

For processing efficiency, use any additional Job Servers for SAP Information Steward processes, including creating data cleansing solutions.

[Configuring a Data Services Job Server for Data Insight \[page 103\]](#)

Before Data Insight can use an SAP Data Services Job Server, add the Job Server to a predefined SAP Information Steward Job Server group.

[Displaying job servers for Information Steward \[page 104\]](#)

View a list of Job Servers in the Central Management Console (CMC) to find out what Job Servers are associated with SAP Information Steward.

[Removing a Job Server from the group \[page 104\]](#)

When you no longer use an SAP Data Services Job Server in the Job Server group, remove the Job Server association with the SAP Information Steward repository.

Parent topic: [Information Steward servers and services \[page 90\]](#)

Related Information

[Verifying Information Steward servers are running \[page 91\]](#)

[Verifying Information Steward services \[page 91\]](#)

[Server license measurement report \[page 93\]](#)

[Configuring Metadata Browsing Service and View Data Service \[page 95\]](#)

[E-mail notification configuration \[page 99\]](#)

[Avoid exceptions in the Metadata Search service \[page 105\]](#)

[Configuring a Data Services Job Server for Data Insight \[page 103\]](#)

[Data Insight can use Data Services Job Servers \[page 102\]](#)

[Displaying job servers for Information Steward \[page 104\]](#)

[Removing a Job Server from the group \[page 104\]](#)

6.6.1 Data Insight can use Data Services Job Servers

For processing efficiency, use any additional Job Servers for SAP Information Steward processes, including creating data cleansing solutions.

An SAP Data Services administrator can install multiple Job Servers on separate computers based on the processing environment. Use the additional Job Servers to run Data Insight processes, such as profiles, rules, and data cleansing solutions.

Note

The Job Servers don't have to be installed on computers that have SAP Information Steward installed.

Before you can use any additional Job Server for profile and rule tasks, add the job server to the predefined Information Steward job server group.

Related Information

[Factors that influence performance and sizing \[page 292\]](#)

6.6.2 Configuring a Data Services Job Server for Data Insight

Before Data Insight can use an SAP Data Services Job Server, add the Job Server to a predefined SAP Information Steward Job Server group.

Use additional Data Services Job Servers to run Data Insight profile and rule tasks or to run Data Cleansing Advisor. To make the Job Servers available to Data Insight, associate each Job Server with the Information Steward repository. The repository association adds the job server to a predefined Information Steward job server group.

Note

You must perform the association on the computer where the Job Server is installed.

To configure a Job Server and associate it with the Information Steward repository, perform the following steps on the computer where the Job Server is installed:

1. Open the Windows Start menu and right-click the Data Services Server Manager application.
2. Select **More** > **Run as Administrator**.
3. In the *Job Server* tab, select *Configuration Editor*.

The *Job Server Configuration Editor* dialog box opens.

4. Select *Add*.

The *Job Server Properties* dialog box opens.

5. Type a name in *Job Server name*.

Note

Make sure that the Job Server name that you enter is different from the Data Services Job Server name.

6. Select *Add* in the *Associated Repositories* group and complete the *Repository Information*.

Enter the following information about the Information steward repository:

- *Database type*
- *Database Server name*
- *Database name*
- *Username*
- *Password*

7. Select *Apply* and select *OK*.
8. Verify that the Job Server displays in the *Job Server Configuration Editor* dialog box
9. Select *OK*.
10. Select *Close and Restart* to restart the job server with the updated configurations.
11. Select *OK* when you receive a message saying that Data Services Service restarts.

For more information about using the Data Services Server Manager, see “Server management” in the *Data Services Administrator Guide*.

6.6.3 Displaying job servers for Information Steward

View a list of Job Servers in the Central Management Console (CMC) to find out what Job Servers are associated with SAP Information Steward.

Job Servers appear in the list of Job Servers for Information Steward only after you set them up with a repository. Associate a Job Server with the Information Steward repository database using the SAP Data Services Server Manager.

1. Log in to the CMC with a user name that belongs to the Metadata Management Administrator group or the Administrator group.
2. Select *Applications* from the *CMC Home* dropdown list.
3. Select *Information Steward Application* in the *Application Name* list.
4. Click *View Data Services Job Server* from the navigation pane at left.

A list of Job Servers appears under *List of Data Service Job Servers for Information Steward* in the main pane. The CMC displays the following information:

- The name of each Data Services Job Server that is associated with the Information Steward repository.
- The computer name and port number for each Job Server, listed in the format `<computer name>:<port number>`.

Related Information

[Configuring a Data Services Job Server for Data Insight \[page 103\]](#)

[Data Insight can use Data Services Job Servers \[page 102\]](#)

6.6.4 Removing a Job Server from the group

When you no longer use an SAP Data Services Job Server in the Job Server group, remove the Job Server association with the SAP Information Steward repository.

To remove a Job Server association from the Information Steward repository and Job Server group, perform the following steps on the computer in which the Job Server resides:

1. Open the Windows Start menu and find the Data Services Server Manager application.
2. Select **More** **Run as Administrator**.
3. In the *Job Server* tab, select *Configuration Editor*.

The *Job Server Configuration Editor* dialog box opens.

4. Choose the name of the Job Server to delete and select *Delete*.
5. Select *Yes* for the question about removing persistent cache tables.
6. View the list of associated repositories in *Job Server Properties* under the *Associated Repositories* section.
7. Choose the Information Steward repository and select *Delete*.
8. Enter the password of your Information Steward repository in *Repository Information*, and select *Apply*.

9. Select *OK*.

The *Job Server Configuration Editor* dialog box opens.

10. Select *OK*.

11. Select *Close and Restart*.

12. Select *OK*.

The system restarts the Data Services Job Server with the updated configuration.

6.7 Avoid exceptions in the Metadata Search service

To prevent the `TooManyClauses` or `OutOfMemory` exceptions when searching in Metadata Management or Metapedia, adjust settings in the EIM Adaptive Processing Server.

The SAP Information Steward Metadata Search service uses the EIM Adaptive Processing Server and the Apache Lucene search engine to conduct searches. The search engine adds current search results to the existing search index. Consequentially, if during a metadata search the existing search index exceeds set limits, the Metadata Search service stops the search, and the Lucene search engine issues the following exceptions:

- `TooManyClauses`: When your search term causes the Lucene search engine to expand the search and attempt to add more than the permitted number of clauses. By default, the permitted number of clauses is 1024.
- `OutOfMemory`: When your search returns so many results that it maximizes your system memory.

Avoid these exceptions by strategically setting the runtime parameters described in the following table.

Parameter	Description
<i>Max Clause Count</i>	<p>Increases the maximum number of clauses permitted by the search engine. The default maximum number of clauses is 1024. Helps to avoid the <code>TooManyClauses</code> exception.</p> <p>-1: Allows the maximum number of search clauses, which is 1024. The default setting is -1. Setting this option to 0 has the same result.</p> <p>Set to greater than 1024: Lucene search engine increases the maximum search clauses by the set number.</p> <div data-bbox="804 1626 1394 1778"><p>❖ Example</p><p>When you set <i>Max Clause Count</i> to 2048, you double the maximum number of search clauses.</p></div>

Parameter	Description
<i>Max Search Result</i>	Limits the number of returned results for the search term to avoid the <code>OutOfMemory</code> exception.

❖ Example

When you set *Max Search Result* to 10, the Lucene search engine returns only the top 10 search results.

Set the two runtime parameters in the Central Management Console for the EIM Adaptive Processing Server.

[Setting runtime parameters for the Metadata Search service \[page 106\]](#)

If the Metadata Search service for Metadata Management or Metapedia stops the search and issues an exception, adjust specific runtime parameter settings and try the search again.

Parent topic: [Information Steward servers and services \[page 90\]](#)

Related Information

[Verifying Information Steward servers are running \[page 91\]](#)

[Verifying Information Steward services \[page 91\]](#)

[Server license measurement report \[page 93\]](#)

[Configuring Metadata Browsing Service and View Data Service \[page 95\]](#)

[E-mail notification configuration \[page 99\]](#)

[Job server group \[page 101\]](#)

6.7.1 Setting runtime parameters for the Metadata Search service

If the Metadata Search service for Metadata Management or Metapedia stops the search and issues an exception, adjust specific runtime parameter settings and try the search again.

To avoid the `TooManyClauses` and `OutOfMemory` search exceptions, adjust the runtime parameters *Max Clause Count* and *Max Search Result* by performing the following steps:

1. Log in to the Central Management Console and open the [Servers](#) page.
2. Expand the [Servers List](#) node at left.
3. Double-click the EIM Adaptive Processing Server at right.

The [Properties](#) page opens.

4. Scroll to the [Information Steward Metadata Search Service](#) group.
5. Set one or both of the following runtime parameters based on the exception to eliminate:

Parameter	Value	Exception
<i>Max Clause Count</i>	<ul style="list-style-type: none"> -1: Allows the maximum number of search clauses, which is 1024. -1 is the default setting. Setting to 0 has the same result. More than 1024: Increases the maximum number of clauses to the set amount. <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 10px; margin-top: 10px;"> <p>→ Tip</p> <p>Set to multiples of 1024, such as 2048, which doubles the maximum number of search clauses. Set to 3072 to triple the maximum number of search clauses.</p> </div>	TooManyClauses
<i>Max Search Result</i>	Positive integer greater than 0: Number of returned results for the search term.	OutOfMemory

6. Select *Save & Close*.

Task overview: [Avoid exceptions in the Metadata Search service \[page 105\]](#)

7 Repository Management

7.1 Viewing repository information

Access your SAP Information Steward repository to view or change settings, such as user name, password, and connection type.

You must have the appropriate credentials to change any repository settings. Additionally, most settings are “view only”.

1. Log into the Central Management Console (CMC) with a user name that is a member of the Administrator group.
2. Choose *Applications* from the navigation list at the top of the CMC Home page.
3. Select *Information Steward Application* from the *Application Name* list.
4. Select *Configure Repository* from the navigation pane at left.

Repository information appears at right. For most database types, the information is “view only”. Any changes that you can make require the database password.

If the database type is Oracle RAC, you can modify the connection string to add another server or tune parameters for failover.

Note

If you have the appropriate credentials to access the repository database, you can change the user name and password. If you change the user name and password, you must restart the Web Application Server and Server Intelligence Agent (SIA).

5. After you're finished, select *Save*. To discard changes, select *Cancel*.

7.2 Viewing Data Cleansing Advisor repository information

The Data Cleansing Advisor stores staged data in a local repository. The staged data allows users to preview cleansing and matching changes to the data without changing their actual data.

When you access the Data Cleansing Advisor repository, you can perform the following tasks:

- Edit the Server port number for the Data Cleansing Advisor local repository.
 - View the Server name. The Server name is the machine name where SAP Information Steward is installed.
1. Log into the Central Management Console (CMC) with a user name that is a member of the Administrator group.
 2. Choose *Applications* from the navigation list at the top of the CMC Home page.
 3. Select *Information Steward Application* in the *Application Name* list.

4. Select [Configure Data Cleansing Advisor](#) from the navigation list at left.
The connection information for the Data Cleansing Advisor local repository is defined when Data Cleansing Advisor is first accessed. You can view the server name here, but you can't change it.
5. To change the server port number, enter the new value in the [Server Port](#) field.

If you change the server port number, you must restart the Enterprise Information Management Adaptive Processing Server (EIM APS). Stopping and restarting the EIM APS stops all services running under EIM APS. Therefore, you must notify all users to stop using the EIM APS, and restart EIM APS only when it isn't in use.

To restart EIM APS in the CMC:

1. Select [Servers](#) from the navigation list at the top of the CMC Home page.
2. Expand the [Service Categories](#) node at left.
3. Select [Enterprise Information Management Services](#) to list the server names at right.
4. Select the [EIM Adaptive Processing Server](#) from the list of servers at right.
5. Use the toolbar buttons to stop and then restart it.

7.3 Enabling Windows authentication and SSL/TLS for Microsoft SQL Server repositories

During installation, you can enable Windows authentication instead of Microsoft SQL Server authentication using the installer, command line, or response file. However, after installation you can enable Windows authentication in the Central Management Console (CMC).

To create a new repository with Windows authentication enabled, set the `ISRepositoryUtility` parameter `repoMSSQLWinAuth` to **true**. For details, see [Creating a new repository \[page 113\]](#).

To change the authentication of your Microsoft SQL Server repository to Windows authentication, perform the following steps:

1. Log into the Central Management Console (CMC) with a user name that is a member of the Administrator group.
2. Choose [Applications](#) from the navigation list at the top of the CMC Home page.
3. Select [Information Steward Application](#) from the Application Name list.
4. Select [Configure Repository](#).
5. Select the [Windows Authentication](#) checkbox.
6. Select [Enable SSL Encryption](#).
7. **Optional:** Select [Verify Server Certificate](#).

Selecting [Verify Server Certificate](#) allows SAP Information Steward to use the client certificate to verify the host name. When you select [Verify Server Certificate](#), also complete the options described in the following table:

Option	Description
Host Name in Certificate	Name used to verify the server's identity.

Option	Description
SSL TrustStore	Absolute location of the SSL TrustStore file.
SSL TrustStore Type	SSL TrustStore file format.
SSL TrustStore Password	Password to access the private key from the TrustStore file.
Unix Trust Store	<div style="border: 1px solid #ccc; padding: 5px; margin-bottom: 5px;"> <p>Note</p> <p>Enabled when you select Enable SSL Encryption.</p> </div> <p>Specifies the location and file name of the trust store.</p> <p>UNIX/Linux only.</p>
Unix Trust Store Password	<div style="border: 1px solid #ccc; padding: 5px; margin-bottom: 5px;"> <p>Note</p> <p>Enabled when you select Enable SSL Encryption.</p> </div> <p>Specifies the password associated with the specified trust store. If you don't password-protect the specified trust store, leave this option blank.</p> <p>UNIX/Linux only.</p>
Database User	Specifies the name of the user or owner of the database or data source.
Database Password	Specifies the password of the user for the database or data source.

8. Select [Save](#).
9. Open the SAP Business Intelligence Central Configuration Manager (CCM) from the Windows Start menu.
10. Select the Web application server, such as Apache Tomcat, and select the [Stop](#) icon.
11. Select the Server Intelligence Agent (SIA) and select the [Stop](#) icon.
12. Select the Web application server and select the [Properties](#) icon.
13. Uncheck the System Account checkbox to disable the System Account.
14. Enter the Windows domain user name and password and select [Apply](#) and [OK](#).
15. Reset the system account settings for the SIA in the same manner.
16. Ensure the associated Data Services Windows service is also configured to use the same domain account configured for the Web application server:
 - a. Select Windows Administrative Tools from the Windows Start menu.
 - b. Open Services.
 - c. Stop the SAP Data Services service.
 - d. Right-click SAP Data Services and select Properties.
 - e. Open the Log On tab.
 - f. Confirm that the account and password are the same domain account configured for the Web application server. Change the settings if necessary.
 - g. Select [Apply](#) and [OK](#).
 - h. Restart the SAP Data Services service and close the Services page.

- Restart the Web application server and SIA, and close the Central Configuration Manager (CCM).

7.4 Configuring SSL/TLS for SAP HANA repository

Include SSL/TLS protection when your SAP Information Steward repository connects to an SAP HANA database.

Before you perform this task, ensure that you've included the bundled JDBC driver when you installed SAP Information Steward, or that you download and install a supported version of the JDBC driver.

To configure SSL/TLS protection for your SAP Information Steward repository that uses an SAP HANA database, perform the following steps:

- Log into the Central Management Console (CMC) and open the SAP Information Steward page.
- Choose *Applications* from the navigation list at the top of the CMC Home page.
- Select *Information Steward Application* from the *Application Name* list.
- Select *Configure Repository* from the navigation pane at left.

Repository information appears at right.

- Select *Enable SSL Encryption*.
- Optional:** Select *Verify Server Certificate*.

Selecting *Verify Server Certificate* allows SAP Information Steward to use the client certificate to verify the host name. When you select *Verify Server Certificate*, also complete the options described in the following table:

Option	Description
<i>Host Name in Certificate</i>	Name used to verify the server's identity.
<i>SSL KeyStore</i>	Absolute location of the Java SSL/TLS KeyStore file.
<i>SSL KeyStore Type</i>	Type of the Java SSL/TLS KeyStore certificate file.
<i>SSL KeyStore Password</i>	Password used to access the private key from the Java SSL/TLS KeyStore file.
<i>SSL TrustStore</i>	Absolute location of the Java TrustStore file.
<i>SSL TrustStore Type</i>	Java TrustStore file format.
<i>SSL TrustStore Password</i>	Password to access the private key from the TrustStore file.
<i>Common Crypto SSL Keystore</i>	Name of the PSE file that you configured in SECUDIR.

The following configuration options are only applicable if you're working on UNIX platforms.

<i>Unix TrustStore</i>	The SSL truststore file to use for establishing a secure connection from a UNIX client to a repository database.
<i>Unix TrustStore Password</i>	Enter the password to access the certificate from the UNIX truststore file.

- Select *Save*.

7.5 Backing up your repository

We recommend that you create a backup policy for your SAP Information Steward repository based on the frequency of the following tasks:

- Add or change Data Insight connections, projects, views, file formats, rules, rule bindings, or scorecards
- Add or change Metadata Management integrator sources or custom attributes
- Define or change Metapedia terms and categories
- Change the include source with failed data setting in an existing rule task

Use the backup utility of your relational database management system that you used to create your Information Steward repository.

7.6 Repository utility overview

Normally, you create the Information Steward repository during installation. However, you can use the IS Repository utility (`ISRepositoryUtility`) in the following situations:

- You want to recover your repository from a backup.
- You want to reset the contents of your repository, such as when moving from a development to production environment or when you want to change the database type.

⚠ Caution

If you reset your repository by using the `create` mode of the Repository utility, you will lose all your existing contents.

- You want to manually upgrade your repository from a prior version of Information Steward to the current version.

Run the Repository utility from the command line.

On Windows platforms:

- `ISRepositoryUtility.bat` is installed to the `<SAP_BusinessObjects>\InformationSteward\MM\bin` directory by default.
- The Repository utility uses the parameter values in the `InstallUtility.properties` file in the `<SAP_BusinessObjects>\InformationSteward\MM\config` directory by default.

On UNIX platforms:

- `ISRepositoryUtility.sh` is installed to the `<SAP_BusinessObjects>/InformationSteward/MM/bin` directory by default.
- The Repository utility uses the parameter values in the `InstallUtility.properties` file in the `<SAP_BusinessObjects>/InformationSteward/MM/config` directory by default.

You can override the default parameter values in `InstallUtility.properties` by either of the following actions:

- As a best practice, make a backup of the `InstallUtility.properties` file and then change the parameter values in `InstallUtility.properties`.
- Enter the parameters with their new values in the command line when you run the Information Steward Repository utility.

[Creating a new repository \[page 113\]](#)

[Recovering the repository \[page 114\]](#)

[Upgrading the repository \[page 116\]](#)

[Completing the repository recovery \[page 117\]](#)

[Repository utility parameters \[page 118\]](#)

7.6.1 Creating a new repository

To create a new SAP Information Steward repository and associate it with the current Central Management Console (CMC) system database:

1. Create a database to use for the Information Steward repository.

For a list of supported database types, consult the *Product Availability Matrix* located on the SAP Customer Portal.

Note

To use an Oracle database for the Information Steward repository, ensure that the JDBC driver exists in the following directory:

- On Windows platforms, `<SAP_BusinessObjects>\InformationSteward\MM\lib\ext`
- On UNIX platforms, `<SAP_BusinessObjects>/InformationSteward/MM/lib/ext`

If the JDBC driver is not in this directory, you must download and install it. For more information, see the *Installation Guide*.

2. Stop your web application to prevent users from accessing the Information Steward repository during recovery.
3. Make a backup of the `InstallUtility.properties` file, which is the default configuration file for the `ISRepositoryUtility` command.
 - On windows platforms, back up the properties file located in `<SAP_BusinessObjects>\InformationSteward\MM\config\InstallUtility.properties`.
 - On UNIX platforms, back up the properties file located in `<SAP_BusinessObjects>/InformationSteward/MM/config/InstallUtility.properties`.
4. Open `InstallUtility.properties` and change the values of the connection parameters to the values of the database you created in step 1.
5. If you want to overwrite an existing repository, change the value of the `-o` parameter in the `InstallUtility.properties` file to `true`.

Caution

If you specify `true` for the `-o` parameter, the utility deletes all objects in the repository.

6. On the machine where you performed the Primary installation type for Information Steward, run `ISRepositoryUtility` from the command line.
 - a. Open a command window and change to the `bin` directory where Information Steward is installed.

For example, on Windows platforms type the following:

```
cd C:\Program Files (x86)\SAP BusinessObjects\InformationSteward\MM\bin
```

- b. Enter the `ISRepositoryUtility` command and specify the `create` mode.
For example, type the following:

```
ISRepositoryUtility create boePassword "myPass1" password "myPass2"
```

Note

The password for both the CMS database and the Information Steward repository must be specified in the command line. Do not store passwords in the configuration or properties files.

7. After you create the repository, restart the Web Application Server and CMS.

Note

If you run Data Insight profiling and rule tasks and you created a new repository with different connection information than in Step 1, you must access the SAP Data Services Server Manager to associate the Information Steward job server with the new repository.

Related Information

[Configuring a Data Services Job Server for Data Insight \[page 103\]](#)

7.6.2 Recovering the repository

To recover an Information Steward repository and synchronize it with the current CMS system database:

1. Stop your web application to prevent users from accessing the Information Steward repository during recovery.
2. Restore your Information Steward repository to a backup that you created with your database backup utility.
3. Make a backup of the `InstallUtility.properties` file, which is the default configuration file for the `ISRepositoryUtility` command.
 - On Windows platforms, back up this file
`<SAP_BusinessObjects>\InformationSteward\MM\config\InstallUtility.properties`.
 - On UNIX platforms, back up this file `<SAP_BusinessObjects>/InformationSteward/MM/config/InstallUtility.properties`.
4. Open the `InstallUtility.properties` file to edit it.

- a. Enter the path and file name of the recovery report file that the utility will generate for this recover operation.

For example: `reportFilePath=<SAP_BusinessObjects> /MM/log/
ISRepoUtil_Recovery_report`

- b. Verify that the values of the following parameters are accurate for the database you are using for your Information Steward repository.

- `repoServer`
- `repoServerPort`
- `repoDatabase`
- `repoConnectionString`
- `repoDatabaseEngine`
- `racString` (if your database is Oracle RAC)

For valid values and other information about these parameters, see [Repository utility parameters \[page 118\]](#).

5. On the machine where you performed the *Primary* installation type for Information Steward, run `ISRepositoryUtility` from the command line.

- a. Open a command window and change to the bin directory where the `ISRepositoryUtility` script is located.

For example, on Windows platforms:

```
cd C:\Program Files (x86)\SAP BusinessObjects\InformationSteward\MM\bin
```

- b. Enter the `ISRepositoryUtility` command, the `recover` mode, the password to connect to the CMS, the repository password, and the path of the recovery report file.

For example:

```
ISRepositoryUtility recover  
boePassword "myPass1" password "myPass2"  
reportFilePath  
"C:<SAP_BusinessObjects>\InformationSteward\MM\log\recovery_report.log"
```

Note

The recovery report file will contain information about the Information Steward repository objects that were updated or created and those objects that require further action.

6. Look for the completion message in the command line or in the log `<SAP_BusinessObjects>\InformationSteward\MM\log\ISRepositoryUtility_<timestamp>.log`.

When the utility completes successfully, the following message appears:

```
Utility has completed successfully.
```

7. Restart the web application.

Note

When the recovery has finished, the Update Search Index utility is automatically scheduled to run five minutes later.

After the `ISRepositoryUtility` utility completes, the following objects will be recovered to their status when the database backup was taken:

- Metadata Management Integrator Sources
- Metadata Management utility configurations (except for Purge, Calculate Scorecard, Computer Report Lineage, and Update Search Index)
- Metadata Management Source groups
- Data Insight profile and rule tasks
- Data Insight Connections
- Data Insight Projects (except for Enterprise Project)

ⓘ Note

Schedules and Data Insight profile and rule tasks that have been deleted after the database backup was created will not be recovered. Therefore, they will be deleted from the recovered repository.

→ Remember

You must complete the repository recovery by taking the actions indicated by the messages in the recovery report file.

7.6.3 Upgrading the repository

To upgrade an Information Steward repository:

1. Stop your web application to prevent users from accessing the Information Steward repository during upgrade.
2. Make a backup of the `InstallUtility.properties` file, which is the default configuration file for the `ISRepositoryUtility` command.
 - On Windows platforms, back up the properties file located in `<SAP_BusinessObjects>\InformationSteward\MM\config\InstallUtility.properties`.
 - On UNIX platforms, back up the properties file located in `<SAP_BusinessObjects>/InformationSteward/MM/config/InstallUtility.properties`.
3. Open `InstallUtility.properties` and change the values of the connection parameters to the values of your repository.
4. On the computer where you performed the Primary installation type for Information Steward, run `ISRepositoryUtility` from the command line.
 - a. Open a command window and change to the `bin` directory where Information Steward is installed.

For example, on Windows platforms type the following:

```
cd C:\Program Files (x86)\SAP BusinessObjects\InformationSteward\MM\bin
```

- b. Enter the `ISRepositoryUtility` command and specify the upgrade mode.
For example, type the following:

```
ISRepositoryUtility upgrade boePassword "myPass1" password "myPass2"
```


Note

The password for both the CMS database and the Information Steward repository must be specified in the command line. Do not store passwords in the configuration or properties files.

5. After you upgrade the repository, restart the Web Application Server and CMS.
6. If you are migrating to new hardware and upgrading Information Steward, use the appropriate tool for the following upgrade paths:
 - If you are upgrading from SAP BusinessObjects Enterprise XI 3.1 and Metadata Management 3.1 to BI Platform 4.0.x and Information Steward 4.1, use the Upgrade Manager to move Information Steward objects (such as Metadata Management Integrator Sources).
 - If you are upgrading from BI Platform 4.0 and Information Steward 4.0 to BI Platform 4.0 and Information Steward 4.1, use the SAP Promotion Management tool to move Information Steward objects (such as cleansing packages).

Related Information

[Configuring a Data Services Job Server for Data Insight \[page 103\]](#)

[Objects supported for migration \[page 328\]](#)

7.6.4 Completing the repository recovery

After you run the Repository utility in `recover` mode, you might need to take additional actions to synchronize the CMS system database.

1. Open the recovery report file that the Repository utility created in the file path that you specified in the `reportFilePath` parameter.
2. If you see the following message in the recovery report file, you can skip the rest of this procedure.

```
All objects are in sync. No further
action required.
```

3. If you see the following message in the recovery report file, review the subsequent messages to obtain the names of the objects that require additional actions.

```
Update the missing information in the
Central Management Console for the placeholder object(s)
listed below:
```

Possible placeholder objects that require additional information include:

- Metadata Management integrator sources
 - Data Insight connections
4. Go to the Central Management Console (CMC) home page and click *Information Steward*.
 5. For Metadata Management integrator sources:
 - a. Expand the *Metadata Management* node and click the *Integrator Sources* node in the navigation tree on the left.

- b. Select the name of the integrator source from the list on the right.
 - c. If you want to keep this integrator source, click **Action > Properties** and enter the connection properties (which include the server, port, database, user name, and password).
 - d. If you want to delete this integrator source, click **Manage > Delete**.
6. For Data Insight connections:
- a. Click the **Connections** node in the navigation tree on the left.
 - b. Select the name of the connection from the list on the right.
 - c. If you want to keep this connection, click **Action > Properties** and enter the connection properties (which include the server, port, database, user name, and password).
 - d. If you want to delete this integrator source, click **Manage > Delete**.

7.6.5 Repository utility parameters

By default, the Repository utility uses the parameter values in the `InstallUtility.properties` file. You can override the default parameter values by either of the following actions:

- As a best practice, make a backup of the `InstallUtility.properties` file and then change the parameter values in `InstallUtility.properties`.
- Enter the parameters with their new values in the command line when you run this utility.

Note

You must always specify the passwords for both the CMS database and the Information Steward repository in the command line when you run the Repository utility.

Parameter	Accepted Values	Required	Description
-c	Absolute path of the configuration file. Default: <code><SAP_BusinessObjects>/MM/config/InstallUtility.properties</code>	No	Location and name of the configuration file whose parameter values this utility uses.

-h	None	No	Specifies all the parameters and their descriptions.
----	------	----	--

Note

All parameters and their descriptions display if you specify no parameters when you run `ISRepositoryUtility`.

Parameter	Accepted Values	Required	Description
-m	create, recover, upgrade Default: create	Yes	Specifies the operation mode.
-o	true, false Default: false	No	Overwrites existing repository when set to true, and you: <ul style="list-style-type: none"> • Create a new repository. • Specify an existing repository.
reportFilePath	Absolute path of the recovery report file.	Yes if mode is "recover"	Location of the recovery report file that this utility generates when running in recover mode. Enclose path with double quotation marks, such as: reportFilePath "C:\Program Files (x86)\SAP BusinessObjects\Information Steward\MM\log\report.log".
boeUser	Any value accepted by the CMS.	Yes	Specifies the administrator user account to connect to the CMS.
boePassword	Any value accepted by the CMS.	Yes	Specifies the administrator user account password to connect to the CMS.
boeServer	Any value accepted by the CMS.	Yes	Specifies the CMS server name. (Optional) Follow server name with a colon and the CMS port number, such as cms1:6400.
boeAuthenticationr	Any value accepted by the CMS. Default: secEnterprise	Yes	Specifies the authentication method to connect user to the CMS.
user	Any value accepted by the database type.	Yes	Specifies the SAP Information Steward repository user name.
password	Any value accepted by the database type.	Yes	Specifies the SAP Information Steward repository password.

Parameter	Accepted Values	Required	Description
repoServer	Any value accepted by the database type.	Yes	Specifies the Information Steward repository server name.
repoServerPort	Integer value	Yes	Specifies the port to communicate with the SAP Information Steward repository server.
repoDatabaseEngine	See list in the Description column	Yes	Specifies the SAP Information Steward repository database type. Database types include the following: <ul style="list-style-type: none"> • IBM DB2 • MS SQL Server • SAP Sybase SQL Anywhere • Oracle • SAP HANA • SAP ASE
repoDatabase	Any value accepted by the database server.	Yes	Specifies the SAP Information Steward repository database name.
repoConnectionString	Any value accepted by the database server.	Yes	Specifies the SAP Information Steward repository connection information. For Microsoft SQL Server, if the SQL Server is an instance name, use this parameter to pass it, such as: <code>repoConnectionString SQLServerExpress.</code>
repoMSSQLWinAuth	true, false Default: false	No	For Microsoft SQL Server repositories, specifies whether to use Windows authentication instead of Microsoft SQL Server authentication when set to true. For example, <code>ISRepositoryUtility create repoMSSQLWinAuth true boePassword test1234 -o true.</code> For more information about how to use Windows authentication with Microsoft SQL Server, see the Microsoft SQL Server documentation.

Note

For the most up-to-date list of supported databases, see the Information Steward Product Availability Matrix (PAM).

Parameter	Accepted Values	Required	Description
racString	Any value accepted by the database server.	No	Specifies additional Oracle RAC connection information.
repoUseSSL	true, false Default: false	No	For SAP HANA, enables SSL/TLS encryption when set to true.
repoSSLVerifyCert	true, false Default: false	No	For SAP HANA, uses the client certificate to verify the host name when set to true.
repoHANASSLHostname	Any value accepted by the database server.	Yes, only when the following options are set to true: repoUseSSL repoSSLVerifyCert	For SAP HANA, specifies host name used to verify the server's identity.
repoHANASSLKeyStore	Any value accepted by the database server.	Yes, only when the following options are set to true: repoUseSSL repoSSLVerifyCert	For SAP HANA, specifies absolute location of the Java KeyStore file.
repoHANASSLKeyStoreType	Any value accepted by the database server.	Yes, only when the following options are set to true: repoUseSSL repoSSLVerifyCert	For SAP HANA, specifies type of KeyStore certificate file.
repoHANASSLKeyStorePassword	Any value accepted by the database server.	Yes, only when the following options are set to true: repoUseSSL repoSSLVerifyCert	For SAP HANA, specifies password used to access the private key from the KeyStore file
repoHANASSLTrustStore	Any value accepted by the database server.	Yes, only when the following options are set to true: repoUseSSL repoSSLVerifyCert	For SAP HANA, specifies the absolute location of the Java TrustStore file.
repoHANASSLTrustStoreType	Any value accepted by the database server.	Yes, only when the following options are set to true: repoUseSSL repoSSLVerifyCert	For SAP HANA, specifies the Java TrustStore file format.

Parameter	Accepted Values	Required	Description
repoHANASSLTrustStorePassword	Any value accepted by the database server.	Yes, only when the following options are set to true: repoUseSSL repoSSLVerifyCert	For SAP HANA, specifies the password to access the private key from the TrustStore file.
repoHANASSLCommonCryptoKeyStore	Any value accepted by the database server.	Yes, only when the following options are set to true: repoUseSSL repoSSLVerifyCert	For SAP HANA, specifies the name of the PSE file that you configured in SECUDIR.

8 Data Insight Administration

8.1 Administration overview for Data Insight

Each deployment of SAP Information Steward supports multiple users in one or more Data Insight projects to assess and monitor the quality of data from various sources. In Data Insight, users work with data within projects and connections:

Data Insight term	Description
Project	Collaborative workspace for data stewards and data analysts to assess and monitor the data quality of a specific domain and for a specific purpose (such as customer quality assessment, sales system migration, and master data quality monitoring).
Connection	Defines the parameters for Information Steward to access a data source. A data source can be a relational database, application, or file.

Data Insight users in a project can perform tasks such as:

- Browse tables and files in connections
- Add tables of interest to the project
- Profile data
- Join data from multiple tables and create expressions to form a specific view of the data
- Execute rules to measure the data quality
- Create data quality scorecards
- Monitor data quality
- Run profile tasks and rule tasks in the Enterprise project to enable profile results and rule results to appear in the Metadata Management module

For more details about what users can do in Data Insight, see the Data Insight section in the *User Guide*.

Before Data Insight users can perform tasks, administrators must set up each user in the Central Management Console (CMC) and perform other setup tasks:

- Define connections to data sources
- Create projects
- Add user names to pre-defined Data Insight user groups to access connections and projects
- Define connections to save all data that failed rules for rule tasks. Each rule task can use a different failed data connection.

After a Data Insight user creates tasks (profile or rule), the administrator can perform the following tasks:

- Create schedules to run the profile task and rule task at regular intervals.
- Modify the default schedule to run utilities if the frequency of the profile and rule tasks warrant it.
- Modify application-level settings to change the default configuration for Data Insight.

Related Information

[Data Insight Connections \[page 125\]](#)

[Data Insight projects \[page 160\]](#)

[Enterprise project \[page 162\]](#)

[User rights in Data Insight \[page 58\]](#)

[Scheduling a utility \[page 280\]](#)

[Information Steward configuration settings \[page 44\]](#)

8.2 Data Insight performance settings

Some of the Data Insight settings in the Information Steward Settings window control task performance by increasing or decreasing the number of needed resources, and controlling the repository size.

The following parameters affect the amount of data that Data Insight needs to process for profiles and rules. You reduce the amount of data to process to conserve resources and to promote efficient processing.

- *Max input size*
- *Max sample data size*
- *Optimization period*

The following parameters affect the size of the Information Steward Data Insight subcategory repository. If you increase these values, the repository size also increases and you might need to free space more often.

- *Max sample data size*
- *Number of distinct values*
- *Number of patterns*
- *Number of words*
- *Results retention period*

8.3 Data Insight Connections

Create connections to databases, applications, and files for storing information from Data Insight processes.

You can create connections for SAP Data Insight for the following reasons:

Reason for connecting	Details
Profiling data	<p>Configure connections to the following types of data sources to collect profile attributes that can help you determine the data quality and structure:</p> <ul style="list-style-type: none">Databases (such as Microsoft SQL Server, IBM DB2, SAP HANA, Oracle, MySQL, Informix IDS, SAP ASE, and ODBC)Applications (such as SAP Business Suite and SAP Business Warehouse)Text files
Failed data	<p>Configure connections to supported databases to store failed data records and related information about data that failed your quality validation rules. You can create multiple failed data connections. For example, create a separate failed data connection for each rule task.</p>

Ensure that you have the proper permissions to access the database connection:

- If you profile and run validation rules on the data, you must have permissions to read the metadata and data from the source tables.
- If you store failed data, you must have permissions to create, modify, and delete tables, and to create stored procedures.


Note

For a complete list of supported databases, applications, and their versions, see the [PAM](#) on the SAP Support Portal.

8.3.1 Defining a Data Insight connection to a database

Create and configure a database connection for Data Insight.

Before following the steps below, log in to the Central Management Console (CMC) with a user name that belongs to the Data Insight administrator group or that has the create rights for *Connections* in Information Steward.

- At the CMC home page, click Information Steward.
- Select the *Connections* node in the tree panel at left.
- Click **Manage > New > Connection** .
The *Create Connection* dialog box opens.

4. Complete the parameters based on the descriptions in the table below.

Note

The parameters vary based on the database type you choose. For database-specific parameters and information, see the applicable connection parameter topic listed for the applicable database type.

Parameter	Description
<i>Connection Name</i>	(Required) Enter a name for the connection. Name requirements: <ul style="list-style-type: none"> • Maximum length is 64 characters • Can be multi-byte • Case insensitive • Can include underscores and spaces • Cannot include other special characters: ?!@#\$%^&*()-+={}[]:;'\^ .,`~
<i>Description</i>	(Optional) Enter a description for the connection.
<i>Connection Type</i>	(Required) Select <i>Database connection</i> .
<i>Purpose</i>	(Required) Select one of the following options: <ul style="list-style-type: none"> • <i>For data profiling</i>: Profiles the data in this connection. • <i>For data that failed rules</i>: Stores all the data that fails specific quality validation rules, the source data, and metadata tables and views. • <i>For data review</i>: For Match Review, contains duplicate records that are the output of a Data Services match transform job.
<i>Bulk Loading</i>	(Required) Select <i>Yes</i> to enable bulk loading for failed data. Bulk loading could improve the processing speed when the software searches for failed data in large database files. Bulk Loading is available when the <i>Purpose</i> is for data that failed rules. Applicable for the following database types: <ul style="list-style-type: none"> • DB2 • Oracle • SAP HANA
<i>Rows per commit</i>	The number of rows processed before a commit. Default is 1000. Setting the rows per commit to a larger number of rows may increase performance.
<i>Database Type</i>	(Required) Select the database type that contains the data that you want to use for profiling or failed data.
<i>Database Version</i>	(Required) Enter current version number for the database that you chose in <i>Database Type</i> .
<i>Server Name</i>	(Required) Enter the server name that the connection will use.

Parameter	Description
<i>Database Name</i>	(Required) Enter the name of the database that you will use for this connection.
<i>Windows Authentication</i>	(Required for Microsoft SQL Server) Select <i>Yes</i> to enable Windows authentication. Select <i>No</i> to disable Windows authentication.
<i>User Name, Password</i>	(Required) Enter the user name and password that will be used to access this connection.
<i>Unsupported Data Types</i>	(Required) Select the action that the system should take if the database contains unsupported data types. <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 5px;"> <p>Note This parameter is only available when the purpose of the connection is for data profiling.</p> </div>
<i>Language, Client Code Page, Server Code Page</i>	(Required) Select a valid option from the drop lists or select Default.
<i>Use NVARCHAR</i>	Use NVARCHAR for VARCHAR data types. <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 5px;"> <p>Note This parameter is only available when the purpose of the connection is for data that failed rules.</p> </div>
<i>Append failed data</i>	(Required) Select <i>Yes</i> to append failed data to the existing table. Select <i>No</i> to save failed data to a new table. <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 5px;"> <p>Note This parameter is only available when the purpose of the connection is for data that failed rules.</p> </div> <p>If you select to include the source data with failed data, and you have a primary key defined in your source data, selecting <i>Yes</i> for this option also appends changed information to the source data as well as failed data. When you select <i>No</i>, the software overwrites the failed data with the data from the last run.</p>
<i>VARCHAR Size</i>	(Required) Maximum size of imported VARCHAR . <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 5px;"> <p>Note This parameter is only available when the purpose of the connection is for data profiling.</p> </div>

5. (Optional) Click *Test Connection* to verify that Information Steward can connect to the newly defined connection.
If the connection fails, correct applicable connection information and click *Test Connection* again.
6. When you are finished with the *Create Connection* dialog box, click *Save*.

The newly configured connection appears in the list of connections in [Information Steward > Connections](#).

After you save the connection, you cannot change the *Connection Name*, *Connection Type*, *Purpose* and connection parameters that uniquely identify a database.

You must also authorize users so that they can perform tasks such as view the data, run profile tasks and run validation rules on the data.

Note

If you configure the system environment variables for the database (for example, ORACLE_HOME, LD_LIBRARY_PATH, SHLIB_PATH, locale settings, ulimit), you must restart the Server Intelligence Agent (SIA). Users cannot access the Central Management Server (CMS) while the SIA is stopped. Therefore, you should consider performing this configuration during scheduled down time to limit the effect on your users.

[Google BigQuery connection parameters \[page 129\]](#)

To help you complete the options for a Google BigQuery connection, view a list of options and gather the necessary information ahead of time.

[IBM DB2 connection parameters \[page 129\]](#)

To help you complete the options for a Data Insight connection to IBM DB2, view a list of options and gather the necessary information ahead of time.

[Informix database connection parameters \[page 131\]](#)

To help you complete the options for a Data Insight connection to Informix, view a list of options and gather the necessary information ahead of time.

[Microsoft SQL Server connection parameters \[page 133\]](#)

To help you complete the options for a Data Insight connection to Microsoft SQL Server, view a list of options and gather the necessary information ahead of time.

[MySQL database connection parameters \[page 135\]](#)

To help you complete the options for a Data Insight connection to MySQL, view a list of options and gather the necessary information ahead of time.

[Netezza connection parameters \[page 136\]](#)

To help you complete the options for a Data Insight connection to Netezza, view a list of options and gather the necessary information ahead of time.

[ODBC connection parameters \[page 138\]](#)

To help you complete the options for a Data Insight connection to ODBC, view a list of options and gather the necessary information ahead of time.

[Oracle database connection parameters \[page 139\]](#)

To help you complete the options for a Data Insight connection to Oracle, view a list of options and gather the necessary information ahead of time.

[SAP HANA connection parameters \[page 140\]](#)

To help you complete the options for a Data Insight connection to SAP HANA, view a list of options and gather the necessary information ahead of time.

[SQL Anywhere connection parameters \[page 143\]](#)

To help you complete the options for a Data Insight connection to SQL Anywhere, view a list of options and gather the necessary information ahead of time.

[SAP ASE connection parameters \[page 144\]](#)

To help you complete the options for a Data Insight connection to SAP ASE, view a list of options and gather the necessary information ahead of time.

[Snowflake connection parameters \[page 145\]](#)

To help you complete the options for a Snowflake connection, view a list of options and gather the necessary information ahead of time.

[Sybase IQ connection parameters \[page 146\]](#)

To help you complete the options for a Data Insight connection to Sybase IQ, view a list of options and gather the necessary information ahead of time.

[Teradata connection parameters \[page 147\]](#)

To help you complete the options for a Data Insight connection to Teradata, view a list of options and gather the necessary information ahead of time.

8.3.1.1 Google BigQuery connection parameters

To help you complete the options for a Google BigQuery connection, view a list of options and gather the necessary information ahead of time.

When you create a Google BigQuery connection in the Central Management Console, create a DSN connection only.

The following table contains descriptions for options specific to Google BigQuery connection parameters with values. For descriptions of more common options, see [Defining a Data Insight connection to a database \[page 125\]](#).

Google BigQuery option	Value	Description
<i>Database Type</i>	Google BigQuery ODBC	Select the database type.
<i>Use Data Source Name (DSN)</i>	Yes	The connection must be DSN.
<i>Data Source Name</i>	Name of data source	Enter the name of the data source to connect to. The data source must already exist.

8.3.1.2 IBM DB2 connection parameters

To help you complete the options for a Data Insight connection to IBM DB2, view a list of options and gather the necessary information ahead of time.

The following table contains descriptions for the IBM DB2 connection parameters with possible values.

DB2 option	Possible values	Description
Database version	DB2 UDB <version number>	Select the version of your DB2 client. This is the version of DB2 that this Data Insight connection accesses.

DB2 option	Possible values	Description
Use Data Source Name (DSN)	Yes	Select whether or not to use DSN to connect to the database.
	No	Default is No, and a server name (also known as DSN-less) connection will be used. For a DSN-less connection, you must fill in <i>Server name</i> , <i>Database name</i> , and <i>Port Number</i> . If you select Yes, you must fill in <i>Data Source Name</i>
Server name	Refer to the requirements of your database	Type the DB2 database server name. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Database name	Refer to the requirements of your database	Type the name of the database defined in DB2. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Port Number	Five digit integer Default: 50000	Type the port number to connect to this database. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Data Source Name	Refer to the requirements of your database	Type the data source name defined in DB2 for connecting to your database. This option is required if <i>Use Data Source Name (DSN)</i> is set to Yes.
User name	The value is specific to the database server and language.	Enter the user name of the account through which SAP Information Steward accesses the database.
Password	The value is specific to the database server and language.	Enter the user's password.
Bulk Loading	Yes	Select Yes to enable bulk loading for failed data. Bulk loading could improve the processing speed when the software searches for failed data in large database files.
	No	
<div style="border: 1px solid #ccc; background-color: #f0f0f0; padding: 10px;"> <p>Note</p> <p>This parameter is only available when the purpose of the connection is for data that failed rules.</p> </div>		
Rows per commit	Default is 1000	For bulk loading. The number of rows processed before a commit. Setting the rows per commit to a larger number of rows may increase performance.
Unsupported Data Types	Import as VARCHAR	Select the action to take to handle unsupported data types.
	Do not import	Default is Import as VARCHAR.
VARCHAR Size	Any positive integer from 1 to 4000	Maximum size of the imported VARCHAR data. Default is 255.

DB2 option	Possible values	Description
Language	default Select the correct language for your database server.	Language abbreviation specified in the ISO 639-2/T standard.
Client Code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i> .	Code page of the database client.
Server code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i> .	Code page of the database server.
Append failed data	Yes or No	Yes: Appends to the existing data. No: Drops and recreates failed data tables. Default is No.

Note

This parameter is only available when the purpose of the connection is for data that failed rules.

8.3.1.3 Informix database connection parameters

To help you complete the options for a Data Insight connection to Informix, view a list of options and gather the necessary information ahead of time.

The following table contains descriptions for the Informix connection parameters with possible values.

Informix option	Possible values	Description
Database version	Informix IDS <version number>	Select the version of your Informix client. This is the version of Informix that this profile connection accesses.
Use Data Source Name (DSN)	Yes or No	Select whether or not to use DSN to connect to the database. Default is No, and a server name (also known as DSN-less) connection will be used. For a DSN-less connection, you must fill in <i>Server name</i> , <i>Database name</i> , and <i>Port Number</i> . If you select Yes, you must fill in <i>Data Source Name</i>
Server name	Refer to the requirements of your database.	Type the Teradata database server name. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.

Informix option	Possible values	Description
Database name	Refer to the requirements of your database.	Type the name of the database defined in Informix. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Port Number	Four digit integer Default: 1526	Type the port number to connect to this database. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Data Source Name	Refer to the requirements of your database.	Type the Data Source Name defined in the ODBC. This option is required if <i>Use Data Source Name (DSN)</i> is set to Yes.
User name	The value is specific to the database server and language.	Enter the user name of the account through which SAP Information Steward accesses the database.
Password	The value is specific to the database server and language.	Enter the user's password.
Unsupported Data Types	Import as VARCHAR Do not import	Select the action to take to handle unsupported data types. Default is Import as VARCHAR.
VARCHAR Size	Any positive integer from 1 to 4000	Maximum size of the imported VARCHAR data. Default is 255.
Language	default Select the correct language for your database server.	Language abbreviation specified in the ISO 639-2/T standard.
Client Code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i>	Code page of the database client.
Server code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i>	Code page of the database server.
Append failed data	Yes or No	Yes: Appends to the existing data. No: Drops and recreates failed data tables. Default is No.

ⓘ Note

This parameter is only available when the purpose of the connection is for data that failed rules.

8.3.1.4 Microsoft SQL Server connection parameters

To help you complete the options for a Data Insight connection to Microsoft SQL Server, view a list of options and gather the necessary information ahead of time.

To use Microsoft SQL Server as a profile source when SAP Information Steward is running on a UNIX platform, you must use an ODBC driver, such as the DataDirect ODBC driver.

For more information about how to obtain the driver, see the [Product Availability Matrix \(PAM\)](#).

The following table contains descriptions for the Microsoft SQL Server connection parameters with possible values.

Microsoft SQL Server option	Possible values	Description
Database version	Microsoft SQL Server <version number>	Select the version of your SQL Server client. This is the version of SQL Server that this profile source accesses.
Server Name	Computer name, fully qualified domain name, or IP address	Enter the name of machine where the SQL Server instance is located.
Database Name	Refer to the requirements of your database.	Enter the name of the database to which the profiler connects.
Windows Authentication	Yes or No Default: No	Indicate whether or not Windows Authentication is used. The default is <i>No</i> , which means Microsoft SQL Server Authentication is used.
User Name	The value is specific to the database server and language.	Enter the user name of the account through which Information Steward accesses the database.
Password	The value is specific to the database server and language.	Enter the user's password.
Use SSL Authentication	Yes or No. The default is No.	Select Yes to enable SSL Authentication to protect network communication.
<div style="border: 1px solid #ccc; background-color: #f0f0f0; padding: 10px; margin: 10px 0;"> <p>Note</p> <p>For UNIX, enter encryption parameters in the SSL Encryption Parameters option. Information Steward does not require SSL encryption parameters for Windows.</p> </div>		
Encryption Parameters	Enter the SSL encryption parameters when you select SSL Authentication and you are on UNIX.	When you are on a UNIX platform and you select to use SSL authentication, enter the encryption parameters in this option.
UnixTrustStorePassword	If you don't password-protect the specified trust store, leave this option blank.	Specifies the password associated with the specified trust store. UNIX/Linux only.

Microsoft SQL Server option	Possible values	Description
Unsupported Data Types	Import as VARCHAR Do not import	Select the action to take to handle unsupported data types. Default is Import as VARCHAR.
VARCHAR Size	Any positive integer from 1 to 4000	Maximum size of the imported VARCHAR data. Default is 255.
Language	Select the correct language for your database server.	Language abbreviation specified in the ISO 639-2/T standard.
Client Code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i> .	Code page of the database client.
Server code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i> .	Code page of the database server.
Append failed data	Yes or No	Yes: Appends to the existing data. No: Drops and recreates failed data tables. Default is No.

Note

This parameter is only available when the purpose of the connection is for data that failed rules.

[SSL encryption parameters for SQL Server on UNIX \[page 134\]](#)

For Microsoft SQL Server connection configuration, select SSL authentication and enter configuration information for UNIX platforms.

8.3.1.4.1 SSL encryption parameters for SQL Server on UNIX

For Microsoft SQL Server connection configuration, select SSL authentication and enter configuration information for UNIX platforms.

Enter the following string in the SSL Encryption Parameters option:

```
<encryption_parameters>ValidateCertificate=yes;CertHostName=vantgvmwin1042;StoredCertPasswd=Password;TrustStore=/build/i311829/cert.pem; </encryption_parameters>
```

8.3.1.5 MySQL database connection parameters

To help you complete the options for a Data Insight connection to MySQL, view a list of options and gather the necessary information ahead of time.

The following table contains descriptions for the MySQL connection parameters with possible values.

MySQL option	Possible values	Description
Database version	MySQL <version number>	Select the version of your MySQL client. This is the version of MySQL that this profile connection accesses.
Use Data Source Name (DSN)	Yes or No	Select whether or not to use DSN to connect to the database. Default is No, and a server name (also known as DSN-less) connection will be used. For a DSN-less connection, you must fill in <i>Server name</i> , <i>Database name</i> , and <i>Port Number</i> . If you select Yes, you must fill in <i>Data Source Name</i>
Server name	Refer to the requirements of your database.	Type the MySQL database server name. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Database name	Refer to the requirements of your database.	Type the name of the database defined in MySQL. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Port Number	Refer to the requirements of your database.	Type the port number to connect to this database. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Data Source Name	Refer to the requirements of your database.	Select or type the Data Source Name defined in the ODBC Administrator for connecting to the database you want to profile. This option is required if <i>Use Data Source Name (DSN)</i> is set to Yes.
User name	The value is specific to the database server and language.	Enter the user name of the account through which the software accesses the database.
Password	The value is specific to the database server and language.	Enter the user's password.
Unsupported Data Types	Import as VARCHAR Do not import	Select the action to take to handle unsupported data types. Default is Import as VARCHAR.
VARCHAR Size	Any positive integer from 1 to 4000	Maximum size of the imported VARCHAR data. Default is 255.

MySQL option	Possible values	Description
Language	default Select the correct language for your database server.	Language abbreviation specified in the ISO 639-2/T standard.
Client Code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i>	Code page of the database client.
Server code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i>	Code page of the database server.
Append failed data	Yes or No	Yes: Appends to the existing data. No: Drops and recreates failed data tables. Default is No.

Note
This parameter is only available when the purpose of the connection is for data that failed rules.

8.3.1.6 Netezza connection parameters

To help you complete the options for a Data Insight connection to Netezza, view a list of options and gather the necessary information ahead of time.

The following table contains descriptions for the Netezza connection parameters with possible values.

Netezza option	Possible values	Description
Database version	Netezza NPS <version number>	Select the version of your Netezza client. This is the version of Netezza that this profile connection accesses.
Use Data Source Name (DSN)	Yes or No	Select whether or not to use DSN to connect to the database. Default is No, and a server name (also known as DSN-less) connection will be used. For a DSN-less connection, you must fill in <i>Server Name</i> , <i>Database Name</i> , and <i>Port Number</i> . If you select Yes, you must fill in <i>Data Source Name</i>
Server Name	Refer to the requirements of your database.	Type the Netezza database server name. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.

Netezza option	Possible values	Description
Database Name	Refer to the requirements of your database.	Type the name of the database defined in Netezza. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Port Number	Refer to the requirements of your database.	Enter the number of the database port. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Data Source Name	Refer to the requirements of your database.	Select or type the Data Source Name defined in the ODBC Administrator for connecting to the database you want to profile. This option is required if <i>Use Data Source Name (DSN)</i> is set to Yes.
User name	The value is specific to the database server and language.	Enter the user name of the account through which the software accesses the database.
Password	The value is specific to the database server and language.	Enter the user's password.
Unsupported Data Types	Import as VARCHAR. Do not import.	Select the action to take to handle unsupported data types. Default is Import as VARCHAR.
VARCHAR Size	Any positive integer from 1 to 4000.	Maximum size of the imported VARCHAR data. Default is 255.
Language	Select the correct language for your database server.	Language abbreviation specified in the ISO 639-2/T standard.
Client Code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i>	Code page of the database client.
Server code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i>	Code page of the database server.
Append failed data	Yes or No	Yes: Appends to the existing data. No: Drops and recreates failed data tables. Default is No.

Note

This parameter is only available when the purpose of the connection is for data that failed rules.

8.3.1.7 ODBC connection parameters

To help you complete the options for a Data Insight connection to ODBC, view a list of options and gather the necessary information ahead of time.

The following table contains descriptions for the ODBC connection parameters with possible values.

ODBC option	Possible values	Description
Data Source Name	Refer to the requirements of your database.	Select or type the Data Source Name defined in the ODBC Administrator for connecting to the database you want to profile.
Additional information	Alphanumeric characters and underscores, or blank	Enter information for any additional parameters that the data source supports (parameters that the data source's ODBC driver and database support). Use the format: <parameter1=value1; parameter2=value2>
User name	The value is specific to the database server and language.	Enter the user name of the account through which the software accesses the database.
Password	The value is specific to the database server and language.	Enter the user's password.
Unsupported Data Types	Import as VARCHAR Do not import	Select the action to take to handle unsupported data types. Default is Import as VARCHAR.
VARCHAR Size	Any positive integer from 1 to 4000	Maximum size of the imported VARCHAR data. Default is 255.
Language	default Select the correct language for your database server.	Language abbreviation specified in the ISO 639-2/T standard.
Client Code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i> .	Code page of the database client.
Server code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i> .	Code page of the database server.
Append failed data	Yes or No	Yes: Appends to the existing data. No: Drops and recreates failed data tables. Default is No.

Note

This parameter is only available when the purpose of the connection is for data that failed rules.

8.3.1.8 Oracle database connection parameters

To help you complete the options for a Data Insight connection to Oracle, view a list of options and gather the necessary information ahead of time.

The following table contains descriptions for the Oracle connection parameters with possible values.

Oracle option	Possible values	Description
Database version	Oracle <version number>	Select the version of your Oracle client. This is the version of Oracle that this profile connection accesses.
Use TNS Name	Yes or No	Select whether or not to use TNS to connect to the database. Default is No, and a server name (also known as DSN-less) connection will be used. For a TNS-less connection, you must fill in <i>Server Name</i> , <i>Instance Name</i> , and <i>Port Number</i> . If you select Yes, you must fill in <i>Database Connection Name</i> , which only displays when you select Yes.
Server Name	Computer name, fully qualified domain name, or IP address	Enter the name of machine where the Oracle Server instance is located. This option is required if <i>Use TNS Name</i> is set to No.
Instance Name	Refer to the requirements of your database.	Enter the System ID for the Oracle database. This option is required if <i>Use TNS Name</i> is set to No.
Port Number	Four digit integer Default: 1521	Enter the port number to connect to this Oracle Server. This option is required if <i>Use TNS Name</i> is set to No.
Database Connection Name	Refer to the requirements of your database.	Enter an existing Oracle connection through which the software accesses sources defined in this profile connection. This option is required if <i>Use TNS Name</i> is set to Yes.
User Name	The value is specific to the database server and language.	Enter the user name of the account through which the software accesses the database.
Password	The value is specific to the database server and language.	Enter the user's password.
Bulk Loading	Yes or No	Select <i>Yes</i> to enable bulk loading for failed data. Bulk loading could improve the processing speed when the software searches for failed data in large database files.
<div style="border: 1px solid #ccc; background-color: #f0f0f0; padding: 5px;"> <p>Note</p> <p>This parameter is only available when the purpose of the connection is for data that failed rules.</p> </div>		
Rows per commit	Default is 1000	For bulk loading. The number of rows processed before a commit. Setting the rows per commit to a larger number of rows may increase performance.

Oracle option	Possible values	Description
Unsupported Data Types	Import as VARCHAR	Select the action to take to handle unsupported data types.
	Do not import	Default is Import as VARCHAR.
VARCHAR Size	Any positive integer from 1 to 4000	Maximum size of the imported VARCHAR data. Default is 255.
Language	Select the correct language for your database server.	Language abbreviation specified in the ISO 639-2/T standard.
Client Code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i> .	Code page of the database client.
Server code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i> .	Code page of the database client.
Append failed data	Yes or No	Yes: Appends to the existing data. No: Drops and recreates failed data tables. Default is No.

Note
This parameter is only available when the purpose of the connection is for data that failed rules.

8.3.1.9 SAP HANA connection parameters

To help you complete the options for a Data Insight connection to SAP HANA, view a list of options and gather the necessary information ahead of time.

The following table contains descriptions for the SAP HANA connection parameters with possible values.

HANA option	Possible values	Description
Database version	HANA <version number>	Select the version of your SAP HANA client. This is the version of SAP HANA that this Data Insight connection accesses.
Use Data Source Name (DSN)	Yes or No	Select whether or not to use DSN to connect to the database. Default is No, and a server name (also known as DSN-less) connection will be used. For a DSN-less connection, you must fill in <i>Server Name</i> , <i>Database Name</i> , and <i>Port Number</i> . If you select Yes, you must fill in <i>Data Source Name</i>

HANA option	Possible values	Description
Server name	Refer to the requirements of your database.	Type the name of the database defined in SAP HANA. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Port Number	Five digit integer Default: 30015	Enter the port number to connect to this SAP HANA Server. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Enable SSL Encryption	Yes or No	Enable or disable SSL encryption.
Encryption Parameters	ValidateCertificate=yes or no CryptoProvider=commoncrypto or openssl CertHostName= KeyStore=	SSL-related parameters that are directly used in the Data Services datastore. Separate the name-value pairs with semicolons. <ul style="list-style-type: none"> • <i>ValidateCertificate</i> – Specify whether the software validates the SAP HANA server SSL certificate. • <i>CryptoProvider</i> – Specify the crypto provider used for SSL/TLS and X.509 communication. The default for HANA SSL implementation is common-crypto. • <i>CertHostName</i> – Specify the host name used to verify the server identity: <ul style="list-style-type: none"> • Leave blank to reuse the value in <i>Server name</i>. • Enter a string that contains the SAP HANA server hostname. • Enter the wildcard character "*" to not validate the certificate host. • <i>KeyStore</i> – Specify the location and file name for your key store file.

Note

If you choose to use X.509 authentication, the key store PSE file must contain both the SSL server certificate and X.509 client certificate. If the key store file name isn't an absolute path, the software assumes the PSE file is located in \$SECUDIR.

Example

```
ValidateCertificate=yes;CertHostName=abcde123.pal.cmp.corp;KeyStore=/abc//388/CRYPTO/sapsrv.pse
```

Data Source Name	Refer to the requirements of your database.	Type the data source name defined in SAP HANA for connecting to your database. This option is required if <i>Use Data Source Name (DSN)</i> is set to Yes.
------------------	---	---

HANA option	Possible values	Description
Bulk Loading	Yes or No	<p>Yes: Use bulk loading. Ensure that the user account has SAVEPOINT ADMIN privileges.</p> <p>No: Do not use bulk loading.</p> <p>Default is No.</p> <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 5px; margin-top: 10px;"> <p>Note</p> <p>This parameter is only available when the purpose of the connection is for data that failed rules.</p> </div>
Rows per commit	Default is 1000	For bulk loading. The number of rows processed before a commit. Setting the rows per commit to a larger number of rows may increase performance.
User name	The value is specific to the database server and language.	Enter the user name of the account through which SAP Information Steward accesses the database.
Password	The value is specific to the database server and language.	Enter the user's password.
Unsupported Data Types	Import as VARCHAR Do not import	<p>Select the action to take to handle unsupported data types.</p> <p>Default is Import as VARCHAR.</p>
VARCHAR Size	Any positive integer from 1 to 4000.	<p>Maximum size of the imported VARCHAR data.</p> <p>Default is 255.</p>
Language	default Select the correct language for your database server.	Language abbreviation specified in the ISO 639-2/T standard.
Client Code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i> .	Code page of the database client.
Server Code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i> .	Code page of the database server.
Append failed data	Yes or No	<p>Yes: Appends to the existing data.</p> <p>No: Drops and recreates failed data tables.</p> <p>Default is No.</p> <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 5px; margin-top: 10px;"> <p>Note</p> <p>This parameter is only available when the purpose of the connection is for data that failed rules.</p> </div>

8.3.1.10 SQL Anywhere connection parameters

To help you complete the options for a Data Insight connection to SQL Anywhere, view a list of options and gather the necessary information ahead of time.

The following table contains descriptions for the SQL Anywhere connection parameters with possible values.

SQL Anywhere option	Possible values	Description
Database version	SQL Anywhere <version number>	Select the version of your SQL Anywhere client. This is the version of SQL Anywhere that this profile connection accesses.
Use Data Source Name (DSN)	Yes or No	Select whether or not to use DSN to connect to the database. Default is No, and a server name (also known as DSN-less) connection will be used. For a DSN-less connection, you must fill in <i>Server name</i> , <i>Database name</i> , and <i>Port Number</i> . If you select Yes, you must fill in <i>Data Source Name</i>
Server name	Refer to the requirements of your database.	Type the SQL Anywhere database server name. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Database name	Refer to the requirements of your database.	Type the name of the database defined in SQL Anywhere. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Port Number	Refer to the requirements of your database.	Type the port number to connect to this database. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Data Source Name	Refer to the requirements of your database.	Select or type the Data Source Name defined in the ODBC Administrator for connecting to the database you want to profile. This option is required if <i>Use Data Source Name (DSN)</i> is set to Yes.
User name	The value is specific to the database server and language.	Enter the user name of the account through which the software accesses the database.
Password	The value is specific to the database server and language.	Enter the user's password.
Unsupported Data Types	Import as VARCHAR Do not import	Select the action to take to handle unsupported data types. Default is Import as VARCHAR.

SQL Anywhere option	Possible values	Description
VARCHAR Size	Any positive integer from 1 to 4000	Maximum size of the imported VARCHAR data. Default is 255.
Language	default Select the correct language for your database server.	Language abbreviation specified in the ISO 639-2/T standard.
Client Code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i>	Code page of the database client.
Server code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i>	Code page of the database server.
Append failed data	Yes or No	Yes: Appends to the existing data. No: Drops and recreates failed data tables. Default is No.

ⓘ Note

This parameter is only available when the purpose of the connection is for data that failed rules.

8.3.1.11 SAP ASE connection parameters

To help you complete the options for a Data Insight connection to SAP ASE, view a list of options and gather the necessary information ahead of time.

The following table contains descriptions for the SAP ASE connection parameters with possible values.

SAP ASE option	Possible values	Description
Database version	SAP ASE <version number>	Select the version of your SAP ASE client. This is the version of Sybase that this profile connection accesses.
Server Name	Computer name	Enter the name of the computer where the SAP ASE instance is located.
Database Name	Refer to the requirements of your database	Enter the name of the database to which the profiler connects.
User Name	The value is specific to the database server and language.	Enter the user name of the account through which the software accesses the database.
Password	The value is specific to the database server and language.	Enter the user's password.

SAP ASE option	Possible values	Description
Unsupported Data Types	Import as VARCHAR	Select the action to take to handle unsupported data types.
	Do not import	Default is Import as VARCHAR.
VARCHAR Size	Any positive integer from 1 to 4000	Maximum size of the imported VARCHAR data. Default is 255.
Language	Select the correct language for your database server.	Language abbreviation specified in the ISO 639-2/T standard.
Client Code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i>	Code page of the database client.
Server code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i>	Code page of the database server.
Append failed data	Yes or No	Yes: Appends to the existing data. No: Drops and recreates failed data tables. Default is No.

Note

This parameter is only available when the purpose of the connection is for data that failed rules.

8.3.1.12 Snowflake connection parameters

To help you complete the options for a Snowflake connection, view a list of options and gather the necessary information ahead of time.

When you create a Snowflake connection in the Central Management Console (CMC), create a server-based (DSN-less) connection.

Note

For a DSN connection, create the connection outside of the CMC, when you create the ODBC connection.

The following table contains descriptions for options specific to Snowflake connection parameters with values. For descriptions of more common options, see [Defining a Data Insight connection to a database \[page 125\]](#).

Snowflake option	Values	Description
<i>Database Type</i>	Snowflake	Select the database type.

Snowflake option	Values	Description
<i>Database Version</i>	Snowflake 3.x Snowflake 5.x	Enter current version number for your Snowflake connection.
<i>Use Data Source Name (DSN)</i>	No	The connection must be server-based (DSN-less).

8.3.1.13 Sybase IQ connection parameters

To help you complete the options for a Data Insight connection to Sybase IQ, view a list of options and gather the necessary information ahead of time.

The following table contains descriptions for the Sybase IQ connection parameters with possible values.

Sybase IQ option	Possible values	Description
Database version	Currently supported versions	Select the version of your Sybase IQ client. This is the version of Sybase IQ that this datastore accesses.
Use Data Source Name (DSN)	Yes or No	Select whether or not to use DSN to connect to the database. Default is No, and a server name (also known as DSN-less) connection will be used. For a DSN-less connection, you must fill in <i>Hostname</i> , <i>Database name</i> , and <i>Port</i> , and <i>Server name</i> .
Host Name	Computer name or IP address	Type the computer name or IP address. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Database Name	Refer to the requirements of your database.	Type the name of the database defined in Sybase IQ. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Port Number	Four digit integer Default: 2638	Type the number of the database port. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Server name	Refer to the requirements of your database.	Type the Sybase IQ database server name. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Data Source Name	Refer to the requirements of your database.	Select or type the Data Source Name defined in the ODBC Administrator for connecting to your database. This option is required if <i>Use Data Source Name (DSN)</i> is set to Yes.

Sybase IQ option	Possible values	Description
User Name	The value is specific to the database server and language.	Enter the user name of the account through which the software accesses the database.
Password	The value is specific to the database server and language.	Enter the user's password.
Unsupported Data Types	Import as VARCHAR	Select the action to take to handle unsupported data types.
	Do not import	Default is Import as VARCHAR.
VARCHAR Size	Any positive integer from 1 to 4000	Maximum size of the imported VARCHAR data. Default is 255.
Language	Select the correct language for your database server.	Language abbreviation specified in the ISO 639-2/T standard.
Client Code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i>	Code page of the database client.
Server code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i>	Code page of the database server.
Append failed data	Yes or No	Yes: Appends to the existing data. No: Drops and recreates failed data tables. Default is No.

Note

This parameter is only available when the purpose of the connection is for data that failed rules.

8.3.1.14 Teradata connection parameters

To help you complete the options for a Data Insight connection to Teradata, view a list of options and gather the necessary information ahead of time.

The following table contains descriptions for the Teradata connection parameters with possible values.

Teradata option	Possible values	Description
Database version	Teradata <version number>	Select the version of your Teradata client. This is the version of Teradata that this datastore accesses.

Teradata option	Possible values	Description
Use Data Source Name (DSN)	Yes or No	Select whether or not to use DSN to connect to the database. Default is No, and a server name (also known as DSN-less) connection will be used. For a DSN-less connection, you must fill in <i>Server name</i> , <i>Database name</i> , and <i>Port Number</i> . If you select Yes, you must fill in <i>Data Source Name</i>
Server name	Refer to the requirements of your database	Type the Teradata database server name. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Database name	Refer to the requirements of your database	Type the name of the database defined in Teradata. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Port Number	Four digit integer Default: 8888	Type the port number to connect to this database. This option is required if <i>Use Data Source Name (DSN)</i> is set to No.
Data Source Name	Refer to the requirements of your database	Type the Data Source Name defined in the ODBC Administrator for connecting to your database. This option is required if <i>Use Data Source Name (DSN)</i> is set to Yes.
Method to Read Data	ODBC or Parallel Transporter Export Operator	Select the method to use to read data. <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 10px; margin-top: 10px;">Note This option is only available for data profiling and cannot be changed after the connection is created.</div>
Log Directory	Directory path	The directory in which to write log files. The directory is optional and only available for Parallel Transporter Export Operator.
Teradata Director Program ID	Alphanumeric characters, underscores, and punctuation	Identifies the name of the Teradata database to load. Required only for Parallel Transporter Export Operator.
User Name	The value is specific to the database server and language	Enter the user name of the account through which the software accesses the database.
Password	The value is specific to the database server and language	Enter the user's password.
Unsupported Data Types	Import as VARCHAR Do not import	Select the action to take to handle unsupported data types. Default is Import as VARCHAR.

Teradata option	Possible values	Description
VARCHAR Size	Any positive integer from 1 to 4000	Maximum size of the imported VARCHAR data. Default is 255.
Language	Select the correct language for your database server	Language abbreviation specified in the ISO 639-2/T standard.
Client Code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i>	Code page of the database client.
Server code page	See "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i>	Code page of the database server.
Append failed data	Yes or No	Yes: Appends to the existing data. No: Drops and recreates failed data tables. Default is No.

Note

This parameter is only available when the purpose of the connection is for data that failed rules.

8.3.2 Defining a Data Insight connection to an application

You must define a connection to any application that contains data that you want to profile to determine the quality and structure of the data.

To create a connection to an application system:

- Log in to the Central Management Console (CMC) with a user name that has the following authorizations:
 - Authorization to read data on the source application system. For authorizations to SAP Applications, see "SAP user authorizations" in the *SAP Data Services Supplement for SAP*.
 - Either belongs to the Data Insight Administrator group or has the Create right on the *Connections* node.
- At the CMC home page, click Information Steward.
- Select the *Connections* node in the navigation tree on the left.
- Click **Manage > New > Data Insight connection**.
- On the *Create Connection* page, enter the following information.

Option	Description
<i>Connection Name</i>	Name that you want to use for this Data Insight source. <ul style="list-style-type: none"> Maximum length is 64 characters Can be multi-byte

Option	Description
	<ul style="list-style-type: none"> • Case insensitive • Can include underscores and spaces • Cannot include other special characters: ?!@#%&*()-+={}[]:";' / \ , . ^ ~ <p>You cannot change the name after you save the connection.</p>
<i>Description</i>	(Optional) Text to describe this Data Insight source.

- In the *Connection Type* drop-down list, select the *Application connection* value.
- In the *Application Type* drop-down list, select one of the following applications that contains the data you want to profile:
 - SAP Business Warehouse
 - SAP Applications

For information about the specific data components that Information Steward can profile, see the [PAM](#).

- Enter the relevant connection information for the application type.
- Complete the parameters based on the descriptions found in [SAP Applications connection parameters \[page 151\]](#). If you want to verify that Information Steward can connect successfully before you save this profile connection, click *Test connection*.
- Click *Save*.
The newly configured connection appears in the list of connections on the right of the *Information Steward* page.

After you create a connection, you must authorize users to perform tasks such as view the data, run profile tasks and run validation rules on the data.

[SAP Business Warehouse connection parameters \[page 150\]](#)

[SAP Applications connection parameters \[page 151\]](#)

8.3.2.1 SAP Business Warehouse connection parameters

The SAP Business Warehouse connection has the same options as the SAP Applications connection type.

Parent topic: [Defining a Data Insight connection to an application \[page 149\]](#)

Related Information

[SAP Applications connection parameters \[page 151\]](#)

[SAP Applications connection parameters \[page 151\]](#)

8.3.2.2 SAP Applications connection parameters

When you set up a connection to an SAP application, you must complete the related parameters in the [Create Connection](#) dialog box in the CMC. To start setting up the connection, define a Data Insight connection to an application.

The table below contains descriptions of the parameters that apply to this type of connection.

Note

The Data Insight module of Information Steward uses the same connections as Data Services to connect to SAP Business Warehouse or SAP ERP Central Component (ECC).

SAP Applications option	Possible values	Description
Server Name	Computer name, fully qualified domain name, or IP address	Name of the remote SAP application computer (host) to which the software connects.
Client Number	000-999	The three-digit SAP client number. Default is 800.
System Number	00-99	The two-digit SAP system number. Default is 00.
User Name	Alphanumeric characters and underscores	Enter the name of the account through which the software accesses the SAP application server.
Password	Alphanumeric characters, underscores, and punctuation	Enter the user's password.
		<div data-bbox="1023 1200 1129 1236" data-label="Section-Header"> <h3>Note</h3> </div> <div data-bbox="1023 1252 1366 1556" data-label="Text"> <p>If you have problems connecting to a system that uses SAP_BASIS version 6.40 and earlier, the system might be expecting an upper-case password. To prevent this issue, install the appropriate kernel patch as described in SAP Note 792850, "Preparing ABAP systems to deal with incompatible passwords."</p> </div>
Application Language	E - English G - German F - French J - Japanese	Select the login language from the drop-down list. You can enter a customized SAP language in this option. For example, you can type S for Spanish or I for Italian.
Code Page		See the section "Supported locales and encodings" in the <i>SAP Data Services Reference Guide</i> .

SAP Applications option	Possible values	Description
<i>ABAP execution option</i>	<p>Execute preloaded</p> <p>Generate and execute</p>	<p>Determines the source for the ABAP program generation.</p> <p><i>Execute preloaded:</i> Runs a previously generated and uploaded ABAP program. Typically, this is used during the production stage of a project.</p> <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 10px; margin: 10px 0;"> <p>Note</p> <p>When you choose <i>Execute Preloaded</i>, you must export ABAP programs used by Information Steward to the SAP system in which the Data Insight view resides so that they can be loaded into the SAP application.</p> </div> <p><i>Generate and execute:</i> Generates a new ABAP program. If you specify this value, specify the <i>Generated ABAP directory</i> option, or leave it blank to use the system default location. Typically, this option is only used in development and test stages of an application project."because in 4.2.3 we're suggesting to use ABAP scripts for large SAP tables.</p>
<i>Execute in background (batch)</i>	Yes, No	<p>Dialog processing or background processing is defined by the user type of the SAP user configured for the communication between Information Steward and the SAP server.</p> <p><i>Yes:</i> Runs in batch mode and does not have a timeout.</p> <p><i>No:</i> Runs in dialog mode and has a timeout.</p>
<i>Job class priority</i>	High, Medium, Low	<p>Sets the order in which jobs from this connection are processed in the SAP system. The default setting is <i>Low</i>. Consider your system and the importance of jobs from this connection compared to the jobs from other defined connections.</p>

SAP Applications option	Possible values	Description
<i>Working directory on SAP server</i>	Directory path	<p>A directory on the SAP application server where the software can write intermediate files. This directory also stores the transport file used by the FTP and shared-directory data transfer methods.</p> <p>By default, the value in this field is blank.</p>
<i>Generated ABAP directory</i>	Directory path	<p>You must specify this directory if any of the following options are selected:</p> <ul style="list-style-type: none"> • If the value for <i>ABAP execution option</i> is <i>Generate and execute</i> • If the value for <i>Data transfer method</i> is <i>RFC</i>
<i>Security Profile</i>	Refer to the requirements of the application	Specify the security profile for Information Steward to read data from the SAP Application.
<i>Number of connection retries</i>	Positive integer	<p>The number of times the software tries to establish a connection with the SAP application server.</p> <p>Defaults to 3.</p>
<i>Interval between retries (sec)</i>	Positive integer	<p>The time delay in seconds between connection retries.</p> <p>Defaults to 10.</p>

SAP Applications option	Possible values	Description
<i>Data transfer method</i>	<p>RFC</p> <p>Shared directory</p> <p>FTP</p>	<p>Define how to retrieve data from the SAP application server to the SAP Data Services server:</p> <p><i>RFC</i>: Use to stream data from the source SAP system directly to the Data Services data flow process using RFC. For secure data transfer, configure <i>SNC</i> authentication with the required SNC quality of protection (in the connection <i>Authentication</i> options). If you select this data transfer method, <i>RFC destination</i> appears.</p> <p><i>Shared directory</i>: Asynchronous processing. Data is written into a file and saved in a shared directory to which the SAP server and Information Steward both have access. If you select this data transfer method, the <i>Application Shared Directory</i> appears.</p> <p><i>FTP</i>: Asynchronous processing. Information Steward (through Data Services) obtains the data from an FTP folder. If you select this data transfer method, the <i>FTP</i> options appear (<i>Local directory</i>, <i>FTP relative path</i>, and so forth).</p> <div data-bbox="1007 1294 1396 1615" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <p>Note</p> <p>If your Data Insight view includes two or more large SAP tables, you can specify the <i>ABAP</i> data transfer method in the <i>View Editor</i> of the Information Steward. web application. For reading large data sets, the <i>ABAP</i> method offers better performance.</p> </div>
<i>RFC destination</i>	SAPDS or <Destination name>	For the <i>RFC</i> data transfer method, enter an RFC destination. You can keep the default name of SAPDS and create a destination of the same name in the source SAP system, or you can enter a destination name for an existing destination.

SAP Applications option	Possible values	Description
<i>Application Shared Directory</i>	Directory path	If you selected the <i>Shared directory</i> data transfer method, specify the SAP application server directory that stores data extracted by SAP Data Services, and to which both Data Services and SAP application server have direct access. After the extraction is completed, Information Steward picks up the file for further processing.
<i>FTP</i>	These options are visible if you selected the <i>FTP</i> data transfer method.	
<i>Local directory</i>	Directory path	If you selected the <i>FTP</i> data transfer method, select a client-side directory to which data from the SAP application server downloads.
<i>FTP relative path</i>	Directory path	Indicate the path from the FTP root directory to the SAP server's working directory. When you select FTP, this directory is required.
<i>FTP host name</i>	Computer (host) name, fully qualified domain name, or IP address	Must be defined to use FTP.
<i>FTP login user name</i>	Alphanumeric characters and underscores	Must be defined to use FTP.
<i>FTP login password</i>	Alphanumeric characters and underscores, or blank	Enter the FTP password.

Parent topic: [Defining a Data Insight connection to an application \[page 149\]](#)

Related Information

[SAP Business Warehouse connection parameters \[page 150\]](#)

[Using SAP applications as a source \[page 306\]](#)

8.3.3 Execute options for ABAP program generation for SAP tables

This topic contains descriptions for the two execute options for generating an ABAP program for Data Insight views with SAP tables.

Administrators choose the source of the ABAP program files while setting up the connection to the SAP table source in the Central Management Console (CMC). There are two options for the *Execute option* parameter in the connection setup dialog box that determine the source for the ABAP program.

Execute options and ABAP program source

Execute option	Source for ABAP program file
<i>Execute preloaded</i>	SAP Business Suite or SAP Business Warehouse application.
<i>Generate and execute</i>	Information Steward

The administrator's selection for the *Execute option* parameter may depend on factors such as the environment (development, testing, production) and the organization's security policies. For more information about ABAP program files, see the *SAP Data Services Supplement for SAP*.

Related Information

[SAP tables and ABAP program migration considerations \[page 333\]](#)

8.3.4 Defining a Data Insight connection to a file

Before you can profile a flat file or Excel spreadsheet file, you must define a connection to the path of the file that will contain the data that you want to examine to determine the quality and structure of the data.

Ensure that the flat file or Excel spreadsheet file meets the following requirements:

- The users running the following services must have permissions on the shared directory where the file resides:
 - Web Application Server (for example, Tomcat)
 - Data Services
 - Service Intelligence Agent
- The file must be located on a shared location that is available to the Data Services Job Server and the View Data Server which are components of SAP Data Services.

Follow these steps to define a Data Insight connection to a file:

1. Open the Information Steward area of the Central Management Console (CMC).
2. Click **Manage > New > Connection**.
The *Create Connection* page opens.
3. Enter the following information in the applicable parameters:

Parameter	Description
<i>Connection Name</i>	Name that you want to use for this data source. Name requirements include: <ul style="list-style-type: none"> • Maximum length is 64 characters • Can be multi-byte • Case insensitive • Can include underscores and spaces • Cannot include other special characters such as ?!@#\$\$%^&*()-+={ } [] : " ; ' ? . ` ~ <p>You cannot change the name after you save the connection.</p>
<i>Description</i>	(Optional) Brief description for the purpose of the connection.

4. In the *Connection Type* drop list, select *File connection*. A file connection type includes both flat files and Excel spreadsheet files.
5. In the Directory path box, enter the path for the file.
6. Click *Save*.

After you create a connection, you must authorize users to perform tasks such as view the data, run profile tasks and run validation rules on the data. For more information see the topic named *User rights in Data Insight*.

Related Information

[User rights in Data Insight \[page 58\]](#)

8.3.5 Displaying and editing Data Insight connection parameters

Situations when you might want to view or change Data Insight connection parameters include:

- You need to view the connection parameters on a development or test system so that you can recreate the Data Insight connection when you move to a production system.
- You have several source systems and you want to ensure that you are connecting to the appropriate source.

To edit connection parameters of your Data Insight project:

1. Ensure that you have *Edit* rights on the connection or you are a member of the Data Insight Administrator group.
2. Go to the Information Steward area of the CMC.
3. Expand the *Data Insight* node in the Tree panel, and select *Connections*.
4. In the list of connections on the right, select the name of your connection, click *Action* in the top menu tool bar, and select *Properties*.
5. Type your changes in the following fields:
 - Description

- Database version
 - User Name
 - Password
 - Language
 - Client Code Page
 - Server Code Page
6. If you want to verify that Information Steward can connect successfully before you save this Data Insight connection, click [Test connection](#).
 7. Click [Save](#).
The edited description appears in the list of projects on the right pane of the [Information Steward](#) page.

8.3.6 Deleting a Data Insight connection

The following table shows the Data Insight objects that you can delete from the Central Management Console (CMC).

For information about dependencies when deleting connections and projects on the Central Management Console (CMC), see “Deleting a connection” and “Deleting a project” in the *Administrator Guide*.

Note

You cannot delete a connection if a table or file is being used in a project. You must remove the table or file from all projects on Information Steward before you can delete the connection or project in the CMC.

Delete object dependents and dependencies

Data Insight object to delete from CMC	Object dependencies that prevent deletion because they are used in another project	Dependent objects that will also be deleted
Profile Task	none	<ul style="list-style-type: none"> • Profile Task Instance
Connection	<ul style="list-style-type: none"> • Referenced project <p>You must delete each table from the Workspace of each project in Data Insight.</p>	<ul style="list-style-type: none"> • Table metadata
Project	Warning message appears, but you can click OK to delete all dependent objects.	<ul style="list-style-type: none"> • Referenced project • Rule • View • Data cleansing solutions

1. Log in to the Central Management Console (CMC) with a user name that belongs to the Data Insight Administrator group or that has the Create right on [Connections](#) in Information Steward.
2. At the CMC home page, click Information Steward.
3. Click the [Connections](#) node in the Tree panel.
4. From the list in the right pane, select the name of the connection and click [Manage](#) [Delete](#).
5. To confirm that you want to delete this connection, click [OK](#) in the warning pop-up window.

6. If the following message appears, you must delete each table from the *Workspace* of each Data Insight project listed in the message.

```
The connection cannot be deleted because it is referenced by the following:  
Project: <projectname>
```

8.3.7 Setting connection type and permissions to access failed data

Administrators determine the connection type on which their users view and export failed data, and they set the specific user permissions for each user in that connection.

You must be an administrator or a Data Insight administrator to perform this task.

Choose to base the failed data on the source or the failed data connection. The default setting is the source data connection.

1. In the Central Management Console (CMC) open *Applications* and select *Information Steward Application*.
2. Select *Configure Application* in the navigation pane at left.
3. Scroll to Permissions in the *Information Steward Settings* page at right.
4. Set the connection option in *Connection on which to base failed data*: Source Connection or Failed Data Connection.
5. Click *Save*.
6. Open the Information Steward page and select *Connections* under the Information Steward node at left.
7. Select the applicable connection. For example, if you chose the failed data connection for the *Access failed data based on* option, select the applicable connection that has the purpose *For data that failed rules*.
8. Select **▸ Manage ▸ User Security ▾** and click *Add Principals*.
9. Select the user or group and click the arrow button to add to the *Selected Users or Groups* list.
10. Click *Add and Assign Security*.
11. Open the *Advanced* tab and click the *Add/Remove Rights* link.
12. Expand *Application* in the navigation pane at left and select *Information Steward Connection*.
13. Select the radio dial under the green check mark icon for each of the following permissions:
 - *Export Data*
 - *View Data*
 - *View Sample Data*
14. Click *OK*.
15. Verify your choices in the *Advanced* tab and click *OK*.

Related Information

[Type-specific rights for source or failed data connections \[page 67\]](#)

8.4 Data Insight projects

You perform all tasks in Data Insight through a project.

The only way to gain access to Data Insight is by selecting a project. Further, you must be granted permission to access a project from your administrator.

A project is a collaborative workspace for data stewards and data analysts. Use a project for the following purposes:

- Access Data Insight
- Monitor data quality of a specific domain
- Monitor data quality for a specific purpose, such as:
 - Customer quality assessment
 - Sales system migration
 - Master data quality

To achieve these purposes, use the tools in Data Insight, such as setting rules for data quality, and configuring a scoring system to indicate good or poor quality.

[Creating a project \[page 160\]](#)

Create a project for Data Insight in the Central Management Console (CMC).

[Editing a project description \[page 161\]](#)

[Enterprise project \[page 162\]](#)

SAP Information Steward provides a special project named Enterprise, which enables the display of Data Insight profiling results and scores in the Metadata Management module in impact and lineage diagrams and in the details for tables and columns.

[Deleting a project \[page 162\]](#)

8.4.1 Creating a project

Create a project for Data Insight in the Central Management Console (CMC).

You must be a member of the Data Insight Administrator group to be able to create projects in the CMC.

1. Log in to the CMC.
2. Select *Information Steward* from the *Organize* list at left.
The Information Steward page opens. There is a file tree in the left pane.
3. Expand the *Data Insight* node and select the *Projects* node in the left pane.

A list of existing projects appears in the main pane.

Note

The **Enterprise Project** is a special project that comes with your *SAP Information Steward* installation.

4. Click  *Manage*  *New*  *Data Insight Project* .

5. On the *New Data Insight Project* page, enter a name for the project in the *Name* box.

Name the project using the following rules:

- Maximum length is 64 characters
- Can be multibyte
- Case insensitive
- Can include underscores and spaces
- Can't include other special characters: ?!@#\$\$%^&*()-+={}[|:~\|.,`~

Note

You can't change the name after you've saved the project.

6. Optional. Enter a description for the project in *Description*.
7. Click *Save*.
The new project appears in the list of projects in the main pane.
8. Highlight the new project and select **► Manage ► User Security**.
9. Select *User Security* in the left pane, and continue configuring the new project by granting access permission to the new project for select users and user groups.

For instructions to grant permissions, see [Assigning users to specific Data Insight objects \[page 69\]](#).

Task overview: [Data Insight projects \[page 160\]](#)

Related Information

[Editing a project description \[page 161\]](#)

[Enterprise project \[page 162\]](#)

[Deleting a project \[page 162\]](#)

[User rights in Data Insight \[page 58\]](#)

8.4.2 Editing a project description

To edit a Data Insight project, you must have *Edit* rights on the project or be a member of the Data Insight Administrator group.

To edit the Description field of your Data Insight project:

1. Log in to the Central Management Console (CMC) with a user name that belongs to the Data Insight Administrator group or that has the Edit right on *Projects* in Information Steward.
2. At the CMC home page, click Information Steward.
3. Expand the *Data Insight* node in the Tree panel, and select *Projects*.
4. In the list of projects on the right, select the name of your project and click **► Action ► Properties**.
5. Type the changes you want in the *Description* field.

6. Click [Save](#) to see the edited description in the list of projects on the right pane of the *Information Steward* page.

Task overview: [Data Insight projects \[page 160\]](#)

Related Information

[Creating a project \[page 160\]](#)

[Enterprise project \[page 162\]](#)

[Deleting a project \[page 162\]](#)

8.4.3 Enterprise project

SAP Information Steward provides a special project named Enterprise, which enables the display of Data Insight profiling results and scores in the Metadata Management module in impact and lineage diagrams and in the details for tables and columns.

You view profile results and data quality scores at the table and column level for tables under the *Data Insight Connection* category in the Metadata Management tab. To enable the display of profiling and rule results on the impact and lineage diagrams, you must run the profiling and rule tasks in the Enterprise Project in Data Insight. For details, see topics “Running a profile task in the Enterprise Project” and “Running a rule task in the Enterprise Project” in the *User Guide*.

Parent topic: [Data Insight projects \[page 160\]](#)

Related Information

[Creating a project \[page 160\]](#)

[Editing a project description \[page 161\]](#)

[Deleting a project \[page 162\]](#)

8.4.4 Deleting a project

To delete a Data Insight project:

1. Log in to the Central Management Console (CMC) with a user name that belongs to the Data Insight Administrator group or that has the Create right on *Projects* in Information Steward.
2. At the CMC home page, click Information Steward.
3. Expand the *Data Insight* node in the Tree panel, and select *Projects*.

4. From the list in the right pane, select the name of the project and click ► **Manage** ► **Delete** ►.
5. To confirm that you want to delete this project, click **OK** in the warning pop-up window.

⚠ Caution

When you delete a project, you delete all the contents of the project which include unapproved rules, tasks, scorecards, profile results, sample data, views, and so forth.

Task overview: [Data Insight projects \[page 160\]](#)

Related Information

[Creating a project \[page 160\]](#)

[Editing a project description \[page 161\]](#)

[Enterprise project \[page 162\]](#)

8.5 Data Insight tasks

Create Data Insight tasks to run immediately in SAP Information Steward or create a schedule for a recurring task in the Central Management Console (CMC).

Create and run tasks for the following Data Insight task types:

- Profiles
- Rules
- Calculate Scorecard utility

Additionally, export task execution commands to run Data Insight tasks using an external scheduler.

[Scheduling a task in the CMC \[page 164\]](#)

Schedule tasks to occur once, on a recurring schedule, or based on an event in the Central Management Console (CMC).

[Export tasks for external scheduler \[page 166\]](#)

To use an external scheduler to run tasks, export generated commands from the Central Management Console (CMC) or execute a command utility.

[Recurrence options \[page 177\]](#)

Choose the frequency to run a scheduled task.

[Notification for task completion \[page 178\]](#)

Send notifications to specified users and groups when a task completes.

[Configuring notification when scheduling a task \[page 179\]](#)

Set up users and groups to receive task completion notifications for specific Information Steward profile and rule tasks that are scheduled to run.

[Rule threshold notification \[page 181\]](#)

[Capturing skipped records from rule filter \[page 182\]](#)

Configure your rule task to capture records that are skipped because of rule filters and to return information about skipped rows and columns.

[Monitoring a task \[page 183\]](#)

[Pausing and resuming a schedule \[page 185\]](#)

[Deleting failed data for a failed rule task \[page 186\]](#)

You can choose to have the software delete conflicting data from a failed data repository under certain circumstances.

[Common runtime parameters for Information Steward \[page 187\]](#)

Descriptions of the common runtime parameters that apply to all metadata integrators, profile tasks, and rule tasks.

8.5.1 Scheduling a task in the CMC

Schedule tasks to occur once, on a recurring schedule, or based on an event in the Central Management Console (CMC).

To perform these steps, you must be an administrator, or a user with the applicable permissions.

1. Log in to the CMC and open the *Information Steward* page.
The *Information Steward* page opens with the *Information Steward* node selected in the pane at left.
2. Expand ► *Data Insight* ► *Projects* ▾ in the left pane, and select the name of your project.
A list of tasks for the selected project appears at right.
3. Click the applicable task and select ► *Action* ► *Schedule* ▾.
The *Schedule* page opens with the *Schedule* node expanded in the left navigation pane.
4. Complete the parameters for each option as applicable.

Navigation option descriptions

Node	Description
<i>Parameters</i>	Set both database and file log levels. Enter JVM arguments and any additional runtime arguments.
<i>Events</i>	Set a schedule to run the task based on already-scheduled events. For more information about <i>Events</i> , see the <i>SAP Business Intelligence Platform Administrator Guide</i> .
<i>Instance Title</i>	Keep the original task name, or enter a different name to specify this particular schedule.

Node	Description
<i>Notification</i>	<p>Set options so the system sends you e-mail notifications when the job runs successfully and when it fails.</p> <div data-bbox="821 436 1402 801" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <p>Note</p> <p>Before Information Steward sends out e-mail notifications, administrators must add the user or user group recipient e-mail information in the user or group settings. For information, see Setting user e-mail address for notifications [page 57]. Administrators must also configure the EIM Adaptive Job Server Destination e-mail settings. For more information, see Setting up e-mail notification configuration [page 100]</p> </div>
<i>Recurrence</i>	<p>Set the interval of time in which to run the task. For example, select to run the task each hour, once a week, and so on.</p> <p>Optionally, set the <i>Number of retries</i> to a value other than the default of 0 , and change the <i>Retry interval in seconds</i> from the default value of 1800.</p>

5. Click *Schedule*.

Task overview: [Data Insight tasks \[page 163\]](#)

Related Information

- [Export tasks for external scheduler \[page 166\]](#)
- [Recurrence options \[page 177\]](#)
- [Notification for task completion \[page 178\]](#)
- [Configuring notification when scheduling a task \[page 179\]](#)
- [Rule threshold notification \[page 181\]](#)
- [Capturing skipped records from rule filter \[page 182\]](#)
- [Monitoring a task \[page 183\]](#)
- [Pausing and resuming a schedule \[page 185\]](#)
- [Deleting failed data for a failed rule task \[page 186\]](#)
- [Common runtime parameters for Information Steward \[page 187\]](#)
- [Pausing and resuming a schedule \[page 185\]](#)
- [Recurrence options \[page 177\]](#)
- [Common runtime parameters for Information Steward \[page 187\]](#)

8.5.2 Export tasks for external scheduler

To use an external scheduler to run tasks, export generated commands from the Central Management Console (CMC) or execute a command utility.

You must be an administrator or a user with scheduling privileges with the right to track activities to export tasks from the CMC.

Some users prefer to use an external scheduler to run all tasks, including their Information Steward tasks. To use an external scheduler, users export task commands and copy them to a batch file that runs their external scheduler.

There are two ways to export tasks from the CMC for an external scheduler:

- Generate a command in the CMC
- Set up and run the `ISTaskExecuteCommand.bat` or `.sh` utility.

You can export task execution commands for the following objects:

- Profile tasks
- Rule tasks
- All profile and rule tasks in a project
- Scorecard calculation

Additionally, you can view task history and download the log file from either Information Steward or the CMC.

[Exporting execution commands for profile and rule tasks \[page 167\]](#)

Use the Central Management Console (CMC) to generate execution commands for profile and rule tasks. Then copy the commands to a batch file to run tasks with an external scheduler.

[Exporting execution commands for calculate scorecard utility \[page 168\]](#)

Use the Central Management Console (CMC) to generate execution commands for the Calculate Scorecard utility. Then copy the commands to a batch file to run tasks in an external scheduler.

[Exporting execution commands for tasks in a project \[page 169\]](#)

Use the Central Management Console (CMC) to generate execution commands for all tasks in a project. Then copy the commands to a batch file to run tasks in an external scheduler.

[ISTaskExecuteCommand utility for Windows and Linux \[page 170\]](#)

Run the task execute command utility for Windows or Linux to schedule tasks on an external scheduler.

[Common parameters for export execution commands \[page 172\]](#)

The generated export execution command (generated command) from the Central Management Console (CMC) contains common parameters that you can also use when you create the `ISTaskExecuteCommand.bat` or `.sh` utility (command utility).

[Include scores in execution commands \[page 174\]](#)

When you export task commands to run rule tasks or calculate scorecards with an external scheduler, you can set up your commands to include scores.

[Including scores in rule task execution command \[page 175\]](#)

Include a score parameter when you generate code to export a task to an external scheduler.

[Including scores in execution commands for calculate scorecard \[page 176\]](#)

Include a score parameter when you generate code to export a calculate scorecard task to an external scheduler.

Parent topic: [Data Insight tasks \[page 163\]](#)

Related Information

- [Scheduling a task in the CMC \[page 164\]](#)
- [Recurrence options \[page 177\]](#)
- [Notification for task completion \[page 178\]](#)
- [Configuring notification when scheduling a task \[page 179\]](#)
- [Rule threshold notification \[page 181\]](#)
- [Capturing skipped records from rule filter \[page 182\]](#)
- [Monitoring a task \[page 183\]](#)
- [Pausing and resuming a schedule \[page 185\]](#)
- [Deleting failed data for a failed rule task \[page 186\]](#)
- [Common runtime parameters for Information Steward \[page 187\]](#)

8.5.2.1 Exporting execution commands for profile and rule tasks

Use the Central Management Console (CMC) to generate execution commands for profile and rule tasks. Then copy the commands to a batch file to run tasks with an external scheduler.

To perform these steps, you must be a user with administrator rights, or a user with scheduling privileges and rights to track activities.

1. Log in to the CMC and open the Information Steward page.

The *Information Steward* page opens with the *Information Steward* node selected in the pane at left.

2. Expand **Data Insight** **Projects** in the tree pane and then select the project name.

All Data Insight tasks appear in the right pane.

3. Right-click the applicable profile task or rule task and select *Export Execution Command*.

The *Export* window opens.

4. Optional. Enter applicable JVM arguments in the *JVM Arguments* text entry field.
5. Optional. Enter applicable additional arguments in the *Additional Arguments* text entry field.
6. Click *Generate*.

The applicable command line parameters appear in the *Command Line Parameters* box. For example, here is a command without JVM arguments or additional arguments:

Sample Code

```
task -configurationTaskObject 31376 boeUser <yourusername> boePassword  
<yourpassword> boeServer <yourserver>
```

Here is a command that includes JVM arguments:

Sample Code

```
task -jvmArgs "-Xms64m -Xmx1024m" -configurationTaskObject 40786 boeUser  
<yourusername> boePassword <yourpassword> boeServer <yourserver>
```

7. Change the following variables from the command with your information:
 - yourusername
 - yourpassword
 - yourserver
8. Copy and paste the completed command into the applicable batch file.

Related Information

[Export tasks for external scheduler \[page 166\]](#)

8.5.2.2 Exporting execution commands for calculate scorecard utility

Use the Central Management Console (CMC) to generate execution commands for the Calculate Scorecard utility. Then copy the commands to a batch file to run tasks in an external scheduler.

To perform these steps, you must be a user with administrator rights, or a user with scheduling privileges and rights to track activities.

1. Log in to the CMC and select *Applications*.
2. Double-click *Information Steward Application*.

The *Information Steward Settings* page opens.

3. Select *Manage Utilities* from the navigation pane at left.
The *Utility Configurations* pane opens at right with a list of configured utilities.
4. Select *Calculate Scorecard*.
5. Select **► Actions ► Export Execution Command ►**.

The *Export* window opens.

6. Optional. Enter applicable JVM arguments in the *JVM Arguments* text entry field.
7. Optional. Enter applicable additional arguments in the *Additional Arguments* text entry field.
8. Click *Generate*.

The applicable command line parameters appear in the *Command Line Parameters* box. For example, here is a command without JVM arguments or additional arguments:

Sample Code

```
scorecard boeUser <yourusername> boePassword <yourpassword> boeServer  
<yourserver>
```

9. Change the following variables from the command with your information:

- yourusername
- yourpassword
- yourserver

10. Copy and paste the generated command into your batch file.

Related Information

[Export tasks for external scheduler \[page 166\]](#)

8.5.2.3 Exporting execution commands for tasks in a project

Use the Central Management Console (CMC) to generate execution commands for all tasks in a project. Then copy the commands to a batch file to run tasks in an external scheduler.

To perform these steps, you must be a user with administrator rights, or a user with scheduling privileges and rights to track activities.

1. Log in to the CMC and open the Information Steward page.

The *Information Steward* page opens with the *Information Steward* node selected in the pane at left.

2. Expand **► Data Insight ► Projects ▾**.

3. Select the project name in the right pane.

4. Select **► Actions ► Export Execution Commands ▾**.

The *Export* window opens.

5. Optional. Enter applicable JVM arguments in the *JVM Arguments* text entry field.

6. Optional. Enter applicable additional arguments in the *Additional Arguments* text entry field.

7. Click *Generate*.

The software assigns a `-configurationTaskObject` identification number to each task in the project, and lists them in the *Command Line Parameters* text box. For example, here is a list of commands for each task in a project without JVM arguments or additional arguments:

Sample Code

```
task -configurationTaskObject 30711 boeUser <yourusername> boePassword
<yourpassword> boeServer <yourserver>
task -configurationTaskObject 30769 boeUser <yourusername> boePassword
<yourpassword> boeServer <yourserver>
task -configurationTaskObject 30886 boeUser <yourusername> boePassword
<yourpassword> boeServer <yourserver>
task -configurationTaskObject 30955 boeUser <yourusername> boePassword
<yourpassword> boeServer <yourserver>
task -configurationTaskObject 31085 boeUser <yourusername> boePassword
<yourpassword> boeServer <yourserver>
task -configurationTaskObject 31108 boeUser <yourusername> boePassword
<yourpassword> boeServer <yourserver>
task -configurationTaskObject 31132 boeUser <yourusername> boePassword
<yourpassword> boeServer <yourserver>
...
```

8. Change the following variables for each command with your information:
 - yourusername
 - yourpassword
 - yourserver
9. Copy and paste the generated commands into your batch file.

Related Information

[Export tasks for external scheduler \[page 166\]](#)

8.5.2.4 ITaskExecuteCommand utility for Windows and Linux

Run the task execute command utility for Windows or Linux to schedule tasks on an external scheduler.

When you use the `ITaskExecuteCommand`, you do not need to generate a command in the Central Management Console (CMC) to obtain an internal task identification number. `ITaskExecuteCommand` only requires object and task names instead of generated task identification numbers.

Note

The `ITaskExecuteCommand` utility is in the `MM\bin` folder of the SAP Information Steward installation directory.

For example:

Sample Code

```
ISTaskExecuteCommand.bat task -configurationProjectObjectName  
<yourprojectname> -configurationTaskObjectName <yourtaskname> boeUser  
<yourusername> boePassword <your password> boeServer <yourserver>
```

You can reuse the same `ISTaskExecuteCommand` with the task name and project name for each run of that task. You can even reuse the command with that task name and project name when you move from one environment to another because the project name and task name do not change from one environment to the next.

There is one exception. When you export all tasks in a project, you use the `-configurationTaskObject` identification number for each task in the `ISTaskExecuteCommand` utility. Therefore, if your task object identification numbers have changed, which could happen when you move to a new environment, regenerate the command in the CMC. You can either use the newly regenerated command with the new identification numbers, or update your existing `ISTaskExecuteCommand` utility for the project with the new identification numbers. The advantages and disadvantages for using the generated command or the utility are as follows:

- If you use the newly generated command from the CMC, you re-enter all of your information, such as user name and password, for each task. You could copy and paste this information to make this process less tedious.
- If you use your existing `ISTaskExecuteCommand`, you replace all task identification numbers with the newly generated identification numbers from the CMC. If none of your information has changed, such as user name and password, you can reuse that information.

The following examples use the Windows `ISTaskExecuteCommand.bat` utility for Windows. Substitute the `.sh` version of the command if you are working in a Linux environment.

Example: Export parameters to run a given named task

Sample Code

```
ISTaskExecuteCommand.bat task -configurationProjectObjectName  
<yourprojectname> -configurationTaskObjectName <yourtaskname> boeUser  
<yourusername> boePassword <your password> boeServer <yourserver>
```

Example: Export parameters for calculate scorecard utility

Sample Code

```
ISTaskExecuteCommand.bat scorecard boeUser <yourusername> boePassword  
<yourpassword> boeServer <yourserver>
```

Example: Export parameters for a project

Sample Code

```
ISTaskExecuteCommand.bat task -configurationTaskObject 18210 boeUser  
<yourusername> boePassword <yourpassword> boeServer <yourserver>  
ISTaskExecuteCommand.bat task -configurationTaskObject 18211 boeUser  
<yourusername> boePassword <yourpassword> boeServer<yourserver>  
ISTaskExecuteCommand.bat task -configurationTaskObject 18212 boeUser  
<yourusername> boePassword <yourpassword> boeServer <yourserver>  
...
```

Example: Export parameters with JVM parameters and additional arguments

Sample Code

```
ISTaskExecuteCommand.bat task -jvmArgs "-Xms64m -Xmx1024m"  
-configurationTaskObject 40786 boeUser <yourusername> boePassword  
<yourpassword> boeServer <yourserver>
```

Related Information

[Include scores in execution commands \[page 174\]](#)

8.5.2.5 Common parameters for export execution commands

The generated export execution command (generated command) from the Central Management Console (CMC) contains common parameters that you can also use when you create the `ISTaskExecuteCommand.bat` or `.sh` utility (command utility).

The following table contains information about the common parameters contained in either the generated command or the utility.

Parameter	Description	Use
<code>-configurationTaskObject</code>	<p>Unique system-generated numeric identification code. The ID changes each time you generate the command in the CMC.</p> <div data-bbox="603 510 991 680" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <p>❖ Example</p> <pre>-configurationTaskObject 18210</pre> </div>	<p>Automatically included in the generated command.</p> <p>Optional or required in the <code>IStaskExecuteCommand</code> based on the following circumstances:</p> <ul style="list-style-type: none"> • Optional when you include the <code>configurationTaskObjectName</code> (task name) and the <code>configurationProjectObjectName</code> (project name) parameters. • Required when you do not include the task name and project name parameters. • Takes precedence over the task name and project name parameters when they are included in the command. The software only uses the <code>configurationTaskObject</code> identification number to identify the task, and ignores the task name and the project name parameters. <div data-bbox="1054 1265 1396 1704" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <p>⚠ Caution</p> <p>The task fails if the <code>configurationTaskObject</code> identification number has changed since the last time you used the <code>IStaskExecuteCommand</code>, even when the task name and the project name parameters are included in the command and are still correct.</p> </div>
<code>boeUser</code>	User name that you use to log in to the CMC	Required
<code>boePassword</code>	Password that corresponds with the <code>boeUser</code> name	Required

Parameter	Description	Use
<code>boeServer</code>	Server name connected with the Information Steward installation	Required
<code>-configurationTaskObjectName</code>	Name of the task being exported	Required if you do not include the <code>-configurationTaskObject</code> identification number parameter
<code>-configurationProjectObjectName</code>	Name of the project being exported	Required if you do not include the <code>-configurationTaskObject</code> identification number parameter
<code>-jvmArgs</code>	Java Machine arguments	Optional
<code>-programArgs</code>	Additional arguments	Optional
<code>-sync</code>	<p>True: Software completes the task execution in the correct runtime order, and the software does not exit until the task completes.</p> <p>False: Default value. Software starts the task but does not wait for the completion before it goes on to another task.</p>	Optional
<code>-timeout</code>	<p>Specifies a timeout value in minutes.</p> <p>When you specify True for <code>-sync</code> in a command line argument, set <code>-timeout</code> to the maximum amount of time in minutes required for any task to complete in this environment.</p> <p>If you don't set a value for <code>-timeout</code>, the default maximum execution time is 60 minutes.</p>	Optional

8.5.2.6 Include scores in execution commands

When you export task commands to run rule tasks or calculate scorecards with an external scheduler, you can set up your commands to include scores.

The software returns a score when the task completes successfully and the task involves one table or view. The score is in the form of an exit code.

The software translates return scores to a range of large negative values to ensure that there is no overlap between existing error codes and return scores.

Note

Error codes take precedence over return codes. The software considers the highest error code to be the most severe.

The software returns a score if you set the `-returnScore` parameter to true in `-programArgs`. The following example shows a generated command with the return score parameter set to true:

Sample Code

```
task -programArgs "-returnScore true" -configurationTaskObject 37801 boeUser  
<yourusername> boePassword <yourpassword> boeServer <yourserver>
```

The software does not return a score under the following circumstances:

- The task involves more than one table or view
- The task returns multiple scores. For example, a task returns multiple scores if a scorecard has more than one key data domain (KDD), or a rule binds to multiple tables.

8.5.2.7 Including scores in rule task execution command

Include a score parameter when you generate code to export a task to an external scheduler.

To perform these steps, you must be a user with administrator rights, or a user with scheduling privileges and rights to track activities.

1. Log in to the CMC and open the Information Steward page.

The *Information Steward* page opens with the *Information Steward* node selected in the pane at left.

2. Expand **Data Insight** > **Projects**.
3. Select the project name in the left pane and then select the applicable task in the right pane.
4. Select **Actions** > **Export Execution Commands**.

The *Export* window opens.

5. Optional. Enter applicable JVM In the *JVM Arguments* field.
6. Enter the following parameter and value in the *Additional Arguments* text entry field: `-returnScore true`.
7. Click *Generate*.

Here is an example of a generated command with return score set to true:

Sample Code

```
task -programArgs "-returnscore true" -configurationTaskObject 34814  
boeUser <yourusername> boePassword <yourpassword> boeServer <yourserver>
```

Related Information

[Include scores in execution commands \[page 174\]](#)

8.5.2.8 Including scores in execution commands for calculate scorecard

Include a score parameter when you generate code to export a calculate scorecard task to an external scheduler.

To perform these steps, you must be a user with administrator rights, or a user with scheduling privileges and rights to track activities.

1. Log in to the CMC and select *Applications*.
2. Double-click *Information Steward Application*.

The *Information Steward Settings* page opens.

3. Select *Manage Utilities* in the navigation pane at left.
The *Utility Configurations* window opens at right with a list of configured utilities.
4. Select *Calculate Scorecard*.
5. Select **▸ Actions ▸ Export Execution Command ▾**.

The *Export* window opens.

6. Optional. Enter applicable JVM In the *JVM Arguments* field.
7. Enter the following parameter and value in the *Additional Arguments* text entry field: **-returnScore true**.
8. Click *Generate*.

Here is an example of a generated command with return score set to true:

Sample Code

```
scorecard -programArgs "-returnscore true" boeUser <yourusername>  
boePassword <yourpassword> boeServer <yourserver>
```

Related Information

[Include scores in execution commands \[page 174\]](#)

8.5.3 Recurrence options

Choose the frequency to run a scheduled task.

When you schedule an integrator source or an SAP Information Steward utility, you can choose the frequency to run it in the *Recurrence* option. The following table describes each recurrence option and shows the additional relevant values that you must select for each recurrence option.

Recurrence options

Recurrence Option	Description
<i>Now</i>	The utility will run when you click <i>Schedule</i> .
<i>Once</i>	The utility will run once only. Select the values for <ul style="list-style-type: none"><i>Start Date/Time</i><i>End Date/Time</i>
<i>Hourly</i>	The utility will run every N hours and X minutes. Select the values for: <ul style="list-style-type: none"><i>Hour(N)</i><i>Minute(X)</i><i>Start Date/Time</i><i>End Date/Time</i>
<i>Daily</i>	The utility will run once every N days. Select the value for <i>Days(N)</i>
<i>Weekly</i>	The utility will run once every week on the selected days. Select the following values: <ul style="list-style-type: none">Days of the week<i>Start Date/Time</i><i>End Date/Time</i>
<i>Monthly</i>	The utility will run once every N months. Select the following values: <ul style="list-style-type: none"><i>Month(N)</i><i>Start Date/Time</i><i>End Date/Time</i>
<i>Nth Day of Month</i>	The utility will run on the Nth day of each month. Select the following values: <ul style="list-style-type: none"><i>Day(N)</i><i>Start Date/Time</i><i>End Date/Time</i>
<i>1st Monday of Month</i>	The utility will run on the first Monday of each month. Select the following values: <ul style="list-style-type: none"><i>Start Date/Time</i><i>End Date/Time</i>
<i>Last Day of Month</i>	The utility will run on the last day of each month. Select the following values: <ul style="list-style-type: none"><i>Start Date/Time</i><i>End Date/Time</i>

Recurrence Option	Description
<i>X Day of Nth Week of the Month</i>	The utility will run on the X day of the Nth week of each month. Select the following values: <ul style="list-style-type: none"> <i>Week(N)</i> <i>Day(X)</i> <i>Start Date/Time</i> <i>End Date/Time</i>
<i>Calendar</i>	The utility will run on the days you specified as "run" days on a calendar you have created in the <i>Calendars</i> management area of the CMC.

Parent topic: [Data Insight tasks \[page 163\]](#)

Related Information

- [Scheduling a task in the CMC \[page 164\]](#)
- [Export tasks for external scheduler \[page 166\]](#)
- [Notification for task completion \[page 178\]](#)
- [Configuring notification when scheduling a task \[page 179\]](#)
- [Rule threshold notification \[page 181\]](#)
- [Capturing skipped records from rule filter \[page 182\]](#)
- [Monitoring a task \[page 183\]](#)
- [Pausing and resuming a schedule \[page 185\]](#)
- [Deleting failed data for a failed rule task \[page 186\]](#)
- [Common runtime parameters for Information Steward \[page 187\]](#)

8.5.4 Notification for task completion

Send notifications to specified users and groups when a task completes.

To ensure that notifications are configured correctly, perform the following checks in the Central Management Console (CMC):

- Open the Users and Groups page in the CMC and view the users or groups that you plan to include in the task completion notification. For more information, see [Setting user e-mail address for notifications \[page 57\]](#).
- Open the Servers page in the CMC, select the Servers List node and open the EIM Adaptive Job Server, and view the properties. For more information, see [Setting up e-mail notification configuration \[page 100\]](#).
- Complete the e-mail notification settings when you schedule a task. Settings that you make in the task schedule override the comparable settings in the EIM Adaptive Job Server properties. For more information, see [Configuring notification when scheduling a task \[page 179\]](#).

Parent topic: [Data Insight tasks \[page 163\]](#)

Related Information

- [Scheduling a task in the CMC \[page 164\]](#)
- [Export tasks for external scheduler \[page 166\]](#)
- [Recurrence options \[page 177\]](#)
- [Configuring notification when scheduling a task \[page 179\]](#)
- [Rule threshold notification \[page 181\]](#)
- [Capturing skipped records from rule filter \[page 182\]](#)
- [Monitoring a task \[page 183\]](#)
- [Pausing and resuming a schedule \[page 185\]](#)
- [Deleting failed data for a failed rule task \[page 186\]](#)
- [Common runtime parameters for Information Steward \[page 187\]](#)

8.5.5 Configuring notification when scheduling a task

Set up users and groups to receive task completion notifications for specific Information Steward profile and rule tasks that are scheduled to run.

Before you complete the following steps, check that you meet the following prerequisites:

- You have the required permissions to perform the task.
- The recipient users or user groups have their e-mail addresses in the users and groups configuration. For information, see [Setting user e-mail address for notifications \[page 57\]](#)
- Your administrator has configured the EIM Adaptive Job Server destination e-mail with the SMTP domain, server address, and port information. For information, see [Setting up e-mail notification configuration \[page 100\]](#).

1. Log in to the Central Management Console (CMC).

The *CMC Home* page opens.

2. Click *Information Steward* from the *Organize* column.
3. In the Information Steward page, expand the Data Insight node and then the Projects node.

The CMC displays a list of tasks for the specific project that you selected.

4. Select the task from the list and select **► Actions ► Schedule ►** from the menu.
5. Choose *Notification* from the navigation list at left.
6. Expand the e-mail Notification node at right.
7. Select one or both options: *A job ran successfully* or *A job failed to run*.

These options determine what activates the e-mail notification.

8. Expand the option node that you chose.
The following table describes the options to configure the e-mail notification.

Option	Description
Use Job Server defaults	Select to use the destination e-mail settings for the EIM Adaptive Job Server. When you select Use Job Server defaults , the options under Set the values to be used here become inactive.
Set values to be used here	Select to add e-mail values.

Note

These settings override any comparable e-mail settings that are defined in the EIM Adaptive Job Server. For example, the value that you enter in [From](#), overrides the value for [From](#) in the EIM Adaptive Job Server e-mail settings.

Options that appear when you select [Set values to be used here](#):

From	Enter the e-mail address of the user or group sending the e-mail notification. Overrides the From setting in the EIM Adaptive Job Server.
To	Enter the e-mail address of the users or groups who subscribe to receive notifications. Overrides the To setting in the EIM Adaptive Job Server.
CC	Enter the e-mail address of the users or groups who receive carbon copies of the notification. Overrides the CC setting in the EIM Adaptive Job Server.
Subject	Specify the default subject heading used in e-mails containing system alerts. Leave blank to use the default subject from the EIM Adaptive Job Server.
Message	Specify the default message to include in e-mails containing system alerts. Leave blank to use the default message from the EIM Adaptive Job Server.

- Click [Schedule](#).

Task overview: [Data Insight tasks \[page 163\]](#)

Related Information

- [Scheduling a task in the CMC \[page 164\]](#)
- [Export tasks for external scheduler \[page 166\]](#)
- [Recurrence options \[page 177\]](#)
- [Notification for task completion \[page 178\]](#)
- [Rule threshold notification \[page 181\]](#)
- [Capturing skipped records from rule filter \[page 182\]](#)
- [Monitoring a task \[page 183\]](#)
- [Pausing and resuming a schedule \[page 185\]](#)

[Deleting failed data for a failed rule task \[page 186\]](#)

[Common runtime parameters for Information Steward \[page 187\]](#)

8.5.6 Rule threshold notification

The Data Insight Administrator or someone in the Administrator group can configure the server to provide email notifications in the Central Management Console (CMC). A Data Insight user must have already created the profile task or rule task on Information Steward (for more information, see the *User Guide*).




Note

For more information about rule threshold e-mail notifications, see the *User Guide*.

When you create a rule task, you can also create email notifications that will alert business users when a rule score meets a specific sensitivity setting. You can send the notifications to an email address, a CMC user, a CMC group, or any combination of these. You can set up multiple email notifications for a single rule task.

You must configure the notification server before processing the task so the server has the correct information to send in the notification. Email notifications will be sent when the task is complete

To configure the notification server to receive processing information:

1. In the CMC Home window, click [Servers](#).
2. Select [Servers List](#) and then highlight the ISJobServer. If you have more than one job server available (in addition to the ISJobServer), you must configure all of them.
3. Choose  [Manage](#)  [Properties](#)  and then select [Destination](#) in the navigation pane.
4. Select [Email](#) from the [Destination](#) drop-down list and then click [Add](#).
5. To set up a notification server for rules, you must enter information into the Domain, Host, Port and From fields. All other fields are optional. The following table describes the fields on the Destination page:

Option	Description
Domain (required)	Enter the fully qualified domain of the SMTP server.
Host (required)	Enter the name of the SMTP server.
Port (required)	Enter the port that the SMTP server is listening on. (This standard SMTP port is 25.)
Authentication	Select Plain or Login if the job server must be authenticated using one of these methods in order to send email.
User Name	Provide the Job Server with a user name that has permission to send email and attachments through the SMTP server.
Password	Provide the Job Server with the password for the SMTP server.
From (required)	Provide the return email address. Users can override this default when they schedule an object.
To	Not used in this scenario. The email address specified when creating the task is used.
CC	Not used in this scenario.

Option	Description
Subject	Not used in this scenario.
Message	Not used in this scenario.
Add placeholder	You can add placeholder variables to the message body using the Add placeholder list. For example, you can add the report title, author, or the URL for the viewer in which you want the email recipient to view the report.
Add Attachment	Not used in this scenario.

6. Click [Save](#).

Task overview: [Data Insight tasks \[page 163\]](#)

Related Information

- [Scheduling a task in the CMC \[page 164\]](#)
- [Export tasks for external scheduler \[page 166\]](#)
- [Recurrence options \[page 177\]](#)
- [Notification for task completion \[page 178\]](#)
- [Configuring notification when scheduling a task \[page 179\]](#)
- [Capturing skipped records from rule filter \[page 182\]](#)
- [Monitoring a task \[page 183\]](#)
- [Pausing and resuming a schedule \[page 185\]](#)
- [Deleting failed data for a failed rule task \[page 186\]](#)
- [Common runtime parameters for Information Steward \[page 187\]](#)
- [Configuring notification when scheduling a task \[page 179\]](#)

8.5.7 Capturing skipped records from rule filter

Configure your rule task to capture records that are skipped because of rule filters and to return information about skipped rows and columns.

Prerequisites:

To be able to capture skipped records from a rule filter, make sure to include the following in the rule editor:

- Select a failed data connection from the [Save all failed data to](#) list.
- Select the option [Include source](#).
- Configure the rule with a filter.

Also, you must bind the rule to at least one column in a source.

To configure a rule task to capture skipped records from a rule filter, perform the following steps:

1. Log in to the Central Management Console and open the [Information Steward](#) page.

2. Expand **Data Insight > Projects** at left.
3. Select the applicable project name.
A list of tasks for the selected project appears at right.
4. Double-click the applicable rule task in the task list.
5. Expand **Schedule > Parameters** at left.
The *Parameters* pane opens at right.
6. Enter the argument `Task.captureRuleFilteredRows true` in *Additional Arguments*.
7. Complete other options as necessary and save.

Results:

Data Insight adds the following files to the failed data repository when you run the rule task:

- `<TABLE_ALIAS>_SR`: Stores all information about skipped columns.
- `<TABLE_ALIAS>_SD`: Stores all information about the skipped rows and columns.

Task overview: [Data Insight tasks \[page 163\]](#)

Related Information

- [Scheduling a task in the CMC \[page 164\]](#)
- [Export tasks for external scheduler \[page 166\]](#)
- [Recurrence options \[page 177\]](#)
- [Notification for task completion \[page 178\]](#)
- [Configuring notification when scheduling a task \[page 179\]](#)
- [Rule threshold notification \[page 181\]](#)
- [Monitoring a task \[page 183\]](#)
- [Pausing and resuming a schedule \[page 185\]](#)
- [Deleting failed data for a failed rule task \[page 186\]](#)
- [Common runtime parameters for Information Steward \[page 187\]](#)

8.5.8 Monitoring a task

To look at the status of or the progress of a profile task or rule task:

1. Login to the Central Management Console (CMC) with a user name that belongs to the Data Insight Administrator group or Administrator group.
2. At the CMC home page, click Information Steward.
The *Information Steward* page opens with the *Information Steward* node selected in the Tree panel.
3. In the Tree panel, expand the *Data Insight* node expand the *Projects* node.
4. Select the name of your project in the Tree panel.
A list of tasks appears in the right panel with the date and time each was last run.

5. To update the *Last Run Status* and *Last Run* columns, click the *Refresh* icon.
6. To view the history of a task, select its name and click **▶ Action ▶ History ▾** in the top menu tool bar. The *Data Insight Task history* pane displays each instance the task was executed with the status, start time end time, and duration. The *Schedule Status* column can contain the following values:

Schedule Status	Description
Failed	Task did not complete successfully.
Pending	Task is scheduled to run one time. When it actually runs, the status changes to "Running".
Recurring	Task is scheduled to recur. When it actually runs, there will be another instance with status "Running".
Running	Task is currently executing.
Success	Task completed successfully.

7. To see the progress of a task instance:
 - a. Select the instance name and click the icon for *View the database log* in the top menu bar of the *Data Insight Task History* page. The *Database Log* window shows the task messages.

ⓘ Note

The *Database Log* shows a subset of the messages in the log file.

- b. To find specific messages in the *Database Log* window, enter a string in the text box and click *Filter*. For example, you might enter **error** to see if there are any errors.
 - c. To close the *Database Log* window, click the *X* in the upper right corner.
8. To save a copy of a task log:
 - a. Scroll to the right of the *Data Insight Task History* page, and click the *Download* link in the *Log File* column in the row of the utility instance you want.
 - b. Click *Save*.
 - c. On the *Save As* window, browse to the directory where you want to save the log and change the default file name if you want.

ⓘ Note

This downloaded log file contains more messages than the *Database Log* because its default logging level is set lower.

9. To close the *Data Insight Task History* page, click the *X* in the upper right corner.

Task overview: [Data Insight tasks \[page 163\]](#)

Related Information

[Scheduling a task in the CMC \[page 164\]](#)

[Export tasks for external scheduler \[page 166\]](#)
[Recurrence options \[page 177\]](#)
[Notification for task completion \[page 178\]](#)
[Configuring notification when scheduling a task \[page 179\]](#)
[Rule threshold notification \[page 181\]](#)
[Capturing skipped records from rule filter \[page 182\]](#)
[Pausing and resuming a schedule \[page 185\]](#)
[Deleting failed data for a failed rule task \[page 186\]](#)
[Common runtime parameters for Information Steward \[page 187\]](#)
[Scheduling a task in the CMC \[page 164\]](#)
[Information Steward logs \[page 340\]](#)
[Changing Metadata Management log levels \[page 341\]](#)

8.5.9 Pausing and resuming a schedule

You can pause a recurring schedule for a task when you do not want to run it at its regularly scheduled time until you resume the schedule.

To pause a schedule:

1. Log in to the CMC with a user name that belongs to the Data Insight Administrator group or that has the Create right on *Projects* in Information Steward.
2. At the CMC home page, click Information Steward.
The *Information Steward* page opens with the *Information Steward* node selected in the Tree panel.
3. In the Tree panel, expand the *Data Insight* node expand the *Projects* node.
4. Select the name of your project in the Tree panel.
5. Select the name of the task from the list on the right panel.
6. Click **▶ Action ▶ History ▾** in the top menu tool bar.
A list of instances for the task appears in the right pane.
7. Select the task instance that has the value "Recurring" in the *Schedule Status* column and click **▶ Action ▶ Pause ▾** in the top menu tool bar.
8. When you are ready to resume the recurring schedule, select the task instance that has the value "Paused" in the *Schedule Status* column and click **▶ Action ▶ Resume ▾** in the top menu tool bar.

Task overview: [Data Insight tasks \[page 163\]](#)

Related Information

[Scheduling a task in the CMC \[page 164\]](#)
[Export tasks for external scheduler \[page 166\]](#)
[Recurrence options \[page 177\]](#)

- [Notification for task completion \[page 178\]](#)
- [Configuring notification when scheduling a task \[page 179\]](#)
- [Rule threshold notification \[page 181\]](#)
- [Capturing skipped records from rule filter \[page 182\]](#)
- [Monitoring a task \[page 183\]](#)
- [Deleting failed data for a failed rule task \[page 186\]](#)
- [Common runtime parameters for Information Steward \[page 187\]](#)

8.5.10 Deleting failed data for a failed rule task

You can choose to have the software delete conflicting data from a failed data repository under certain circumstances.

When a rule task fails because of structural changes to the source data, or because someone changed the rule task set up between runs, delete failed data using the *Delete Failed Data and Run* option in the Central Management Console (CMC).

Note

You must have the appropriate permissions to perform these steps.

To delete failed data from a failed data repository using the *Delete Failed Data and Run* option, perform the following steps:

1. Log in to the CMC and open the *Information Steward* page.
2. Expand **Data Insight** > **Projects** at left.
A list of your projects appears.
3. Select the applicable project name.
A list of all tasks for that project appears at right.
4. Right-click the applicable rule task and select *Delete Failed Data and Run*.

An Error Reason message appears describing why the rule task failed, and it includes the failed data error code.

Note

If the *Delete Failed Data and Run* option isn't active in the list, the rule task failed for other reasons. Look at the log file to find out why the task failed. See the *User Guide* for a list of failed data error codes and descriptions.

5. Select *Delete Failed Data and Run* to continue with the deletion.

The software deletes the conflicting data from the failed data repository.

Note

To automatically clean up your failed data on a set schedule, use the Clean Up Failed Data utility. The schedule is turned off (set to 0) by default so that you don't lose any history when you upgrade an installation. However, you can enter a number that represents the frequency in days for the utility to run.

Task overview: [Data Insight tasks \[page 163\]](#)

Related Information

- [Scheduling a task in the CMC \[page 164\]](#)
- [Export tasks for external scheduler \[page 166\]](#)
- [Recurrence options \[page 177\]](#)
- [Notification for task completion \[page 178\]](#)
- [Configuring notification when scheduling a task \[page 179\]](#)
- [Rule threshold notification \[page 181\]](#)
- [Capturing skipped records from rule filter \[page 182\]](#)
- [Monitoring a task \[page 183\]](#)
- [Pausing and resuming a schedule \[page 185\]](#)
- [Common runtime parameters for Information Steward \[page 187\]](#)
- [Utilities overview \[page 276\]](#)

8.5.11 Common runtime parameters for Information Steward

Descriptions of the common runtime parameters that apply to all metadata integrators, profile tasks, and rule tasks.

When you schedule a metadata integrator, profile task, or rule task, you can change the default values of the runtime parameters. Schedule tasks in the Central Management Console (CMC). Change default parameters in the [Parameters](#) page for Information Steward.

The following table describes the common runtime parameters that are applicable to all metadata integrators, profile tasks, and rule tasks.

Runtime parameters for Information Steward

Runtime parameter	Description
Database Log Level	<p>The level of database information to provide in the log.</p> <p>The default log level is <i>Information</i>. You can change the default log level to a different level based on your needs. For example, change the default level to <i>Integrator trace</i> to provide detailed trace information about the integrator run.</p> <p>View the database log in the Information Steward repository after a task has completed. Or view the log while the Data Insight task or metadata integrator is running.</p>

Runtime parameter	Description
File Log Level	<p>The level of file information to provide in the log.</p> <p>The default logging level is <i>Configuration</i>. You can change the default log level to a different level based on your needs. For example, change the default level to <i>Integrator Trace</i> to include trace information to debug your task or integrator run.</p> <p>Download and view the file log from the file repository server after the Data Insight task or metadata integrator task completes.</p>
JVM Arguments	<p>Runtime parameters for Java processes.</p> <p>The Information Steward Job Server creates a Java process to perform the profile task, rule task, or metadata collection.</p> <p>For example, if the metadata source is large, increase the default memory size using a JVM argument.</p>
Additional Arguments	<p>Optional runtime parameters for the metadata integrator source or Data Insight task.</p>

Parent topic: [Data Insight tasks \[page 163\]](#)

Related Information

- [Scheduling a task in the CMC \[page 164\]](#)
- [Export tasks for external scheduler \[page 166\]](#)
- [Recurrence options \[page 177\]](#)
- [Notification for task completion \[page 178\]](#)
- [Configuring notification when scheduling a task \[page 179\]](#)
- [Rule threshold notification \[page 181\]](#)
- [Capturing skipped records from rule filter \[page 182\]](#)
- [Monitoring a task \[page 183\]](#)
- [Pausing and resuming a schedule \[page 185\]](#)
- [Deleting failed data for a failed rule task \[page 186\]](#)
- [Log levels \[page 340\]](#)
- [Runtime parameters for integrator sources \[page 229\]](#)
- [User collection parameters \[page 233\]](#)

9 Metadata Management Administration

9.1 Administration overview for Metadata Management

The Metadata Management module of SAP Information Steward collects metadata about objects from different source systems and stores the metadata in a repository.

Source systems include the following types:

- Business Intelligence: SAP BusinessObjects Enterprise and SAP Business Warehouse
- Data Modeling
- Data Integration: SAP Data Services and SAP Data Federator
- Relational database systems

When you access Information Steward as a user that belongs to the Metadata Management Administrator group, you can perform the following tasks in the Central Management Console (CMC):

- Configure integrator sources
- Run integrators to collect metadata
- View the status of metadata integrator runs
- Organize metadata integrator sources into groups for relationship analysis
- Manage user security for integrator sources, source groups, and Metapedia
- Enable BI launch pad users to view of documents
- View Metapedia terms and definitions in data lineage documents
- Compute and store end-to-end impact and lineage information for reporting
- Manage the Metadata Management search indexes

9.2 Configure metadata integrator sources

SAP Information Steward metadata integrators collect metadata from repository sources and populate the local repository with the collected metadata.

Information Steward supports the following metadata integrators:

- SAP BusinessObjects Enterprise
- SAP Business Warehouse
- SAP HANA
- SAP ERP Central Component
- Apache Hadoop
- Common Warehouse Model (CWM)
- SAP Data Federator

- SAP Data Services
- Relational Database systems
- MITI — Meta Integration Model Bridge (MIMB) applications
- SAP PowerDesigner
- Excel
- Amazon Redshift

During installation, the installer has all integrators selected for installation by default. If you uncheck any integrator, you won't be able to configure the integrator for use in Information Steward.

[Configuring SAP BusinessObjects Enterprise integrator sources \[page 191\]](#)

The SAP BusinessObjects Enterprise integrator collects metadata for objects from OLAP workspace, application users, dashboards, and more.

[Configuring SAP Business Warehouse integrator sources \[page 194\]](#)

The SAP Business Warehouse (BW) metadata integrator collects attributes and metadata for objects such as InfoObjects, InfoCubes, InfoProviders, and Crystal Reports.

[Configuring SAP Data Services metadata integrator sources \[page 195\]](#)

The SAP Data Services metadata integrator collects metadata for objects such as tables, columns, file formats, and data flows.

[Configuring SAP ERP Central Component integrator sources \[page 197\]](#)

The SAP ERP Central Component (ECC) integrator enables you to collect information from the SAP System about data models for objects such as tables, structures, data elements, and data domains.

[Configuring Apache Hadoop integrator sources \[page 199\]](#)

The Apache Hadoop integrator supports the collection of metadata for standard objects from Hive and HDFS.

[Configuring SAP HANA integrator sources \[page 204\]](#)

The SAP HANA metadata integrator collects metadata from Attribute, Analytical, and Calculation views.

[Configuring SAP PowerDesigner integrator sources \[page 206\]](#)

The SAP PowerDesigner metadata integrator collects metadata for objects in the PowerDesigner repository, and details about physical data models, including architectural and conceptual information that defines a physical data model.

[Configuring Common Warehouse Model integrator sources \[page 210\]](#)

The Common Warehouse Model (CWM) metadata integrator collects data models, data model instances, files, and metadata from SAP PowerDesigner.

[Configuring SAP Data Federator integrator sources \[page 212\]](#)

The SAP Data Federator metadata integrator collects metadata about projects, datasources, target schemas, and more.

[Configuring sources for Meta Integration Metadata Bridge Metadata Integrator \[page 212\]](#)

[Configuring relational database system integrator sources \[page 213\]](#)

The relational database systems metadata integrator collects metadata for server instances, databases, schemas, tables, views, and procedures when they exist in the database.

[Configuring Excel Content integrator sources \[page 219\]](#)

The Excel Content metadata integrator collects metadata from a relational database source or a Business Intelligence source.

[Configuring Amazon Redshift integrator sources \[page 220\]](#)

The Amazon Redshift metadata integrator supports the collection of metadata from an Amazon Redshift database, such as server instances, databases, schemas, and tables.

9.2.1 Configuring SAP BusinessObjects Enterprise integrator sources

The SAP BusinessObjects Enterprise integrator collects metadata for objects from OLAP workspace, application users, dashboards, and more.

To configure the SAP BusinessObjects Enterprise integrator, you must belong to the Metadata Management Administrator group or have the rights to add objects to the integrator source folders.

1. Log into the Central Management Console (CMC) and select Information Steward from the CMC Home page.
2. Select **Manage > New > Integrator Source**.
3. Select *SAP BusinessObjects Enterprise* in the *Integrator Type* drop-down list.
4. Enter the information described in the following table in the *New Integrator Source* page.

Option	Description
<i>Name</i>	Name for this metadata integrator source. The maximum length is 128 characters.
<i>Description</i>	Optional. Description for this metadata integrator source.
<i>CMS Server Name</i>	Host name of the Central Management Server (CMS). The host name is required.
Note The version of SAP BusinessObjects Enterprise installed on the SAP Information Steward host server must match the version of SAP BusinessObjects Enterprise that the CMS manages.	
<i>User Name</i>	Your CMS server user name. The default value is Administrator.
<i>Password</i>	Password associated with the specified user name.

Option	Description
<i>Authentication Method</i>	<p>Specify the process that the CMS uses to verify the identity of a user who attempts to access the system. See the <i>Business Intelligence Platform Administrator's Guide</i> for available processes.</p> <p>The default value is <code>Enterprise</code>.</p> <p>If you specify <code>Windows AD</code>, you must set JVM parameters in the the following areas:</p> <ul style="list-style-type: none"> Enterprise Information Management Services server Runtime parameters for the metadata integrator <p>For more information about using Windows AD, see Configuring sources with Windows AD authentication [page 193].</p>
<i>BI Launch pad Integration User</i>	User for BI Launchpad integration with Metada Management.
<i>Password</i>	Password for BI Launchpad integration user.
<i>InfoView Integration User</i>	<p>Name of the BI Launchpad (formerly known as InfoView) user. Calls the Information Steward lineage diagrams when you select View Lineage for selected documents in the Documents List page of BI Launchpad.</p> <p>The default value is <code>Administrator</code>.</p>
<i>Password</i>	Password associated with the specified BI Launchpad user.

- Click [Test connection](#).

Verifies that Information Steward can connect to the source.

- Click [Save](#).

The newly configured source appears in the list of [Integrator Sources](#) on the right of the Metadata Management page.

[Checkpointing \[page 193\]](#)

Checkpointing is an automatic mechanism that sets “safe start” points during job processing from which to restart jobs if there are unexpected failures.

[Configuring sources with Windows AD authentication \[page 193\]](#)

If you select to use Windows Active Directory (AD) authentication for the SAP BusinessObjects Enterprise metadata integrator, you must configure Java Virtual Machine (JVM) parameters for the EIM Adaptive Processing Server.

9.2.1.1 Checkpointing

Checkpointing is an automatic mechanism that sets “safe start” points during job processing from which to restart jobs if there are unexpected failures.

To collect large quantities of objects, SAP Information Steward is able to run the SAP BusinessObjects Enterprise metadata integrator for extended periods of time. However, sometimes an unexpected failure occurs, such as a power outage or an accidental shutdown. The failure stops the job from completing.

In addition to setting “safe start” points, the checkpointing feature records warning, error, and failure incidents for later analysis. If unexpected failures occur during object collection, analyze the log file to understand what caused the failure. When you rerun the job, the software doesn't have to start processing from the beginning. It finds the nearest “safe start” point to continue the job, and it completes object collection in a shorter amount of time.

9.2.1.2 Configuring sources with Windows AD authentication

If you select to use Windows Active Directory (AD) authentication for the SAP BusinessObjects Enterprise metadata integrator, you must configure Java Virtual Machine (JVM) parameters for the EIM Adaptive Processing Server.

To perform the following process, you must be a member of the administrator group, or have permission to configure servers.

To set the JVM parameters in the EIM Adaptive Processing Server, perform the following steps:

1. Log into the Central Management Console (CMC) and select [Servers](#) from the CMC Home page.
2. Expand the [Service Categories](#) node at left and select [Enterprise Information Management Services](#).
3. Right-click the EIM Adaptive Processing Server in the right pane and select [Stop Server](#).
4. Double-click the EIM Adaptive Processing Server.

The properties page opens.

5. Enter the following JVM parameters in the [Command Line Parameters](#) text box and set them to your file locations.

- `java.security.auth.login.config`
- `java.security.krb5.conf`

For example, enter the following text:

```
-Djava.security.auth.login.config=c:/winnt/bsclogin.conf  
-Djava.security.krb5.conf=c:/winnt/krb5.ini
```

6. Click [Save & Close](#).
7. Right-click the EIM Adaptive Processing Server and click [Start Server](#).

When you run the SAP BusinessObjects Enterprise metadata integrator source using Windows AD authentication, you must specify the required runtime parameters.

Related Information

[Running SAP BusinessObjects Enterprise Metadata Integrator with Windows AD authentication \[page 229\]](#)

9.2.2 Configuring SAP Business Warehouse integrator sources

The SAP Business Warehouse (BW) metadata integrator collects attributes and metadata for objects such as InfoObjects, InfoCubes, InfoProviders, and Crystal Reports.

To configure an SAP BW integrator source, you must have the create or add permission on the integrator source.

To configure sources for the SAP BW metadata integrator, perform the following steps:

1. Log in to the Central Management Console and open the Information Steward page.
2. Select **Manage > New > Integrator Source**.

The *New Integrator Source* page opens.

3. Select *SAP Business Warehouse* in the *Integrator Type* drop-down list.
4. Complete the options as described in the following table:

Option	Description
<i>Name</i>	Specifies a name for this integrator source. The maximum length is 128 characters.
<i>Description</i>	Optional. Specifies a description for the source.
<i>Connection Type</i>	Specifies a connection type for this source: <ul style="list-style-type: none">• <i>Custom Application Server</i>• <i>Group/Server selection</i>
<i>Application Server</i>	Specifies the SAP Application Server host name. Applicable only when you select <i>Custom Application Server</i> for <i>Connection Type</i> .
<i>Message Server</i>	Specifies the SAP BW Message Server host name. Applicable when you select <i>Group/Server Selection</i> for <i>Connection Type</i> .
<i>Group/Server</i>	Specifies the SAP group name. Applicable when you select <i>Group/Server Selection</i> for <i>Connection Type</i> .
<i>Client</i>	Specifies the identification number for the SAP BW client.

Option	Description
<i>System Number</i>	Specifies the number for the SAP BW system.
<i>SAProuter String</i>	Optional. Specifies the SAProuter connection string that connects the metadata integrator to the SAP BW system. The string contains the host name, the service port, and the password, if one was given. For more information, see Route String Entry for SAP-router in the SAP NetWeaver documentation.
<i>SAP User</i>	Specifies the user name for connecting to the SAP BW system.
<i>SAP Password</i>	Specifies the password for the specified <i>SAP User</i> name.
<i>Language</i>	Specifies the 2-character ISO code for the language to use for the descriptions of the SAP BW objects.

5. Select *Test connection*.
6. Select *Save* after the connection test is successful.

The SAP BW source appears in the list of integrator sources.

9.2.3 Configuring SAP Data Services metadata integrator sources

The SAP Data Services metadata integrator collects metadata for objects such as tables, columns, file formats, and data flows.

To perform the following process, you must have the create or add rights to the integrator source.

To configure SAP Data Services metadata integrator sources, perform the following steps:

1. Log into the Central Management Console (CMC) and open the Information Steward page.
2. Select **Manage > New > Integrator Source**.

The *New Integrator Source* page opens.

3. Select *Data Services* from the *Integrator Type* drop-down list.
4. Complete the options as described in the following table:

Option	Description
<i>Name</i>	Specifies a name for this integrator source. The maximum length is 128 characters.
<i>Description</i>	Optional. Specifies a description for the source.

Option	Description
<i>Database Type</i>	<p>Specifies the database type of the Data Services local repository:</p> <ul style="list-style-type: none"> • DB2 • Microsoft SQL Server • MySQL • Oracle • SAP Sybase • SAP HANA
<i>Computer Name</i>	Specifies the name of the computer where the Data Services repository resides.
<i>Database Port Number</i>	Specifies the port number for the source database.
<i>Datasource, Database Name, or Service Name</i>	<p>Specifies the name of the database, data source, or SID/service for the Data Services repository.</p> <ul style="list-style-type: none"> • DB2: Specify the data source name • Microsoft SQL Server: Specify the database name • MySQL: Specify the database name • SAP Sybase: Specify the database name • SAP HANA: Specify the database name • Oracle: Specify the SID/Service name
<i>Windows Authentication</i>	<p>Specifies whether to use Windows authentication to log in to Microsoft SQL Server.</p> <ul style="list-style-type: none"> • Yes: Microsoft SQL Server uses Windows authentication. • No: Microsoft SQL Server uses Microsoft SQL Server authentication. No is the default setting.
<i>Database User</i>	<p>Specifies the user name for the Data Services repository.</p> <p>Not applicable when the Data Services repository is Microsoft SQL server and you set <i>Windows Authentication</i> to Yes</p>
<i>Database Password</i>	<p>Specifies the password associated with the specified database user for the Data Services repository.</p> <p>Not applicable when the Data Services repository is Microsoft SQL server and you set <i>Windows Authentication</i> to Yes</p>
<i>HTTP Scheme</i>	<p>Specifies the protocol to connect to the SAP Data Services Management Console web application to access Auto Documentation reports.</p> <p>The default setting is HTTP.</p>
The following configuration options are only available for database type Microsoft SQL Server.	
<i>Use SSL Authentication</i>	Enables or disables SSL authentication.

Option	Description
<i>Verify Server Certificate</i>	Specifies whether the client certificate should be used to verify the host name.
<i>Host Name in Certificate</i>	Name used to verify the server's identity.
<i>SSL TrustStore</i>	Specifies the absolute location of the Java truststore file.
<i>SSL TrustStore Type</i>	Specifies the Java TrustStore file format.
<i>SSL TrustStore Password</i>	Specifies the password to access the private key from the truststore file.
The following configuration options are only available for database type SAP HANA.	
<i>Use SSL Authentication</i>	Enables or disables SSL authentication.
<i>Verify Server Certificate</i>	Specifies whether the client certificate should be used to verify the host name.
<i>Host Name in Certificate</i>	Name used to verify the server's identity.
<i>SSL KeyStore</i>	Absolute location of the Java KeyStore file.
<i>Java KeyStore file format</i>	Specifies the Java keystore file format.
<i>SSL KeyStore Password</i>	Password used to access the private key from the KeyStore file.
<i>SSL TrustStore</i>	Specifies the absolute location of the SSL truststore file.
<i>SSL TrustStore Type</i>	Specifies the SSL TrustStore file format.
<i>SSL TrustStore Password</i>	Specifies the password to access the private key from the truststore file.

5. Select *Test Connection*.
6. Click *Save* after the connection test is successful.
The new Data Services source appears in the integrator sources list.

9.2.4 Configuring SAP ERP Central Component integrator sources

The SAP ERP Central Component (ECC) integrator enables you to collect information from the SAP System about data models for objects such as tables, structures, data elements, and data domains.

To perform the following process, you must have the create or add rights to the integrator source.

To configure SAP ECC metadata integrator sources, perform the following steps:

1. Log into the Central Management Console (CMC) and open the Information Steward page.
2. Select **Manage > New > Integrator Source**.

The *New Integrator Source* page opens.

3. Select *SAP ERP Central Component* from the *Integrator Type* drop-down list.
4. Complete the options as described in the following table:

Option	Description
<i>Name</i>	Specifies a name for this integrator source. The maximum length is 128 characters.
<i>Description</i>	Optional. Specifies a description for the source.
<i>Connection Type</i>	Specifies the type of connection to use for this source: <ul style="list-style-type: none"> <i>Custom Application Server</i> <div data-bbox="874 566 1398 719" data-label="Text"> <p>Note If you select <i>Custom Application Server</i>, enter a value for <i>Application Server</i>.</p> </div> <i>Group/Server Selection</i> <div data-bbox="874 770 1398 922" data-label="Text"> <p>Note If you select <i>Group/Server Selection</i>, enter a value for <i>Message Server</i> and <i>Group/Server</i>.</p> </div>
<i>Application Server</i>	Specifies the host name of the application server. Applicable when you select <i>Custom Application Server</i> for <i>Connection Type</i> .
<i>Message Server</i>	Specifies the host name for the SAP Message Server. Applicable when you select <i>Group/Server Selection</i> for <i>Connection Type</i> .
<i>Group/Server</i>	Specifies the name for the SAP group. Applicable when you select <i>Group/Server Selection</i> for <i>Connection Type</i> .
<i>Client</i>	Specifies the SAP System client ID number.
<i>System Number</i>	Specifies the number for the SAP system.
<i>SAProuter String</i>	Specifies the string that contains the information to set up a connection between the SAP ECC metadata integrator and the SAP system. The string contains the host name, the service port, and, if applicable, the password. Specifying an SAProuter string is optional.
<i>SAP User</i>	Specifies the name of the user authorized to connect to the SAP system.
<i>SAP Password</i>	Specifies the password associated with the specified SAP user.
<i>Language</i>	Specifies the 2-character ISO code for the language in which the SAP system objects are described.

5. Select *Test Connection*.

6. Click [Save](#) after the connection test is successful.
The new SAP ECC source appears in the integrator sources list.

9.2.5 Configuring Apache Hadoop integrator sources

The Apache Hadoop integrator supports the collection of metadata for standard objects from Hive and HDFS.

Before performing the following task, export the Knox security certificate from the Linux computer that is hosting Knox. The keystore is located at `'/var/lib/knox/data/security/kestores':keytool -export -alias gateway-identity -file gateway-identity.crt -keystore gateway.jks`.

After you export the Knox security certificate, perform the following steps:

1. Copy the certificate from the Linux computer to the computer that hosts Information Steward.
 1. For integrator runtime: `SAP BusinessObjects\SAP BusinessObjects Enterprise XI 4.0\win64_x64\sapjvm\jre\lib\security`
 2. For the test connection: `\SAP BusinessObjects\Data Services\ssl\trusted_certs\jssecacerts`
2. Import the certificate using the following command: `keytool -import -trustcacerts -alias gateway-identity -file gateway-identity.crt -keystore cacerts`

To test the Apache Hadoop integrator, we used the dependencies listed in the following table. Ensure that you use the relevant or corresponding files and versions that are supported for your specific environment.

ⓘ Note

- If you run the Hadoop metadata integrator only for HDFS, you don't need 'hive*' jar file dependencies from this list.
- If you run the Hadoop metadata integrator only for Hive, or for Hive and HDFS, all of the dependencies are required.

Dependency files

- slf4j-api-1.6.1.jar
- slf4j-log4j12-1.6.1.jar
- antlr-runtime-3.5.2.jar
- aspectjrt-1.8.2.jar
- aspectjweaver.jar
- commons.io_2.0.1.jar
- commons-cli-1.3.1.jar
- commons-codec-1.10.jar
- commons-collections-3.2.1.jar
- commons-configuration-1.9.jar
- commons-lang3-3.5.jar
- freemarker-2.3.9.jar
- gateway-shell-0.5.1.jar
- groovy-json-2.4.3.jar
- hadoop-auth-2.2.0.jar
- hadoop-common-2.4.1.jar
- hadoop-hdfs-2.2.0.jar
- hive-exec-0.13.0.jar
- hive-hcatalog-core-0.13.0.jar
- hive-jdbc-1.1.1.jar
- hive-metastore-0.13.0.jar
- hadoop-mapreduce-client-core-2.4.1.jar
- hive-service-0.14.0.jar
- hive-webhcat-java-client-0.13.1.jar
- httpclient-4.5.2.jar
- httpcore-4.4.4.jar
- ion-java-1.0.2.jar
- jackson-annotations-2.6.0.jar
- jackson-core-2.6.6.jar
- jackson-databind-2.6.6.jar
- jackson-dataformat-cbor-2.6.6.jar
- javax.jdo-jdo2-api-2.3-eb.jar
- javax.mail-api-1.4.6.jar
- javax.ws.rs.jar
- jdo.jar
- jersey-bundle-1.6.jar
- jetty-util-6.1.25.jar
- jamespath-java-1.11.118.jar
- joda-time-2.8.1.jar
- libfb303-0.9.0.jar
- log4j-1.2.16.jar

To configure sources for the Apache Hadoop metadata integrator, perform the following steps:

1. Log into the Central Management Console (CMC) and open the Information Steward page.
2. Select **Manage > New > Integrator Source**.

The *New Integrator Source* page opens.

3. Select *Apache Hadoop* from the *Integrator Type* drop-down list.
4. Complete the options as described in the following table:

Note

Some subsequent options appear only when you select specific values for previous options. For example, the options you see after the *Entry Point* option depend on the selection you make for the *Entry Point* option.

Option	Description
<i>Name</i>	Specifies a name for this integrator source. The maximum length is 128 characters.
<i>Description</i>	Optional. Specifies a description for the source.

Option	Description
<i>Collect Hive</i>	Specifies whether to collect metadata for Hive databases, schemas, tables, or columns. <ul style="list-style-type: none"> Yes: Collect metadata No: Doesn't collect metadata
<i>Database/Schemas</i>	Specifies the specific database and schemas to collect: <ul style="list-style-type: none"> Enter a list of databases and schemas separated by a comma: Collects the listed databases and schemas. Leave blank: Collects all databases and schemas.
<i>Collect HDFS</i>	Specifies whether to collect metadata for HDFS directories, files, and file fields for CVS files. <ul style="list-style-type: none"> Yes: Collects metadata No: Doesn't collect metadata
<i>Paths</i>	Specifies the paths for collection, separated by commas.
<i>Libraries directory</i>	Specifies the name of the directory on the Information Steward server that contains all of the required Hadoop libraries and dependencies.
<i>Entry point</i>	Specifies the entry point for the Hadoop cluster. <ul style="list-style-type: none"> Knox: Cluster uses the Knox gateway to access the REST API. Hosts and Ports: <ul style="list-style-type: none"> Select if the cluster doesn't use the Knox gateway to access the REST API. Select if the cluster uses the SAP Big Data Services (Altiscale) gateway to access the REST API.

Note

Also set *HttpFS Used* to **Yes** and complete the *HttpFS User* option with your Big Data Services gateway user name.

If you set *Entry point* to *Knox*, complete the following options:

<i>Username</i>	Specifies the authenticated user name.
<i>Password</i>	Specifies the password associated with the authenticated user name.
<i>Host</i>	Specifies the name of the Knox host.
<i>Port</i>	Specifies the port number for the Knox gateway.
<i>Gateway</i>	Specifies the file path for the Knox gateway.

If you set *Entry point* to *Hosts and Ports*, complete the following options:

Option	Description
<i>NameNode host</i>	<p>Specifies the fully qualified domain name of the Hadoop Name Node.</p> <p>If you don't use Kerberos, the software uses the service user to get HDFS metadata.</p> <p>If you set <i>Collect HDFS</i> to <i>No</i> and <i>Collect Hive</i> to <i>Yes</i>, provide the Name Node host name to collect the Hive table to the HDFS file lineage.</p> <p>If you leave <i>NameNode host</i> blank, the software doesn't collect the lineage.</p>
<i>HDFS port</i>	<p>Specifies the Hadoop HDFS Name Node port number.</p> <p>If you set <i>Collect HDFS</i> to <i>No</i> and <i>Collect Hive</i> to <i>Yes</i>, provide the HDFS port number to collect the Hive table to the HDFS file lineage.</p> <p>If you leave <i>HDFS port</i> blank, the software doesn't collect the lineage.</p>
<i>HttpFS used</i>	<p>Specifies whether the HttpFS gateway is configured to access HDFS.</p> <ul style="list-style-type: none"> • <i>Yes</i>: Specifies that you use the SAP Big Data Services as the gateway. • <i>No</i>: Specifies the HttpFS gateway isn't configured to access HDFS.
<i>HttpFS user</i>	<p>Specifies the name of the user who is accessing HDFS through HttpFS. Leave blank to indicate that you aren't using HttpFS.</p>
<i>Metastore Node host name</i>	<p>Specifies the fully qualified domain name of the Hadoop Data Node that is running the Hive Metastore service.</p>
<i>Hive Metastore port</i>	<p>Specifies the port number on which the Hive Metastore service is running.</p>
<i>Hive Server2 port</i>	<p>Specifies the port number on which Hive Server2 is running.</p>
<i>Hive user</i>	<p>Specifies the user name for Hive.</p> <p>If you set <i>Collect Hive</i> to <i>No</i>, leave <i>Hive user</i> blank.</p> <p>The software ignores the setting in <i>Hive user</i> when you use Kerberos authentication and both of the following statements are true:</p> <ul style="list-style-type: none"> • Hive uses the Kerberos user principal to authenticate. • Hive uses the Hive service user principal to collect metadata.
<i>Hive password</i>	<p>Specifies the password for the Hive user.</p>

Option	Description
<i>Truststore location & name</i>	Specifies the full path and name of the truststore. Applicable only if the Hive server uses SSL.
<i>Truststore password</i>	Specifies the password for accessing the truststore. <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 5px; margin-top: 10px;"> <p>Note The Truststore must contain the client certificate.</p> </div>
<i>Kerberos authentication</i>	Specifies whether to use Kerberos for authentication. <ul style="list-style-type: none"> Yes: Specifies to use Kerberos for authentication. No: Specifies that the software collects HDFS meta-data using the service user and the software collects Hive metadata using the user provided.
If you set <i>Kerberos authentication</i> to <i>Yes</i>, complete the following options:	
<i>Kerberos user principle</i>	Specifies the user principal that you use for Kerberos.
<i>Kerberos authentication method</i>	Specifies the authentication method to use: <ul style="list-style-type: none"> Keytab: Uses a keytab to perform the Kinit. Password: Uses a password to perform Kinit.
If you set <i>Kerberos authentication method</i> to <i>Keytab</i>, complete the following options:	
<i>Keytab location</i>	Specifies the path for the keytab file on the Information Steward client. If the keytab isn't present, create it for a user principal that is added to the list of authenticated principals using the kadmin.local tool. Applicable when you select <i>Keytab</i> for the <i>Kerberos authentication method</i> option.
<i>Kerberos configuration file path</i>	Specifies the path for the Kerberos configuration file, krb5.ini/krb5.conf, on the Information Steward client. The default location is /etc/krb5.conf. Applicable when you select either <i>Keytab</i> or <i>Password</i> for the <i>Kerberos authentication method</i> option.
If you set <i>Kerberos authentication method</i> to <i>Password</i>, complete the following options:	
<i>Kerberos user password</i>	Specifies the user password for the user principal you entered for the <i>Kerberos user principle</i> option. Applicable when you select <i>Password</i> for the <i>Kerberos authentication method</i> option.
<i>Kerberos configuration file path</i>	Specifies the path for the Kerberos configuration file, krb5.ini/krb5.conf, on the Information Steward client. The default location is /etc/krb5.conf. Applicable when you select either <i>Keytab</i> or <i>Password</i> for the <i>Kerberos authentication method</i> option.

5. Select *Test Connection*.

- Click [Save](#) after the connection test is successful.
The new Apache Hadoop source appears in the integrator sources list.

9.2.6 Configuring SAP HANA integrator sources

The SAP HANA metadata integrator collects metadata from Attribute, Analytical, and Calculation views.

Before you perform the following task, ensure that you have permission to create configuration sources, and that you have the following permissions:

- GRANT SELECT ON SCHEMA _SYS_BIC TO `<ROLE_NAME>`;
- GRANT SELECT ON SCHEMA _SYS_BI TO `<ROLE_NAME>`;
- GRANT CATALOG READ TO `<ROLE_NAME>`;
- GRANT REPO.READ ON ".REPO_PACKAGE_ROOT" TO `<ROLE_NAME>`;

To configure a new SAP HANA database source, perform the following steps:

- Log into the Central Management Console (CMC) and open the Information Steward page.
- Select **Manage > New > Integrator Source**.

The *New Integrator Source* page opens.

- Select *SAP HANA* from the *Integrator Type* drop-down list.
- Complete the options as described in the following table.

Options are listed in alphabetical order.

Option	Description
<i>Computer Name</i>	Specifies the host name of the computer on which the SAP HANA database server is running.
<i>Database Password</i>	Specifies the password for the database user authorized to connect to the SAP HANA database.
<i>Database Port Number</i>	Specifies the port number of the computer on which the SAP HANA database server is running.
<i>Database User</i>	Specifies the name of the user authorized to connect to the SAP HANA database
<i>Description</i>	Optional. Specifies a description for the source.

Option	Description
<i>HANA HDI Containers</i>	<p>Restricts collection of SAP HANA Data Integrator (HDI) calculation views to specific containers.</p> <p>To list multiple containers, use a comma delimiter.</p> <p>Leave blank: Collects Calculation views from all SAP HANA XSA schemas regardless of user access control.</p> <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 10px; margin-top: 10px;"> <p>Note</p> <p>If a container name isn't valid or doesn't exist, SAP Information Steward stops the execution and writes an error message to the log.</p> </div>
<i>Host Name in Certificate</i>	Specifies the host name used to verify the server's identity.
<i>SSL KeyStore type</i>	Specifies the type of the keystore certificate file.
<i>Name</i>	<p>Specifies a name for this integrator source.</p> <p>The maximum length is 128 characters.</p>
<i>Packages</i>	<p>Optional. Specifies the name of the package or packages. To enter multiple package names, separate each with a comma.</p> <p>Entering the name of the package or packages limits the collection of SAP HANA Integrator source configuration to the named package or packages, and any child packages.</p>
<i>SSL KeyStore</i>	Specifies the absolute location of the Java keystore file.
<i>SSL KeyStore Password</i>	Specifies the password used to access the private key from the keystore file.
<i>SSL TrustStore</i>	Specifies the absolute location of the Java truststore file.
<i>SSL TrustStore Type</i>	Specifies the Java TrustStore file format.
<i>SSL TrustStore Password</i>	Specifies the password to access the private key from the truststore file.
<i>Table Schema</i>	<p>Optional. Specifies the name of the table schema.</p> <p>Entering a table schema name limits metadata collection to the named schema. Information Steward collects all tables and views that belong to the specified schema.</p> <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 10px; margin-top: 10px;"> <p>Note</p> <p>If a view in the specified schema is a relational view or Information Model view that references a different table schema, Information Steward also collects the metadata from the referenced table schema.</p> </div>
<i>Use SSL Authentication</i>	Enables or disables SSL authentication.

Option	Description
<i>Verify Server Certificate</i>	Specifies whether the client certificate should be used to verify the host name.

5. Select *Test Connection*.
6. Click *Save* after the connection test is successful.
The new SAP HANA source appears in the integrator sources list.


9.2.7 Configuring SAP PowerDesigner integrator sources

The SAP PowerDesigner metadata integrator collects metadata for objects in the PowerDesigner repository, and details about physical data models, including architectural and conceptual information that defines a physical data model.

Before you can perform the following task, ensure that you have permission to view Metapedia users and approvers and that you have permission to create integrator sources.

In addition, perform all of the preparation tasks listed in the topic [Prepare to configure SAP PowerDesigner integrator sources \[page 209\]](#).

To configure a source for the PowerDesigner metadata integrator:

1. Log in to the CMC and access the Information Steward page.
2. Click **Manage** > **New** > **Integrator Source**. 
The *New Integrator Source* page opens.
3. Select *SAP Sybase PowerDesigner* from the *Integrator Type* drop-down list.
4. Complete the options described in the following table.

Option	Description
<i>Name</i>	Specifies a name for this integrator source. The maximum length is 128 characters.
<i>Description</i>	Optional. Specifies a description for the source.
<i>Repository Connection</i>	Specifies the name of the definition for the PowerDesigner repository connection.
<i>User</i>	Specifies the user name for the PowerDesigner application.
<i>Password</i>	Specifies the password for the specified user of the PowerDesigner application.

Option	Description
<i>Collect Glossary</i>	<p>Specifies whether to obtain glossary entries from the PowerDesigner repository.</p> <ul style="list-style-type: none"> • <i>True</i>: Obtains glossary entries from the PowerDesigner repository through the PowerDesigner application. <div data-bbox="874 504 1396 739" style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 10px;"> <p>Note</p> <p>Applicable for SAP PowerDesigner version 16.1 and higher.</p> <p>The software sets all collected terms to the approved state in Metapedia.</p> </div> <ul style="list-style-type: none"> • <i>False</i>: Disables glossary collection. All other glossary collection settings are removed from the setup. <div data-bbox="874 824 1396 974" style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 10px;"> <p>Note</p> <p>When you select <i>False</i>, there are no other options to complete.</p> </div>
<p>The following options are applicable when you select <i>True</i> for <i>Collect Glossary</i>:</p>	
<i>Metapedia Author</i>	<p>Specifies the Metapedia author that the software assigns to imported glossary terms collected by the PowerDesigner integrator.</p> <p>Select the applicable author from the drop-down list.</p>
<i>Metapedia Approver</i>	<p>Specifies the Metapedia approver that the software assigns to imported glossary terms collected by the PowerDesigner integrator.</p> <p>Select the applicable approver from the drop-down list.</p>
<i>Metapedia Observer</i>	<p>Specifies the Metapedia observer that the software assigns to imported glossary terms collected by the PowerDesigner integrator.</p> <p>The observer receives an email notification when a term is approved or deleted. The observer must have Metapedia view rights or at least be a member of the Metapedia users group.</p>

Option	Description
<i>Force update from PowerDesigner</i>	<p data-bbox="813 315 1402 421">Specifies whether to collect all terms from the PowerDesigner repository and add to Metapedia, even when they already exist in Metapedia.</p> <ul data-bbox="813 421 1402 896" style="list-style-type: none"> <li data-bbox="813 421 1402 526">• <i>True</i>: Collects terms from the PowerDesigner repository and adds them to Metapedia, even when they already exist in Metapedia. <ul data-bbox="877 526 1402 896" style="list-style-type: none"> <li data-bbox="877 526 1402 638">• If the term already exists in Metapedia, the software overwrites the applicable components in Metapedia with the components collected from PowerDesigner. <li data-bbox="877 638 1402 772">• If the term component already exists in Metapedia but it doesn't exist in PowerDesigner, the software doesn't change the existing component in Metapedia. <li data-bbox="877 772 1402 896">• If the term in Metapedia is assigned to a different category than the term in PowerDesigner, the software doesn't change the category in Metapedia. <div data-bbox="874 913 1402 1205" style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p data-bbox="874 913 1402 958">Note</p> <p data-bbox="874 958 1402 1205">After a term is added to Metapedia through running the PowerDesigner integrator, you can't edit the Metapedia term using PowerDesigner. However, you can update the term in the PowerDesigner application and rerun the PowerDesigner integrator with the <i>Force update from PowerDesigner</i> set to <i>True</i>.</p> </div> <ul data-bbox="813 1205 1402 1332" style="list-style-type: none"> <li data-bbox="813 1205 1402 1332">• <i>False</i>: Collects terms that are unique to PowerDesigner and adds them to Metapedia. Doesn't collect terms from the PowerDesigner repository that already exist in Metapedia. <div data-bbox="826 1350 1402 1541" style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p data-bbox="826 1350 1402 1395">Note</p> <p data-bbox="826 1395 1402 1541">To see where a term in Metapedia came from, look at the comment in the upper right corner of the <i>Metapedia Term Properties</i> page.</p> </div>

5. Select *Test Connection*.

→ Remember

If you haven't followed the steps to change your DCOM user information, and completed other important prerequisites, your connection will be unsuccessful and you'll receive an error.

6. Click *Save* after the connection test is successful.
The new SAP PowerDesigner source appears in the integrator sources list.

9.2.7.1 Prepare to configure SAP PowerDesigner integrator sources

Before you configure sources for the SAP PowerDesigner metadata integrator, perform prerequisite steps to ensure the correct location for the application, define a repository, and change the DCOM identity of the application computer.

Perform the following prerequisites before following the steps to configure sources for the PowerDesigner metadata integrator:

- Ensure that you've installed the integrator, and that you have the correct permissions and rights.
- Make sure that your SAP PowerDesigner application is installed on the same machine as where you plan to run the PowerDesigner metadata integrator.
- Define a repository connection in the PowerDesigner application. Make note of the following information, which is required when you configure sources for the PowerDesigner integrator:
 - PowerDesigner application user name and password
 - Name of the PowerDesigner repository

Note

For information about defining a PowerDesigner repository, see your SAP PowerDesigner documentation.

- Change the identity information for the PowerDesigner Microsoft Distributed Component Object Model (DCOM) on the machine that contains both the PowerDesigner and SAP Information Steward application installations.

Related Information

[Changing DCOM user information for PowerDesigner \[page 209\]](#)

[Configuring SAP PowerDesigner integrator sources \[page 206\]](#)

9.2.7.2 Changing DCOM user information for PowerDesigner

The Microsoft Distributed Component Object Model (DCOM) enables a computer to run applications over the network on a different computer, as if the application is running locally.

Before you can run the SAP PowerDesigner metadata integrator in SAP Information Steward, change the user for the PowerDesigner DCOM.

Note

The PowerDesigner application isn't available when you make this change.

Before you perform the following task, make sure that both Information Steward and PowerDesigner applications are on the same machine.

To change the user information on the PowerDesigner DCOM, perform the following steps:

1. In Windows, type “Administrative Tools” in the Start search box and select **Windows Administrative Tools** from the results.
2. Open **Component Services**.
3. Expand the Component Services node and select **Computers > My Computer > DCOM Config**.
4. Right-click *PowerDesigner* in the *DCOM Config* list and select *Properties*.
The *PowerDesigner Properties* window opens.
5. Open the *Identity* tab.
6. Select *This user* and enter the applicable user name in the *User* text box.
7. Enter the applicable password in the *Password* text box and the *Confirm password* text box.
8. Select *OK*.

Before you can use the SAP PowerDesigner application installed on the same machine as the SAP Information Steward application, reset the user and password in the *Identity* tab to the previous values.

9.2.8 Configuring Common Warehouse Model integrator sources

The Common Warehouse Model (CWM) metadata integrator collects data models, data model instances, files, and metadata from SAP PowerDesigner.

Before you perform the following task, ensure that you have the right to add objects to the integrator sources folder.

1. Log into the Central Management Console (CMC) and open the Information Steward page.
2. Select **Manage > New > Integrator Source**.

The *New Integrator Source* page opens.

3. Select *Common Warehouse Model* from the *Integrator Type* drop-down list.
4. Complete the options as described in the following table.

Option	Description
<i>Source Name</i>	Specifies a name for this integrator source. The maximum length is 128 characters.
<i>Description</i>	Optional. Specifies a description for the source.

Option	Description
<i>File Name</i>	<p>Specifies the location and file name of the file that contains the CWM content. For example: C:\data\cwm_export.xml</p> <p>Select the Browse button to find and select the file.</p> <p>Ensure that the file is available from the computer where the Metadata Management web browser is running.</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p>Note</p> <p>The software places a copy of the specified file in the Input File Repository Server on the SAP Business Intelligence Platform. Therefore, if the original file is subsequently updated, update the specified file before you run the CWM Integrator again. For instructions, see Obtaining updates from the original CWM content file [page 211].</p> </div>

5. Click [Save](#).
The newly configured source appears in the list of integrator sources list.

9.2.8.1 Obtaining updates from the original CWM content file

If the original Common Warehouse Model (CWM) content file is updated after you configure the CWM integrator, update the integrator source CWM content file before running the integrator again.

When you create the CWM metadata integrator source, you specify the name and location of the CWM content file. The software copies the CWM content file to the SAP Business Intelligence Platform Input File Repository. Therefore, if the original file is later updated, obtain the updates before you run the CWM integrator again.

Perform the following steps in the SAP Information Steward page of the Central Management Console:

1. Update the configuration to re copy the CWM file:
 - a. Select the CWM integrator source in the Integrator list.
 - b. Select [Action](#) [Properties](#).

In the Properties dialog box, the file name displays in Comments under the [File Name](#) text box. The file name contains the preface “frs.”
 - c. Select [Browse](#) and browse to the original file location.
 - d. Select [Save](#).
2. Create a new schedule for the CWM integrator:

A new schedule overwrites the existing schedule that contains the copy of the previous file.

 - a. With the CWM integrator source name still selected in the Integrator list, select [Action](#) [Schedules](#).
 - b. Select the applicable values for the [Recurrence](#) and [Parameter](#) options.

- c. Select *Schedule*.

9.2.9 Configuring SAP Data Federator integrator sources

The SAP Data Federator metadata integrator collects metadata about projects, datasources, target schemas, and more.

To perform the following process, you must have the create or add rights to the integrator source.

To configure SAP Data Federator integrator sources, perform the following steps:

1. Log into the Central Management Console (CMC) and open the Information Steward page.
2. Select ► *Manage* ► *New* ► *Integrator Source* ►.

The *New Integrator Source* page opens.

3. Select *Data Federator* from the *Integrator Type* drop-down list.
4. Complete the options as described in the following table:

Option	Description
<i>Name</i>	Specifies a name for this integrator source. The maximum length is 128 characters.
<i>Description</i>	Optional. Specifies a description for the source.
<i>DF Designer Server Address</i>	Specifies the IP address of the computer where the Data Federator Designer resides.
<i>DF Designer Server Port</i>	Specifies the port number for the Data Federator Designer. The default port is 3081.
<i>User Name</i>	Specifies the name of the user that has permission to access Data Federator Designer.
<i>Password</i>	Specifies the password for the specified user in <i>User Name</i> .

5. Select *Test Connection*.
6. Click *Save* after the connection test is successful.
The new SAP Data Federator source appears in the integrator sources list.

9.2.10 Configuring sources for Meta Integration Metadata Bridge Metadata Integrator

This section describes how to configure the Metadata Integrator for Meta Integration Metadata Bridge (MIMB). For a description of the objects collected by the MIMB Integrator, see the MIMB documentation at <http://www.metaintegration.net/Products/MIMB/Documentation/>.

Note

Ensure that you selected the Meta Integration Metadata Bridge (MIMB) Metadata Integrator when you installed SAP Information Steward.

To configure the MIMB Integrator, you must have the *Create* or *Add* right on the integrator source.

1. Log in to the Central Management Console (CMC) and access the Information Steward area.
2. Click the down arrow next to *Manage* in the top menu tool bar and select **► New ► Integrator Source ▾**.
3. In the *Integrator Type* drop-down list, select *Meta Integration Model Bridge*.
4. On the *New Integrator Source* page, enter values for *Name* and *Description*. The maximum length of an integrator source name is 128 characters.
5. In the *Bridge* drop-down list, select the type of integrator source from which you want to collect metadata and follow the instructions on the user interface to configure the connection information.
6. Click *Save*.
The newly configured source appears in the list of Integrator Sources on the right of the page.

9.2.11 Configuring relational database system integrator sources

The relational database systems metadata integrator collects metadata for server instances, databases, schemas, tables, views, and procedures when they exist in the database.

To perform the following process, you must have the create or add rights to the relational database source.

Configure and run the Metadata Integrator for the following types of databases:

- Universe Connections
- DB2
- Microsoft SQL Server
- MySQL
- Oracle
- JDBC
- Sybase ASE
- Sybase IQ

To configure the sources for a Relational Database System metadata integrator, perform the following steps:

1. Log into the Central Management Console (CMC) and open the Information Steward page.
2. Select **► Manage ► New ► Integrator Source ▾**.
The *New Integrator Source* page opens.
3. Complete the options as described in the following table.

Note

When you select the *Connection Type*, the remaining options change based on the connection type you choose.

Option	Description
<i>Name</i>	<p>Specifies a name for this integrator source.</p> <p>The maximum length is 128 characters.</p>
<i>Description</i>	Optional. Specifies a description for the source.
<i>Connection Type</i>	<p>Specifies the type of database.</p> <p>Select one of the following database types from the drop-down list:</p> <ul style="list-style-type: none"> • Universe Connection • DB2 • Microsoft SQL Server • MySQL • Oracle • JDBC (for databases such as Teradata) • SAP ASE (formerly Sybase ASE) • SAP IQ (formerly Sybase IQ)
	<div style="background-color: #f0f0f0; padding: 10px; border: 1px solid #ccc;"> <p>Note</p> <p>For Oracle and MySQL, you must install the applicable JDBC driver. For information, see Configuring sources for JDBC connections [page 216].</p> </div>
	<div style="background-color: #f0f0f0; padding: 10px; border: 1px solid #ccc;"> <p>Note</p> <p>For more information about the Universe Connection configuration, see Configuring sources for universe connections [page 218].</p> </div>
The remaining options that appear are based on the connection type you choose.	
<i>Connections</i>	<p>Specifies the name of the Central Management Server (CMS) connection.</p> <p>Select a secure connection from the drop-down list.</p> <p>Applicable when you select <i>Universe Connection</i> for <i>Connection Type</i>.</p>
<i>Computer Name</i>	Specifies the name of the host on which the database server is running.
<i>Database Port Number</i>	Specifies the port number on which the database server is running.
<i>Use SSL Authentication</i>	<p>Enables or disables SSL authentication.</p> <p>Only applicable when you select <i>Microsoft SQL Server</i> for <i>Connection Type</i>.</p>

Option	Description
<i>Verify Server Certificate</i>	<p>Specifies whether the client certificate should be used to verify the host name.</p> <p>Only applicable when you select <i>Microsoft SQL Server</i> for Connection Type.</p>
<i>Host Name in Certificate</i>	<p>Name used to verify the server's identity.</p> <p>Only applicable when you select <i>Microsoft SQL Server</i> for Connection Type.</p>
<i>SSL TrustStore</i>	<p>Absolute location of the SSL TrustStore file.</p> <p>Only applicable when you select <i>Microsoft SQL Server</i> for Connection Type.</p>
<i>SSL TrustStore Type</i>	<p>SSL TrustStore file format.</p> <p>Only applicable when you select <i>Microsoft SQL Server</i> for Connection Type.</p>
<i>SSL TrustStore Password</i>	<p>Password to access the private key from the TrustStore file.</p> <p>Only applicable when you select <i>Microsoft SQL Server</i> for Connection Type.</p>
<i>Database Name, or Service name</i>	<p>Specifies the name of the database for DB2, Microsoft SQL Server, or MySQL. Specifies the Service Identification (SID) for Oracle.</p>
<i>Windows Authentication</i>	<p>Specifies whether to use Windows Authentication to log into Microsoft SQL Server.</p> <ul style="list-style-type: none"> • Yes: Uses Windows Authentication to log into Microsoft SQL Server. • No: Uses Microsoft SQL Server authentication to log into Microsoft SQL Server. No is the default setting.
<i>Database User</i>	<p>Specifies the name of the user or owner of the database or data source.</p>
<i>Database Password</i>	<p>Specifies the password of the user for the database or data source.</p>

Option	Description
<i>Table Schema</i>	<p>Optional. Specifies the schema name, or multiple schema names, to import from this database source.</p> <p>Specify multiple schemas by separating each schema name with a comma.</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> <p>Note</p> <p>When the schema name includes a comma, use double quotation marks around the schema name.</p> </div> <p>If you don't specify a schema name or names for DB2 or Microsoft SQL Server, the software imports all available schemas.</p> <p>If you don't specify a schema name or names for Oracle, the software uses the user name to import the schema.</p>

4. Select *Test Connection*.
5. Click *Save* after the connection test is successful.
The new configured source appears in the integrator sources list.

[Configuring sources for JDBC connections \[page 216\]](#)

When you select a relational database source that uses a JDBC connection, such as Teradata, configure the JDBC driver.

[Configuring sources for universe connections \[page 218\]](#)

When you select a universe connection for the relational database metadata integrator, configure the source based on the type of driver required for the universe connection.

9.2.11.1 Configuring sources for JDBC connections

When you select a relational database source that uses a JDBC connection, such as Teradata, configure the JDBC driver.

To configure a JDBC driver for relational database sources, perform the following steps:

1. Obtain the JDBC driver from your database server web site or utilities CD.
2. Unzip the JDBC driver into a folder such as:

```
c:\temp\teradata
```

3. Log into the Central Management Console (CMC) and open the Information Steward page.
4. Select **Manage > New > Integrator Source**.
The *New Integrator Source* dialog box opens.
5. Select *Relational Database* from the *Integrator Type* drop-down list.
6. Specify the JDBC connection parameters described in the following table.

Option	Description
<i>Name</i>	Specifies a name for this integrator source. The maximum length is 128 characters.
<i>Description</i>	Optional. Specifies a description for the source.
<i>Connection Type</i>	Select <i>JDBC</i> from the drop-down list.
<i>Driver</i>	Specifies the name of the JDBC driver class. Find the driver class name when you obtain the JDBC driver.
<i>URL</i>	Specifies the URL for the JDBC connection to the database.
<i>Catalog</i>	Optional. Specifies the name of the catalog in the relational database.
<i>Database User</i>	Specifies the name of the user or owner who has permission to access the database.
<i>Database Password</i>	Specifies the password for the user or owner who has permission to access the database.
<i>Table Schema</i>	Optional. Specify the name of the schema, or the names of multiple schemas to import from this source database. Specify multiple schema names by including a comma between each name. <div data-bbox="826 1077 1398 1232" style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 5px;"> <p>Note</p> <p>When the schema name includes a comma, use double quotation marks around the schema name.</p> </div>
<i>Library files</i>	Specifies the jar files required for this database. Separate each jar file name with a semicolon. For example: <code>c:\temp\teradata\tdgssjava.jar;c:\temp\teradata\terajdbc4.jar;c:\temp\teradata\tdgssconfig.jar</code> <div data-bbox="826 1431 1398 1568" style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 5px;"> <p>Note</p> <p>In a distributed deployment, set <i>Library Files</i> to the classpath on the computer where the integrator runs.</p> </div>

7. Select *Test Connection*.
8. Click *Save* after the connection test is successful.
The new source appears in the integrator sources list.

Task overview: [Configuring relational database system integrator sources \[page 213\]](#)

Related Information

[Configuring sources for universe connections \[page 218\]](#)

9.2.11.2 Configuring sources for universe connections

When you select a universe connection for the relational database metadata integrator, configure the source based on the type of driver required for the universe connection.

Before you configure a source for a universe connection that uses an ODBC driver, ensure that the ODBC datasource exists on the computer where the integrator runs.

The relational database metadata integrator collects metadata from secured universe connections that use JDBC or ODBC drivers.

To configure a universe connection source that uses a JDBC or an ODBC driver, perform the following steps:

1. Perform the following substeps for JDBC drivers:
 - a. Obtain the JDBC driver from your database server web site or utilities CD.
 - b. Unzip the JDBC driver into a folder such as:

```
c:\temp\teradata
```

2. Log into the Central Management Console (CMC) and open the Information Steward page.
3. Select **Manage > New > Integrator Source**.
The *New Integrator Source* dialog box opens.
4. Select *Relational Database* from the *Integrator Type* drop-down list.
5. Complete the options as described in the following table.

Option	Description
<i>Name</i>	Specifies a name for this integrator source. The maximum length is 128 characters.
<i>Description</i>	Optional. Specifies a description for the source.
<i>Connection Type</i>	Select <i>Universe Connection</i> from the drop-down list.
<i>Connections</i>	Specifies the name of the Central Management System (CMS) connection. The drop-down list displays the secure connections defined in the CMS.

Option	Description
<i>Table Schema</i>	<p>Optional. Specifies the name of the schema, or multiple schema names, to import from this source database.</p> <p>To specify multiple schemas, separate the name of each schema with a comma.</p> <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 5px;"> <p>Note</p> <p>When the schema name includes a comma, use double quotation marks around the schema name.</p> </div> <p>If you don't specify a schema name for Microsoft SQL Server or DB2 databases, the software imports all available schemas.</p> <p>If you don't specify a schema name for Oracle, the software uses the user name to import the schema.</p>
<i>Library Files</i>	<p>Specifies the full path to the Java library files. When there are multiple Java library files, enter each path and file name separated by semicolons.</p> <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 5px;"> <p>Note</p> <p>In a distributed deployment, set <i>Library Files</i> to the classpath on the computer where the integrator runs.</p> </div>

6. Select *Test Connection*.
7. Click *Save* after the connection test is successful.
The new source appears in the integrator sources list.

Task overview: [Configuring relational database system integrator sources \[page 213\]](#)

Related Information

[Configuring sources for JDBC connections \[page 216\]](#)

9.2.12 Configuring Excel Content integrator sources

The Excel Content metadata integrator collects metadata from a relational database source or a Business Intelligence source.

To perform the following process, you must belong to the Metadata Management administrator group and have the add and create rights to the integrator source.

1. Log into the Central Management Console (CMC) and open the Information Steward page.

2. Select **Manage > New > Integrator Source**.

The *New Integrator Source* page opens.

3. Select *Excel Metadata Content* from the *Integrator Type* drop-down list.
4. Enter a name for the integrator source in the *Name* text box, up to 128 characters.
5. Optional. Enter a description for the integrator source in the *Description* text box.
6. Either select an existing Excel workbook or download the template and create a copy.

To select an existing Excel workbook that you created previously based on an original template file:

- a. Select *Browse* next to *File Name* and select the Excel file.
- b. Select the applicable template model type from the *Model Name* drop-down list. Options are *Relational Database* or *Business Intelligence*.

Select:

- *Relational Database*
 - *Business Intelligence*
- c. Select *Save*.

To download the template and create a copy:

- a. Select the type of model from the *Model Name* drop-down list.

Select:

- *Relational Database*
 - *Business Intelligence*
- b. Select *Download template*.

A Windows download dialog box opens.

- c. Select **Save > Save as**.
 - d. Enter a new name for the template and select a location.
 - e. Select *Save*.
7. In the *New Integrator Source* dialog box, select *Browse* and select the Excel template that you created.
 8. Select *Save*.

The Excel Content metadata integrator source appears in the list integrator sources.

Configure the Excel workbook before you run the Excel Content metadata integrator source.

9.2.13 Configuring Amazon Redshift integrator sources

The Amazon Redshift metadata integrator supports the collection of metadata from an Amazon Redshift database, such as server instances, databases, schemas, and tables.

To perform the following process, you must have the create or add rights to the integrator source.

1. Log into the Central Management Console (CMC) and open the Information Steward page.
2. Select **Manage > New > Integrator Source**.

The *New Integrator Source* page opens.

3. Select *Amazon Redshift* from the *Integrator Type* drop-down list.

- Complete the options as described in the following table.

Option	Description
<i>Name</i>	Specifies a name for this integrator source. The maximum length is 128 characters.
<i>Description</i>	Optional. Specifies a description for the source.
<i>Server</i>	Specifies the name of the Amazon Redshift server.
<i>Database Port Number</i>	Specifies the port number for the Amazon Redshift database server.
<i>Database Name</i>	Specifies the name of the Amazon Redshift database to use as the source.
<i>Database User</i>	Specifies the name of the user authorized to access the Amazon Redshift database.
<i>Database Password</i>	Specifies the password for the user authorized to access the Amazon Redshift database.
<i>Table Schema</i>	Specifies the table schema name or names. When you specify more than one table schema name, separate the names with a comma.
<i>Library Files</i>	Specify the required library files for the Amazon Redshift source. Separate the library file names with a semicolon.

- Select *Test Connection*.
- Click *Save* after the connection test is successful.
The new Amazon Redshift source appears in the integrator sources list.

9.2.13.1 Defining an Amazon Redshift connection

Before you access any Amazon Redshift sources, create a connection to your Amazon Redshift database.

Ensure that you have database permission to read the metadata from Amazon Redshift tables. You can create connections in the Central Management Console (CMC) only when you're an administrator or have sufficient privileges.

SAP Information Steward supports Amazon Redshift engine 1.0 on Redhat Linux and the following Windows servers: 2008, 2008R2, 2012, and 2012R2 x64.

To create an Amazon Redshift connection, perform the following steps:

- Log into the CMC and open the SAP Information Steward page.
- Select the *Connections* node in the tree panel at left.
- Select **► Manage ► New ► Connection ►**.

The *Create Connection* page opens.

- Complete the options described in the following table.

Option	Description
<i>Connection Name</i>	Specifies the name for this Amazon Redshift connection. <ul style="list-style-type: none"> Maximum length is 64 characters Can be multibyte Case insensitive Can include underscores and spaces Can't include other special characters: ?!@#\$%^&*()-+={}[]:"';/\ ., `~ You can't change the name after you save the connection.
<i>Description</i>	Optional. Specifies a description for the connection.
<i>Connection Type</i>	Select <i>Database connection</i> .
<i>Purpose of connection</i>	Select <i>For data profiling</i> .
<i>Database Type</i>	Select <i>Amazon Redshift</i> .
<i>Database version</i>	Select <i>Redshift 8.X</i> or applicable version.

- Enter the relevant connection information for the database type.
- Select *Test connection*.
- Select *Save*.

Note

After you save the connection, you can't change its name, connection type, purpose and connection parameters, which uniquely identify a database.

The new connection appears in the list of connections.

Grant permission to specific users to access and use the connection.

9.3 Manage integrator sources and instances

Managing your integrator sources includes tasks such as adding or removing authorized users, purging source runs, and deleting sources.

Perform all management tasks for integrator sources in the Central Management Console (CMC).

The following table describes the management tasks to perform with integrator sources and instances, listed by the option you select in the *Manage* menu.

Manage option	Description
<i>New</i>	Create a new integrator source or source group.
<i>User Security</i>	Manage user security for integrator sources, source groups, or Metapedia objects.
<i>Refresh</i>	Obtain the latest integrator sources information.

Manage option	Description
Delete	Delete the selected integrator source configuration and its associated schedules, source runs, and logs.
Purge	Remove all source runs for the selected integrator. This option doesn't remove the source configuration, file logs, and schedules.

The following table describes the management tasks to perform with integrator sources and instances, listed by the option you select in the [Actions](#) menu.

Actions options	Description
History	View the current and previous executions of the selected integrator source.
Properties	View and edit the configuration information of the selected integrator source.
Schedule	Run the integrator at regular intervals.
Run now	Run the integrator immediately.

[Viewing and editing an integrator source \[page 223\]](#)

View and modify an integrator source description, connection information, and other information.

[Deleting an integrator source \[page 224\]](#)

When you no longer need the integrator source, or you want to rename the source, delete the source.

[Changing limit for storing integrator runs \[page 225\]](#)

Administrators can change the default limit of 100 runs for the number of integrator runs stored in the repository.

[Integrator run required \[page 226\]](#)

The software requires you to run an integrator source when an update installation of SAP Information Steward changes the metadata integrator.

9.3.1 Viewing and editing an integrator source

View and modify an integrator source description, connection information, and other information.

Before you perform the following steps, ensure that you have the view and edit rights for the metadata integrator source.

1. Log into the Central Management Console (CMC) and open the SAP Information Steward page.
2. Expand the [Metadata Management](#) node in the Tree panel at left and select [Integrator Sources](#).
3. From the list in the right pane, select the applicable integrator source.
4. Select **▸ Actions ▸ Properties ▾**.

The [Properties](#) page opens for the selected integrator source.

5. Change any property for the integrator source, except its type and name.
6. Select [Test Connection](#).

The software verifies the database connection information for the integrator source and displays a verification message.

7. Select *Save* when you're finished viewing and editing the integrator source.

The *Properties* page closes.

8. To edit an SAP BusinessObjects Enterprise metadata integrator source, perform the following steps:
 - a. Expand the *Schedule* node in the Tree panel at right.
 - b. Select Parameters.
 - c. Change settings for *Log Level* or *Update Option*.

Task overview: [Manage integrator sources and instances \[page 222\]](#)

Related Information

[Deleting an integrator source \[page 224\]](#)

[Changing limit for storing integrator runs \[page 225\]](#)

[Integrator run required \[page 226\]](#)

9.3.2 Deleting an integrator source

When you no longer need the integrator source, or you want to rename the source, delete the source.

Before performing the following task, make sure that you belong to the Metadata Management Administrator user group or that you have the right to delete integrator sources.

When you delete an integrator source, you also delete the collected metadata from the repository.

To delete an integrator source, perform the following steps:

1. Log into the Central Management Console and open the SAP Information Steward page.
2. Expand the *Metadata Management* node in the tree pane at left.
3. Select the *Integrator Sources* node.
A list of configured integrator sources appears at right with the date and time of the last run.
4. Select the integrator source and select ► *Manage* ► *Delete* ▾.

The software issues a confirmation.

5. Select *OK* to close the confirmation and delete the integrator source.

The integrator source is no longer listed in the list of integrator sources.

Task overview: [Manage integrator sources and instances \[page 222\]](#)

Related Information

[Viewing and editing an integrator source \[page 223\]](#)

[Changing limit for storing integrator runs \[page 225\]](#)

[Integrator run required \[page 226\]](#)

9.3.3 Changing limit for storing integrator runs

Administrators can change the default limit of 100 runs for the number of integrator runs stored in the repository.

Each time you run a metadata integrator or SAP Information Steward utility, the software creates a new instance and log files for the instance. By default, the software stores up to 100 runs in the repository. When this maximum number is exceeded, the software deletes the oldest instance and the associated log file to make room for the next run.

Note

The Purge utility deletes the database log in the repository for each instance that the software deletes.

Before you can perform the following task, ensure that you have *Full Control* access level on the Metadata Management folder in the Central Management Console (CMC).

To change the limits to delete integrator source instances, perform the following steps:

1. Log into the CMC and open the Information Steward page.
2. Expand the *Metadata Management* node in the Tree pane at left.
3. Select **► Actions ► Limits ►**.
The *Limits* page for Metadata Management opens.
4. To change the default value of 100 maximum number of instances to keep, perform the following substeps:
 - a. Select the checkbox for *Delete excess instances when there are more than N instances*.
 - b. Enter a new number in the text box under this option.
 - c. Select *Update* to save your changes.
5. To specify a maximum number of instances to keep for a specific user or group, perform the following substeps:
 - a. Select *Add* next to *Delete excess instances for the following users/groups*.
 - b. Select the user or group name from the *Available users/groups* pane and select the right-pointing arrow.
 - c. Select *OK*.
 - d. Enter a number in the text box under *Maximum instance count per object per user*.
 - e. Select *Update*.
6. To specify a maximum number of days to keep instances for a specific user or group:
 - a. Select *Add* next to *Delete instances after N days for the following users or groups*.
 - b. Select the user or group name from the *Available users/groups* pane and select the right-pointing arrow.

- c. Select *OK*.
- d. Enter a number in the text box under *Maximum instance count per object per user*.
- e. Select *Update*.

Task overview: [Manage integrator sources and instances \[page 222\]](#)

Related Information

[Viewing and editing an integrator source \[page 223\]](#)

[Deleting an integrator source \[page 224\]](#)

[Integrator run required \[page 226\]](#)

9.3.4 Integrator run required

The software requires you to run an integrator source when an update installation of SAP Information Steward changes the metadata integrator.

If you don't run an integrator source after an update installation that changes the metadata integrator, the software doesn't display metadata objects collected by the integrator. Information Steward notifies you to run the integrator source again in two locations:

- In SAP Information Steward, the directory page of the *Metadata Management* tab: A message states: "Run required: Ask your Information Steward administrator to run the integrator source in the CMC."
- In the Integrator Sources node of the Information Steward page in the Central Management Console (CMC): The *Last run* column, a message states "Run required for upgrade."

To collect the metadata in the repository and view the metadata objects in Information Steward, run the applicable integrator immediately, or wait for the next scheduled run.

Parent topic: [Manage integrator sources and instances \[page 222\]](#)

Related Information

[Viewing and editing an integrator source \[page 223\]](#)

[Deleting an integrator source \[page 224\]](#)

[Changing limit for storing integrator runs \[page 225\]](#)

[Running a Metadata Integrator \[page 227\]](#)

9.4 Running a Metadata Integrator

Run the Metadata Integrator to collect the metadata for each source that you configured.

When you select *Integrator Sources*, all configured integrator sources appear.

When you select an integrator source from the list in the right pane, you can run it immediately or define a schedule to run it.

To find Metadata Management integrator sources, follow these steps:

1. Open the Information Steward page in the CMC and select the Metadata Management node from the left pane.
2. Expand the *Integrator Sources* node. All configured integrator sources appear in the center pane.

Note

For the SAP Business Warehouse integrator to have authorization rights to collect metadata objects from SAP Business Intelligence (BI), you must install the correct support package based on the version of SAP BI you are running. For a list of the SAP BI versions and the applicable support package related to each, see the *Installation Guide*.

[Running a Metadata Integrator immediately \[page 227\]](#)

Instead of scheduling a metadata ingegrator, run it immediately.

[Defining a schedule to run a Metadata Integrator \[page 228\]](#)

To run a Metadata Integrator at regular intervals, define a schedule for it.

[Running SAP BusinessObjects Enterprise Metadata Integrator with Windows AD authentication \[page 229\]](#)

9.4.1 Running a Metadata Integrator immediately

Instead of scheduling a metadata ingegrator, run it immediately.

1. Log in to the Central Management Console (CMC) with a user name that belongs to the Metadata Management Administrator or Administrator user group.
2. Click Information Steward.
The *Information Steward* page opens with the *Information Steward* node selected in the tree panel at left.
3. Expand the *Metadata Management* node and click *Integrator Sources*.
4. From the list of configured sources that appears in the center pane, select the integrator source that you want by clicking anywhere on the row except its type.

Note

If you click the integrator *Source Type*, you display the version and customer support information for the integrator source. If you double-click the row, you open the *Properties* dialog box for the integrator source.

5. Click  *Actions*  *Run Now*  in the top menu tool bar.

→ Tip

You can also click the icon *Run selected object(s) now* in the icon bar under *Manage* and *Actions*.

- To view the progress of the integrator run, select the integrator source, and click **▶ Action ▶ History ▾**.

→ Tip

If you select *Now* in the *Run object* option under **▶ Action ▶ Schedule ▶ Recurrence ▾** and click *Schedule*, the *Integrator History* page automatically displays.

- Click the *Refresh* icon to update the status.
- If you use impact and lineage reports on the *Reports* option in the *View* drop-down menu, you must recompute the contents of the lineage staging table to incorporate changes from the integrator runs.

9.4.2 Defining a schedule to run a Metadata Integrator

To run a Metadata Integrator at regular intervals, define a schedule for it.

Ensure that you have the proper user or group rights to perform the following task.

- From the Central Management Console (CMC) click Information Steward.
The *Information Steward* page opens with the *Information Steward* node selected in the tree pane at left.
- Expand the *Metadata Management* node.
- Select the *Integrator Sources* node.
A list of configured integrator sources appears in the center pane with the date and time each was last run.
- From the list of configured sources that appears on the right, select the source from which you want to collect metadata by clicking anywhere on the row except its type.

ⓘ Note

If you click the *Source Type*, you display the version and customer support information for the metadata integrator.

- Click **▶ Actions ▶ Schedule ▾**.
- (Optional) To change the default instance title, select the *Instance Title* node on the left and change it to a unique name that describes this schedule.
- Select the *Recurrence* node on the left to choose the frequency in the *Run object* list.
Choose the additional relevant values as applicable.
- (Optional) To send notifications when the integrator has run, select the *Notification* node on the left. For more information about *Notification*, see the *SAP Business Intelligence platform Administrator Guide*.
- (Optional) To trigger the execution of a Metadata Integrator when an event occurs, select the *Events* node on the left. For more information about *Events*, see the *SAP Business Intelligence platform Administrator Guide*.
- (Optional) To change the default values for runtime parameters for the integrator, select the *Parameters* node on the left. (For runtime parameter information, see the topic Common runtime parameters for metadata integrators.)
- Click *Schedule*.

12. If you use impact and lineage reports on the [Reports](#) option in the [View](#) menu list of Information Steward, you must recompute the contents of the lineage staging table to incorporate changes from the integrator runs. Similar to setting up a regular schedule to run an integrator, you can set up a schedule to compute the lineage staging table at regular intervals.

9.4.3 Running SAP BusinessObjects Enterprise Metadata Integrator with Windows AD authentication

Before performing the steps below, ensure you have configured the SAP EIM Services server with the applicable JVM parameters for Windows Active Directory (AD) authentication.

Follow these steps to run the SAP BusinessObjects Enterprise Metadata Integrator with Windows AD authentication:

1. From the Central Management Console (CMC) click [Information Steward](#) in the [Organize](#) list and expand the [Metadata Management](#) node in the tree pane at left.
2. Click [Integrator Sources](#) and select the applicable integrator source name that appears in the main pane at right.
3. Select [Actions](#) > [Schedule](#).
4. Type the runtime parameters for the metadata integrator and select [Parameters](#) from the navigation tree.
5. In the [Command Line Parameters](#) text box enter the following JVM parameters using your file locations:
java.security.auth.login.config and java.security.krb5.conf.
For example, if your file location is c:/winnt/bsclogin.conf, the JVM parameters would be:

```
-Djava.security.auth.login.config=c:/winnt/bsclogin.conf  
-Djava.security.krb5.conf=c:/winnt/krb5.ini
```

6. (Optional) To change the default instance title, select [Instance Title](#) from the tree pane and change it to a unique name that describes this schedule.
7. Select [Recurrence](#) from the tree pane and choose the frequency in the [Run object](#) drop-down list.
8. Choose additional relevant values for the selected recurrence option as applicable.
9. (Optional) To receive a notification email when the integrator task has completed, select [Notification](#) from the tree pane. For more information about email notifications, see the [Business Intelligence Platform Administrator Guide](#).
10. (Optional) To trigger the execution of a metadata integrator when an event occurs, select [Events](#) from the tree pane. For more information about scheduling, see the [Business Intelligence Platform Administrator Guide](#).
11. Click [Schedule](#).

9.5 Runtime parameters for integrator sources

Runtime parameters change the normal or default actions when you run an integrator source.

Users with administrator rights can set runtime parameters for different integrator sources.

[Common runtime parameters for metadata integrators \[page 230\]](#)

There are common runtime parameters that apply to all metadata integrators.

[Runtime parameters for specific metadata integrators \[page 230\]](#)

There are runtime parameters that apply to specific metadata integrators.

[Changing runtime parameters for integrator sources \[page 247\]](#)

9.5.1 Common runtime parameters for metadata integrators

There are common runtime parameters that apply to all metadata integrators.

The following table describes common runtime parameters that are applicable to all metadata integrators.

Runtime parameter	Description
Database Log Level	<p>The level of database information to provide in the log.</p> <p>The default logging level is <i>Information</i>. You can change the default log level to a different level based on your needs. For example, change the level to <i>Integrator Trace</i> to provide more detailed information about the integrator run.</p> <p>View the database log in the Metadata Management repository after a task has completed. Or, view the log while the metadata integrator is running.</p>
File Log Level	<p>The level of file information to provide in the log.</p> <p>The default logging level is <i>Configuration</i>. You can change the default log level to a different level based on your needs. For example, change the level to <i>Integrator Trace</i> to debug the integrator run.</p> <p>Download and view the file log from the file repository server after the metadata integrator task completes.</p>
JVM Arguments	<p>Runtime parameters for Java processes.</p> <p>The Metadata Management job server creates a Java process to perform metadata collection.</p> <p>For example, if the metadata source is large, increase the default memory size.</p>
Additional Arguments	<p>Optional runtime parameters for the metadata integrator source.</p>

9.5.2 Runtime parameters for specific metadata integrators

There are runtime parameters that apply to specific metadata integrators.

The following table describes the integrators and the integrator-specific runtime parameters.

Integrator	Parameters
SAP BusinessObjects Enterprise	<p>Parameters to collect a subset of Central Management Console (CMC) objects.</p> <p>Universe Folders:</p> <ul style="list-style-type: none"> • Folder Name Expression • Collect Universes <p>Public Folders group and Personal Folders group:</p> <ul style="list-style-type: none"> • Folder Name Expression • Collect Web Intelligence documents and source universes • Collect Desktop Intelligence documents and source universes • Collect Crystal Reports and associated universes • Collect dashboards objects and associated universes
SAP Business Warehouse	<p>Update Options:</p> <ul style="list-style-type: none"> • Update existing objects and add newly selected objects • Delete Existing objects before starting object collection <hr/> <ul style="list-style-type: none"> • JVM Arguments • Additional Arguments
SAP HANA	<ul style="list-style-type: none"> • Parameters that adjust the number of threads used for the metadata collection. • Parameters to filter the queries or workbooks to collect. <p>Parameters that specify the SAP HANA Studio version to use for the configured HANA database.</p> <div style="border: 1px solid #ccc; background-color: #f0f0f0; padding: 10px; margin-top: 10px;"> <p>Note</p> <p>Applicable for users who have multiple versions of SAP HANA Studio installed, and who want to change the version that was set by the CMC during installation.</p> </div>
SAP Sybase PowerDesigner Metadata Integrator	<p>Update Option parameters determine which physical data models to collect upon connecting with the repository.</p> <p>The pdHomeDir parameter specifies the Power Designer installation location.</p>
Relational Database Amazon Redshift	<p>Parameters allowing you to enter simple expressions that include or exclude specified table names or view names at runtime.</p>

Integrator

SAP ERP Central Component

Parameters

Parameter to configure source information.

Note

The default for *JVM Arguments* is `-Dint_name=SAPIntegrator -Xmx6072m` (m indicates megabytes). This memory setting should cover most medium to large size ECC systems. Adjust the value if the default amount of memory is not enough.

[Runtime parameters for SAP BusinessObjects Enterprise integrator \[page 232\]](#)

Use several types of runtime parameters for the SAP BusinessObjects Enterprise metadata integrator.

[Runtime parameters for an SAP Business Warehouse integrator source \[page 239\]](#)

[Runtime parameters for SAP Sybase PowerDesigner Metadata Integrator \[page 241\]](#)

[Runtime parameters for Relational Database Metadata Integrator \[page 242\]](#)

[Runtime parameters for SAP ERP Central Component integrator source \[page 244\]](#)

Aside from the common runtime parameters, you can use `ProgramArgument []` `defineProgramArguments ()`, which allows you to configure source information.

[Runtime parameters for Amazon Redshift \[page 246\]](#)

The runtime parameters for the Amazon Redshift Integrator include parameters in which you can enter simple expressions that include or exclude specified table names or view names at runtime.

9.5.2.1 Runtime parameters for SAP BusinessObjects Enterprise integrator

Use several types of runtime parameters for the SAP BusinessObjects Enterprise metadata integrator.

The runtime parameters for SAP BusinessObjects Enterprise metadata integrator include the following:

- User collection parameters
- Selective Central Management Server (CMS) object collection
- Separate processes parameters
- SAP Enterprise XI 3.x system
- Metadata collection control for reports

[User collection parameters \[page 233\]](#)

User collection parameters for SAP BusinessObjects Enterprise Metadata Integrator adjust memory usage when collecting user permissions.

[Selective object collection \[page 234\]](#)

Selective collection of SAP BusinessObjects Enterprise Metadata Integrator objects in the SAP Central Management Server (CMS) can reduce processing time and provide a specific view of your Enterprise system objects.

[Separate processes parameters \[page 237\]](#)

The SAP BusinessObjects Enterprise (BOE) metadata integrator has a separate process for each applicable collection.

[SAP BusinessObjects Enterprise 3.x metadata collection with SSL \[page 238\]](#)

[Control metadata collection for reports \[page 239\]](#)

Parent topic: [Runtime parameters for specific metadata integrators \[page 230\]](#)

Related Information

[Runtime parameters for an SAP Business Warehouse integrator source \[page 239\]](#)

[Runtime parameters for SAP Sybase PowerDesigner Metadata Integrator \[page 241\]](#)

[Runtime parameters for Relational Database Metadata Integrator \[page 242\]](#)

[Runtime parameters for SAP ERP Central Component integrator source \[page 244\]](#)

[Runtime parameters for Amazon Redshift \[page 246\]](#)

9.5.2.1.1 User collection parameters

User collection parameters for SAP BusinessObjects Enterprise Metadata Integrator adjust memory usage when collecting user permissions.

The following table describes the runtime parameters for user collection parameters for SAP BusinessObjects Enterprise Metadata Integrator.

Runtime parameter	Description	Default value
collectUserPermissions	<p>Enable or disable user permissions collection</p> <p>Specify true to enable user permissions collection.</p>	false

Note

If you specify true, increase the memory with the `-Xmx` parameter in *JVM Arguments*. If the amount of memory is available, you can set this value as high as `-Xmx1500m`, where "m" indicates megabytes.

Set this parameter in *Additional Arguments* on the *Parameters* page.

9.5.2.1.2 Selective object collection

Selective collection of SAP BusinessObjects Enterprise Metadata Integrator objects in the SAP Central Management Server (CMS) can reduce processing time and provide a specific view of your Enterprise system objects.

To set options for selective collections, schedule the applicable integrator source in the Central Management Console (CMC) and complete the collection options in the *Parameters* page.

Use a selective CMS object collection strategy to run the integrator multiple times, and each time collect a different component. Use the specific update options to control the content of your collections and to control how to update existing collection objects. You could also incrementally collect CMS metadata for a large SAP BusinessObjects Enterprise deployment by using selective collection.

[Collecting objects selectively \[page 234\]](#)

When you schedule a metadata integrator source in the Central Management Console (CMC), choose options in the *Parameters* page to selectively collect objects.

[Collect metadata from multiple folders \[page 235\]](#)

Collect objects from universe, public, or personal folders when you schedule to run Enterprise integrator in the Central Management Console (CMC).

[Parameters page option descriptions \[page 235\]](#)

Collect objects from Universe, Public, or Personal folders by setting corresponding collection options under each group in the Parameters page in the Central Management Console (CMC).

9.5.2.1.2.1 Collecting objects selectively

When you schedule a metadata integrator source in the Central Management Console (CMC), choose options in the *Parameters* page to selectively collect objects.

There are two update options to choose in the *Parameters* page as described in the following table.

Update options

Option	Description
<i>Update existing objects and add newly selected objects</i>	Adds objects to a previous object collection for the integrator source, and updates existing objects that are new or have changed since the last run.
<i>Delete Existing objects before starting object collection</i>	Deletes the objects collected from previous runs of the integrator source and adds only the objects that you select for this run.

❖ Example

Update existing objects and add newly selected objects

For the first run of your integrator source, you select to collect only Web Intelligence document objects. On the next run, you select to collect Crystal Reports and associated universes in addition to the Web Intelligence document objects. With the *Update existing objects and add newly selected objects* option selected, the software does the following:

- Adds the Crystal Reports objects to the existing Web Intelligence document objects in the *Metadata Management* tab
- Collects only the Web Intelligence document objects that are new or have changed since the last run and adds them to the Metadata Management tab

❖ Example

Delete existing objects before starting object collection

You run the Enterprise metadata integrator on a specific source multiple times, and the software has collected numerous objects in the *Metadata Management* tab in Information Steward. To view only the Crystal Reports and associated universe objects in the *Metadata Management* tab, select *Delete Existing objects before starting object collection* when you schedule the next run in CMC. Then select the *Collect Crystal Reports and associated Universes* option, and deselect all other collection options. When the job runs, the software does the following:

- Deletes the current objects in the *Metadata Management* tab for the specific source
- Collects the Crystal Reports and associated universe objects to display in the *Metadata Management* tab of Information Steward for the specific source

9.5.2.1.2.2 Collect metadata from multiple folders

Collect objects from universe, public, or personal folders when you schedule to run Enterprise integrator in the Central Management Console (CMC).

Anyone using the *Metadata Management* tab in Information Steward can see the contents of the universe, public, or personal folders. However, the software requires the proper permissions before you can run collections on the objects in these folders. For example, only you or an administrator can run a collection on objects in your personal folder.

ⓘ Note

Metadata Integrator does not collect metadata from inboxes or categories.

9.5.2.1.2.3 Parameters page option descriptions

Collect objects from Universe, Public, or Personal folders by setting corresponding collection options under each group in the Parameters page in the Central Management Console (CMC).

The following table describes the options in the *Parameters* page of the CMC where you can select collection options, including the options described for selective collections.

Option	Description
<i>Folder Name Expression</i>	<p>Specifies the names of the folders that you want in the collection using a Java Regular Expression. There is a <i>Folder Name Expression</i> for Universe, Public, and Personal folders.</p> <p>The following list contains folder name expression examples:</p> <ul style="list-style-type: none"> • an asterisk (*) means collect all folders • <code>folderName</code> collects metadata within a specific folder. It also collects metadata in any associated sub-folders • <code>folderName*</code> an asterisk at the end of a string value includes all folders with that string • <code>folderName folderName</code> a pipe (also called vertical bar separator) between folder names includes one or the other folder. For example, <code>Sales Finance</code> means either the Sales or the Finance folder. • <code>^(?!folderName\$)</code> excludes the folder with this specific folder name • <code>^(?!folderName)</code> excludes any folders that start with this specific folder name • <code>(?!folderName)</code> excludes any folders that has this specific folder name within the name
<i>Collect Universes</i> (Universe Folders)	<p>Collects the SAP Enterprise universe metadata from the folders specified in the <i>Folder Name Expression</i> option.</p> <div data-bbox="804 1227 1396 1637" style="background-color: #f0f0f0; padding: 10px; border: 1px solid #ccc;"> <p>Note</p> <ul style="list-style-type: none"> • If you uncheck this option, and choose to collect any report that uses a universe, the integrator collects the universe and Lumira stories metadata as well. • If you select only this option and uncheck all of the report options, the integrator also collects Lumira datasets. • Only datasets built in Lumira 1.1.8 contain the universe references collected by the integrator. </div>
<i>Collect Web Intelligence Documents and source Universes</i>	<p>Collects Web Intelligence documents and source universes from Public or Personal folders.</p>
<i>Collect Desktop Intelligence Documents and source Universes</i>	<p>Collects Desktop Intelligence documents and source universes from Public or Personal folders.</p>
<i>Collect Crystal Reports and associated Universes</i>	<p>Collects Crystal Reports and associated universes from Public or Personal folders.</p>

Option	Description
<i>Collect Dashboards Objects and associated Universes</i>	Collects Dashboards Objects metadata from public or personal folders with UNX universe and BEX query as data sources.
<i>Collect Analysis OLAP Workspace</i>	Collects metadata from SAP BusinessObjects Analysis OLAP workspaces from public folders.
<i>Collect Design Studio Application</i>	Collects metadata from SAP BusinessObjects Design Studio Applications from public folders.
<i>Database Log Level</i>	Sets the level of detail in the database log located in the Metadata Management Repository.
<i>File Log Level</i>	Sets the level of detail in the file log located in the SAP installation directory.
<i>Update existing objects and add newly selected objects</i>	Collects a subset of physical data models based on the changes made since the last run. This option is selected by default. Using this option may reduce processing time.
<i>Delete Existing objects before starting object collection</i>	Collects all physical data models, overwriting models in the repository.
<i>JVM Arguments</i>	(Optional) Configures runtime parameters for the Java process.
<i>Additional Arguments</i>	(Optional) Configures runtime parameters for the metadata integrator source.

9.5.2.1.3 Separate processes parameters

The SAP BusinessObjects Enterprise (BOE) metadata integrator has a separate process for each applicable collection.

The BOE metadata integrator uses a separate process for each of the following metadata collections:

- Users, groups, and folders
- Reports
- Universes

Each process uses separate resources, memory and computer, to improve performance and throughput during the Metadata Integrator run.

The BOE metadata integrator provides the following runtime parameters for running separate processes. Set these parameters in *Additional Arguments* on the *Parameters* page.

Runtime parameter	Description	Default value
univProcessWorkLimit	Specifies the number of universes that are collected by each universe process. If the SAP BusinessObjects Enterprise Metadata Integrator requires more memory to run, decrease this value.	50
reportProcessWorkLimit	Specifies the number of reports that are collected by each report process. If the SAP BusinessObjects Enterprise Metadata Integrator requires more memory to run, decrease this value.	300

9.5.2.1.4 SAP BusinessObjects Enterprise 3.x metadata collection with SSL

You can configure an integrator source on SAP Information Steward to collect metadata from an SAP BusinessObjects Enterprise XI 3.x system. The Information Steward installation program installs a remote job server component on SAP BusinessObjects Enterprise XI 3.x to enable this collection. If SSL is enabled between the Information Steward and the remote job server systems, you must set the `businessobjects.migration` runtime parameter when you schedule the integrator source.

[Setting the runtime parameter for SAP Enterprise 3.x integrator with SSL \[page 238\]](#)

9.5.2.1.4.1 Setting the runtime parameter for SAP Enterprise 3.x integrator with SSL

When you schedule an integrator source to collect metadata from a BusinessObjects Enterprise XI 3.x system that has SSL enabled, you must set the `businessobjects.migration` runtime parameter

Note

You cannot run remote integrators if *Federal Information Processing Standards (FIPS)* mode is enabled on BusinessObjects Enterprise XI 3.x.

Follow these steps to set the runtime parameter for the SAP BusinessObjects Enterprise XI 3.x integrator source when SSL is enabled:

1. From the Central Management Console (CMC) click *Information Steward* in the *Organize* list and expand the *Metadata Management* node in the tree pane at left.
2. Expand *Metadata Management* and click *Integrator Sources*.
3. From the list of configured sources that appears on the right, select the source by clicking anywhere on the row except its type.

Note

If you click the source type, you display the version and customer support information for the Metadata Integrator.

4. Click **Actions** > **Schedule**.
5. In the **Schedule** window, click **Parameters** in the navigation tree.
6. In the text box for **JVM Arguments**, enter the following parameter:
`-Dbusinessobjects.migration=on`
7. Click **Recurrence** in the navigation tree to schedule this integrator source.
8. Click **Schedule**.

Related Information

[Recurrence options \[page 177\]](#)

9.5.2.1.5 Control metadata collection for reports

The SAP BusinessObjects Enterprise Metadata Integrator provides the following run-time parameters to control metadata collection for reports.

Set these parameters in **Additional Arguments** in the **Parameters** page of the Central Management Console (CMC).

Run-time parameter	Description	Default value
collectRptFldVar	You can disable collection of field levels for Crystal Reports and Web Intelligence documents. Field levels include: <ul style="list-style-type: none">• Report fields• Formula fields (Crystal Reports only)• SQL expression fields• Running total fields• Variables (Web Intelligence only)	true
collectReportsByUniverse	Restricts the collection of reports. When set to true, the integrator collects only reports referencing collected universes.	true

9.5.2.2 Runtime parameters for an SAP Business Warehouse integrator source

The SAP Business Warehouse metadata integrator provides runtime parameters for the following reasons:

- to adjust the number of threads to use when collecting metadata from the SAP system
- to filter the queries or workbooks to collect
- to collect APAP CDS Views datasources

The following table describes each runtime parameter.

Runtime parameters for SAP Netweaver Business Warehouse integrator source

Runtime parameter	Description
<i>Number of Threads</i>	<p>Specifies the number of threads to use when collecting metadata from the SAP system.</p> <div data-bbox="820 636 1402 801" style="background-color: #f0f0f0; padding: 10px;"> <p>❖ Example</p> <p>If your SAP Business Warehouse system has a large number of objects and it has available work processes, increase the number of threads.</p> </div> <p>The default value is 5. In a multiprocessor environment, you can increase this value.</p> <p>The number of threads is limited by the number of processors configured on the SAP Business Warehouse server. If the number of threads is greater than the available processors, the thread is put on a queue and will be processed when a processor becomes available.</p>
<i>Query Name Expression</i>	<p>Specifies the names of the queries that you want in the collection using a Java regular expression.</p> <div data-bbox="820 1173 1402 1834" style="background-color: #f0f0f0; padding: 10px;"> <p>❖ Example</p> <p>The following are examples of a Java regular expressions:</p> <ul style="list-style-type: none"> • *: An asterisk (*) means that all queries are collected. • queryName*: An asterisk at the end of a string value includes all queries with that string. • queryName queryName: A pipe, also called vertical bar separator, between query names includes one or the other query. For example, Sales Finance means either the Sales or the Finance query. • ^ (?!queryName\$): Excludes the query with this specific query name. • ^ (?!queryName): Excludes any queries that start with this specific query name. • (?!queryName): Excludes any queries that have this specific query name within the name. </div>

Runtime parameter	Description
<i>Workbook Name Expression</i>	Specifies the names of the workbooks that you want the integrator source to collect using a Java Regular Expression. Examples are similar to list of examples in the description for <i>Query Name Expression</i>
collectCDS	Adds the collection of ABAP CDS View DataSources when you run the metadata integrator with BW/4HANA. Include <code>collectCDS true</code> in the <i>Additional Arguments</i> option when you schedule the integrator.

Note

For the SAP Business Warehouse integrator to have authorization rights to collect metadata objects from SAP NetWeaver Business Intelligence, you must install the correct support package based on the version of SAP NetWeaver Business Intelligence you are running. For a list of the SAP NetWeaver Business Intelligence versions and the applicable support package related to each, see the *Installation Guide*.

Parent topic: [Runtime parameters for specific metadata integrators \[page 230\]](#)

Related Information

[Runtime parameters for SAP BusinessObjects Enterprise integrator \[page 232\]](#)

[Runtime parameters for SAP Sybase PowerDesigner Metadata Integrator \[page 241\]](#)

[Runtime parameters for Relational Database Metadata Integrator \[page 242\]](#)

[Runtime parameters for SAP ERP Central Component integrator source \[page 244\]](#)

[Runtime parameters for Amazon Redshift \[page 246\]](#)

9.5.2.3 Runtime parameters for SAP Sybase PowerDesigner Metadata Integrator

The PowerDesigner Metadata Integrator has runtime parameters for selecting which physical data models are collected when the Central Management Server (CMS) connects to the PowerDesigner repository and for specifying the Power Designer installation location.

Parameters	Description
Update Option	<p>These parameters are located in the <i>Parameters</i> page of the CMS and apply to only physical data models.</p> <ul style="list-style-type: none"> • <i>Delete existing objects before starting object collection</i> : Collects all physical data models, overwriting models in the repository. • <i>Update existing objects and add newly selected objects</i>: Collects a subset of physical data models based on the changes made since the last run. This is selected by default. <p>Using this option may reduce processing time.</p>
pdHomeDir	<p>Enter this parameter in the <i>Additional Arguments</i> field on the <i>Parameters</i> page to specify the Power Designer installation location.</p> <p>You would use this parameter if the integrator is unable to automatically detect the Power Designer installation or if the installation is using the wrong location. For example, to override the installation location you would enter <code>pdHomeDir "C:/Sybase/Power Designer SP03/"</code></p>

Parent topic: [Runtime parameters for specific metadata integrators \[page 230\]](#)

Related Information

[Runtime parameters for SAP BusinessObjects Enterprise integrator \[page 232\]](#)

[Runtime parameters for an SAP Business Warehouse integrator source \[page 239\]](#)

[Runtime parameters for Relational Database Metadata Integrator \[page 242\]](#)

[Runtime parameters for SAP ERP Central Component integrator source \[page 244\]](#)

[Runtime parameters for Amazon Redshift \[page 246\]](#)

9.5.2.4 Runtime parameters for Relational Database Metadata Integrator

The runtime parameters for the Relational Database Metadata Integrator include parameters in which you can enter simple expressions that include or exclude specified table names or view names at runtime.

Parameter	Description
Table Name Expression	<p>Enter a regular expression to include or exclude named tables.</p> <p>For example the following regular expression excludes all table names that contain "PRO":</p> <pre>^(?!.*(PRO)).*\$</pre> <p>The default setting is an asterisk (*), which means the expression collects all tables.</p> <div style="border: 1px solid #ccc; padding: 5px; background-color: #f9f9f9;"> <p>Note</p> <p>Even if the regular expression excludes the tables, all dependent objects appear in the Metadata Management tab, and all dependencies and data objects appear in the impact and lineage diagrams.</p> </div>
View Name Expression	<p>Enter a regular expression to include or exclude named views. For example, the following regular expression includes views with the names specified:</p> <pre>(.*(sales cost dates)).*\$</pre> <p>The default setting is an asterisk (*), which means the expression collects all views.</p> <div style="border: 1px solid #ccc; padding: 5px; background-color: #f9f9f9;"> <p>Note</p> <p>Even if the regular expression excludes the views, all dependent objects appear in the Metadata Management tab, and all dependencies and data objects appear in the impact and lineage diagrams.</p> </div>
Include public synonyms (For Oracle only)	<p>For Oracle databases only, specifies whether all public synonyms are to be collected. It is selected by default. When selected, all public synonyms are collected. This option does not appear for any other database types.</p>

Parent topic: [Runtime parameters for specific metadata integrators \[page 230\]](#)

Related Information

[Runtime parameters for SAP BusinessObjects Enterprise integrator \[page 232\]](#)

[Runtime parameters for an SAP Business Warehouse integrator source \[page 239\]](#)

[Runtime parameters for SAP Sybase PowerDesigner Metadata Integrator \[page 241\]](#)

[Runtime parameters for SAP ERP Central Component integrator source \[page 244\]](#)

[Runtime parameters for Amazon Redshift \[page 246\]](#)

9.5.2.5 Runtime parameters for SAP ERP Central Component integrator source

Aside from the common runtime parameters, you can use `ProgramArgument[]` `defineProgramArguments()`, which allows you to configure source information.

The following table shows the available options for this parameter:

Arguments for `ProgramArgument[]` `defineProgramArguments()`

Argument name	Type	Description
<code>SAPConnType</code>	String	The connection type for the source: <ul style="list-style-type: none">• <code>AppServer</code>• <code>GroupServer</code>
<code>SAPClientId</code>	String	The SAP client number.
<code>SAPSysId</code>	String	The SAP system ID.
<code>SAPHost</code>	String	The AppServer host name.
<code>SAPMHost</code>	String	The host name of the message server.
<code>SAPGroup</code>	String	The group name when in GroupServer mode.
<code>SAProuter</code>	String	The SAP router string.
<code>SAPSysNum</code>	String	The system number.
<code>SAPUserId</code>	String	The name of the user connecting to the SAP system.
<code>SAPPassword</code>	String	The password for the user connecting to the SAP system.
<code>SAPLang</code>	String	The ISO Language code. For example, <code>en</code> .

Parent topic: [Runtime parameters for specific metadata integrators \[page 230\]](#)

Related Information

[Runtime parameters for SAP BusinessObjects Enterprise integrator \[page 232\]](#)

[Runtime parameters for an SAP Business Warehouse integrator source \[page 239\]](#)

[Runtime parameters for SAP Sybase PowerDesigner Metadata Integrator \[page 241\]](#)

[Runtime parameters for Relational Database Metadata Integrator \[page 242\]](#)

[Runtime parameters for Amazon Redshift \[page 246\]](#)

[Common runtime parameters for metadata integrators \[page 230\]](#)

[Configuring SAP ERP Central Component integrator sources \[page 197\]](#)

9.5.2.5.1 Selective object collection

To execute your scheduled SAP ERP Central Component metadata integrator jobs more quickly and efficiently, configure the job to collect updated objects instead of all objects.

To add new objects, modify updated objects, and remove deleted objects from the Metadata Management repository, schedule executions of your SAP ERP Central Component metadata integrator jobs, and select objects to collect based on your requirements.

Selective object collection makes sense, especially since several SAP ERP objects, such as ECC and SAP S/4HANA objects, remain static for long periods of time. Selective object collection focuses your jobs on objects that were added, updated, or deleted since the last job execution. Selective object collection significantly reduces processing time and reduces the load on the Metadata Management repository database.

Note

Currently, selective object collection is for SAP DDIC (ABAP server data dictionary) tables and views and their child objects.

Collection options

Before you select collection options, complete the following options based on the descriptions.

Option	Description
<i>Number of Threads</i>	The number of processors to run at the same time to process the job. The default setting is 5.
<i>Root Data Model</i>	Enter the ERP technical data model or models to which processing is restricted. To include multiple values, separate each model with a comma, excluding any spaces. Example SD , CA , ICM For more information about technical data models, see your ERP documentation.

Configure your scheduled SAP ERP Central Component metadata integrator jobs in the *Parameters* dialog of the Central Management Console. The following list contains the collection options:

- *Collect Data Elements*
- *Collect Domains*
- *Collect View Lineage*
- *Collect Domain Value Extractor*
- *Collect Function Model Extractor*
- *Collect CDS Views*
- *Collect All Tables*

In addition to selecting the objects to collect to run jobs more quickly and efficiently, select an *Update Option* in the *Parameters* dialog. The following table describes the options.

Option	Description
<i>Delete existing objects before starting object collection</i>	<p>Collects all current objects after first deleting existing objects from the repository.</p> <p>Select this option for the first time you run the SAP ERP integrator. When you select this option, the job takes a lot longer to process than if you select <i>Update existing objects and add newly selected objects</i>.</p>
<i>Update existing objects and add newly selected objects</i>	<p>Collects a subset of objects based on the changes made since the last run. Using this option can reduce processing time.</p> <p>Select this option for delta runs. This option is selected by default.</p>

9.5.2.6 Runtime parameters for Amazon Redshift

The runtime parameters for the Amazon Redshift Integrator include parameters in which you can enter simple expressions that include or exclude specified table names or view names at runtime.

Parameter	Description
Table Name Expression	<p>Enter a regular expression to include or exclude named tables.</p> <p>For example, to exclude object names containing word "temp", the expression will be:</p> <p><code>^(?!.*(temp)).*\$</code></p> <p>The default setting is an asterisk (*), which means the expression collects all tables.</p>

Note

Even if the regular expression excludes the tables, all dependent objects appear in the Metadata Management tab, and all dependencies and data objects appear in the impact and lineage diagrams.

Parameter	Description
View Name Expression	<p>Enter a regular expression to include or exclude named views. For example, the following regular expression includes object names such as svn, cvs, nvs, mvc:</p> <p><code>.*(svn cvs nvs mvc).*</code></p> <p>The default setting is an asterisk (*), which means the expression collects all views.</p> <div style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 10px; margin-top: 10px;"> <p>Note</p> <p>Even if the regular expression excludes the views, all dependent objects appear in the Metadata Management tab, and all dependencies and data objects appear in the impact and lineage diagrams.</p> </div>

Parent topic: [Runtime parameters for specific metadata integrators \[page 230\]](#)

Related Information

[Runtime parameters for SAP BusinessObjects Enterprise integrator \[page 232\]](#)

[Runtime parameters for an SAP Business Warehouse integrator source \[page 239\]](#)

[Runtime parameters for SAP Sybase PowerDesigner Metadata Integrator \[page 241\]](#)

[Runtime parameters for Relational Database Metadata Integrator \[page 242\]](#)

[Runtime parameters for SAP ERP Central Component integrator source \[page 244\]](#)

[Common runtime parameters for metadata integrators \[page 230\]](#)

9.5.3 Changing runtime parameters for integrator sources

When you schedule an integrator source, you can change the default values of the runtime parameters in the *Parameters* page in the Central Management Console (CMC).

Before you follow the steps below, make sure you understand all of the settings in the *Parameters* page.

To change the runtime parameters for an integrator source:

1. Log in to the Central Management Console (CMC) with a user name that belongs to the Metadata Management Administrator group and click *Information Steward* on the CMC home page.
2. Expand the *Metadata Management* node.
3. Select the *Integrator Sources* node, and select the applicable integrator source.
4. Click **► Actions ► Schedule ►**.
The *Schedule* page opens.

5. Click [Parameters](#) in the navigation tree at left.
The [Parameters](#) page opens in the center pane.
6. Complete the applicable parameters.
7. (Optional) In the [JVM Arguments](#) box, enter the runtime parameter and the value you want to change.
8. (Optional) In the [Additional Arguments](#) box, specify a runtime parameter for a specific integrator source.
9. Click [Recurrence](#) in the navigation tree.
The [Recurrence](#) dialog opens.
10. Enter schedule options to schedule this integrator source.
11. Click [Schedule](#).

Related Information

[Runtime parameters for specific metadata integrators \[page 230\]](#)

[Recurrence options \[page 177\]](#)

9.6 Viewing integrator run progress, history, and log files

You can view date information, history, and database logs for an integrator source. Open the Central Management Console (CMC) and navigate to the [Information Steward](#) area and choose [Metadata Management](#) [Integrator Sources](#).

A list of all existing integrator sources appears in the main pane. Each integrator source contains additional information as applicable in the following columns:


- Name
- Type
- Category
- Last Run
- Description
- Created

To view more information about a single integrator source:

1. Select the name of the applicable integrator source from the list of all configured integrator sources and click [Actions](#) [History](#).
The [Integrator History](#) dialog opens with the following schedule information shown in columns:

Column name	Description
Schedule Name	Name provided in the Instance Title box when the instance was scheduled.

Column name	Description
Schedule Status	The status of the integrator run. Values include: <ul style="list-style-type: none"> • Success • Failed • Pending • Running • Paused • Recurring • Stopped
Start Time	Time the run started.
End Time	Time the run ended.
Duration	Total time for the run.
Log File	Link to the log file that contains progress messages for the run.

- View the database log file while the integrator source run is in progress. Either select **Actions** > **Database Log** or click the *View the database log* icon () in the menu bar at the top of the *Integrator History* dialog. Use the *Refresh* option in the database log to reveal additional information as the run progresses.
- After the integrator source has finished or has failed, use the methods from step 2 to view information, or click the *Download* link in the *Log File* column.
- Click the “X” icon in the upper right corner of the database log page to close the log file.

By default, Metadata Management writes high-level messages (number of universes processed, number of reports processed, and so on) to the log. You can change the message level on the configuration page for the integrator source.

Related Information

[Changing Metadata Management log levels \[page 341\]](#)

[Information Steward logs \[page 340\]](#)

9.7 Other options in the Integrator History page






There are other options in the Integrator History page that enables you to obtain history information from the integrator run.

At the top of the Integrator History page in the main pane, there is a menu bar that includes addition actions:

- An *Actions* list menu
- Various icons for specific actions

Action drop list menu	Description
Run Now	Run the selected integrator immediately.
Pause	Pauses the instance that is scheduled to run or scheduled as a recurring instance. The selected instance is not canceled.
Reschedule	Define a new schedule to execute the selected integrator.
Database Log	View the database log file.
Refresh	Refresh the status of the selected integrator instance.
Delete	Delete the selected integrator instance from the history list.

The icons that are shown in the top menu bar are slightly different than the options in the *Actions* menu.

Icon	Action	Description
	Rerun an instance	Rerun an existing instance.
	Stop the selected instance	Stop running the instance.
	Pause selected instance	Pause the instance that is scheduled to run or scheduled as a recurring instance. The selected instance is not canceled.
	Resume paused instance	Restart the paused instance. Only available when the instance has been paused.
	View the database log	View messages that indicate which metadata objects have been collected from the integrator source.

9.8 Troubleshooting

You may encounter some error and warning messages as you view database logs and file logs for each Metadata Integrator run. The error and warning messages are related to the following items:

- Crystal Reports
- Desktop Intelligence document
- Out of memory error
- Parsing failure errors
- SQL parse errors
- Data Federator Designer error

9.8.1 Crystal Report message

```
Crystal Report [reportname]. Unable to find class for universe object
[objectname]. Universe class cannot be uniquely identified for object
[objectname]. Data association cannot be established directly through the object.
Reference will be established directly to the column.
```

Cause: The SAP BusinessObjects Enterprise Metadata Integrator cannot uniquely identify a universe object that is used to create a Crystal Report when the object has the same name as another object in a different universe class. Therefore, the integrator cannot establish the correct relationship between the universe object and the report. However, the integrator can establish a relationship between the source table or column and the report because the SQL parser can find the source column used by the universe object.

Action: Name the objects uniquely across the classes in a universe.

9.8.2 Desktop Intelligence document error

DeskI document [reportname]. Universe [] not found.

Cause: Data providers in the Desktop Intelligence document refer to an invalid or non-existent universe.

Action: Open the Desktop Intelligence document and edit the data provider to specify a valid universe.

9.8.3 Out of memory error

Error occurred during initialization of VM
Could not reserve enough space for object heap

Cause: The SAP BusinessObjects Enterprise Metadata Integrator does not have enough memory to run.

Action: Decrease the value of the MaxPermSize run-time parameter in the *JVM Arguments* on the *Parameters* page when you schedule the integrator source. For example, enter -XX:MaxPermSize=256m and rerun the Metadata Integrator.

9.8.4 Parsing failure

Parsing failure. Unable to find columns for SQL with select *.

Cause: The SAP BusinessObjects Enterprise Metadata Integrator cannot collect the metadata for a universe derived table if the SQL used in the derived table is of the form SELECT * FROM TABLE.

Action: Always use the fully-qualified column names in the projection list of the SELECT clause.

9.8.5 Parsing failure for derived table

Parsing failure for derived Table <table_name>. Unable to find table associated with column <column_name>

Cause: The SQL parser in the SAP BusinessObjects Enterprise Metadata Integrator requires column names in a derived table to be qualified by the table name: <table_name.column_name>. If you do not qualify the column name, the Metadata Integrator cannot associate the column to the correct table.

Action: Fully qualify the column reference or add the tables used by the derived tables to the universe. The Metadata Integrator treats the universe tables as a system catalog to find the table and column references.

9.8.6 Unable to parse SQL

Unable to parse SQL ... `<error message>`

Cause: The SQL parser in Metadata Management has limited parsing capabilities to extract column names and table names to build the relationships. For example, if a Metadata Integrator fails to parse the SQL for a view, it cannot build the source-target relationship between the view and the table or tables upon which the view is based.

However, the Metadata Integrators collect the SQL statement and the Metadata Management Explorer displays it.

Action: Analyze the SQL statement in the Metadata Management Explorer and establish a user-defined relationship for these tables and columns.

9.8.7 Unable to retrieve SQL to parse

Unable to retrieve SQL to parse. `<error message>`

The cause of this warning message can be one of the following:

Cause: You do not have sufficient privilege to extract metadata about Web Intelligence documents.

Action: Take one of the following actions:

- Have your administrator change your security profile to give you permission to refresh Web Intelligence documents.
- Run the Metadata Integrator with a different user id that has permission to refresh Web Intelligence documents.

Cause: If a database connection is not configured for Trusted Authentication in SAP BusinessObjects Business Intelligence (BI) Platform, you must supply the user id and password at runtime. If you try to collect metadata for a report that uses a non-Trusted connection to the database, the report collection fails.

Action: Configure both your SAP BusinessObjects BI Platform server and client to enable Trusted Authentication. For details, see the *SAP BusinessObjects Business Intelligence Platform Administrator Guide*.

Cause: The extract for Web Intelligence documents fails if you create your Web Intelligence documents with the *Refresh on Open* option and the computer on which you run the SAP BusinessObjects Enterprise Metadata Integrator does not have connection to the source database on which the reports are defined.

Action: Take one of the following actions:

- Run SAP Information Steward on the computer where SAP BusinessObjects Enterprise is installed
- Define the database connection on the computer where you run the SAP Metadata Management.

9.8.8 Connection with Data Federator Designer

If the Data Federator Integrator connects successfully to the Data Federator Designer, but Data Federator returns an error:

1. Login to the Data Federator Designer.
2. Within Data Federator Designer, you can obtain a more detailed error message.

9.9 Grouping Metadata Integrator sources

SAP Information Steward provides the capability to group metadata sources into groups such as Development System, Test System, and Production System. After the groups are defined, you can view impact and lineage diagrams for a specific Source Group.

Related Information

[Creating source groups \[page 253\]](#)

[Modifying source groups \[page 253\]](#)

9.9.1 Creating source groups

You must have been assigned the right to add objects before you can follow these steps.

1. Log in to the Central Management Console (CMC) and access the *Information Steward* area.
2. Select **Metadata Management > Integrator Sources** from the tree in the left pane.
A list of existing integrator sources appears in the main pane.
3. Click **Manage > New > Source Group**.
The *Source Group* page opens.
4. Enter a name and description for the new source group in the *Name** and *Description* boxes.
5. Select the checkbox to the left of each integrator source that you want to add to the new source group and click *Save*.

You can see the new source group listed with the other source groups by choosing **Metadata Management > Source Groups**.

9.9.2 Modifying source groups

You must have been assigned the right to edit objects before you can follow these steps.

1. Log in to the Central Management Console (CMC) and access the *Information Steward* area.
2. ▶ *Metadata Management* ▶ *Source Groups* ▶.
3. Right click the source group that you want to modify and select *Properties*.
The *Source Group* dialog box opens.
4. You can change the *Name* and *Description* of the source group, and you can deselect any source that you don't want to include in the group. When you are finished making edits, click *Save*.
5. Right-click the applicable source group and click *User Security* to do any of the following tasks:
 - Add or remove principals to this source group.
 - View security for a selected principal.
 - Assign security for a selected principal.
6. Click *Save*.

9.9.3 Deleting source groups

You must have been assigned the right to delete objects before you can follow these steps.

1. Log in to the Central Management Console (CMC) and access the *Information Steward* area.
2. Select ▶ *Metadata Management* ▶ *Source Groups* ▶.
A list of source groups appears in the main pane.
3. Right-click the source group that you want to delete, and then select *Delete* from the list.
4. Click *OK* to confirm the deletion.

9.10 Configuring BI Launch Pad for Information Steward

BI Launch Pad enables users on SAP BusinessObjects Business Intelligence (BI) platform to access Information Steward on a different computer with SAP BusinessObjects Information platform services (IPS).

Follow these steps when your Information Steward installation is on IPS, and your BI Launch Pad users view the following items on a different computer:

- Lineage reports, such as Crystal or Web Intelligence
 - Metapedia terms or categories associated with the lineage report
1. On the computer where IPS is installed, log in to the Central Management Console (CMC).

ⓘ Note

The software doesn't support using the localhost URL. When you open the CMC web interface, make sure that you use the webserver name instead of "localhost". For example, instead of using `http://localhost:<port_number>/BOE/CMC`, use `http://<webserver>:<port_number>/BOE/CMC`.

2. Select *Applications* from the navigation dropdown list located at the top left of the CMC.
3. Right-click on *Information Steward Application* in the *Application Name* list and select *BI Launch Pad Integration* from the dropdown menu.

4. Enter the connection information for the BI platform Central Management Server (CMS) for your BI Launch Pad application and click [Save](#).

Grant view access to the user who accesses Information Steward from BI Launch Pad. For instructions, read about assigning principals to an access control list in the *SAP BusinessObjects Business Intelligence platform 4.1 Administrator Guide*.

Additionally, for information about relationships for universe objects and reports, see the *User Guide*.

9.11 Displaying user-defined reports in Information Steward

You can display Crystal reports in tables in the Information Steward repository.

Reports must be created in Crystal Reports version 12 or later.

Use a properties file to define the business question that report results will answer (for example, "which tables have similar names in different integrator sources").

1. Create a custom reports directory on a single web application server or on multiple servers.

Create the directory on a single web application server when the directory path should be available on the web application server machine.

Create the directory on multiple web application servers when the same directory path and content is maintained on each server machine or when an NFS-shared directory path that is available from all application servers is used. (In the latter case, the user account under which the application server is running must have read permissions on the directory contents.)

2. For each Crystal report, create a properties file and name it using one of the following conventions:

- When the Crystal report is in the default language of the operating system, use `reportfilename.rpt.properties`.
- When the Crystal report uses a different preferred viewing locale than the operating system, use `reportfilename.rpt_<language>.properties`.

For example, a properties file might be named `TableNamesInMultipleSources.rpt.properties` when the associated Crystal report is in the same language as the operating system or named `TableNamesInMultipleSources.rpt_fr.properties` when the report is in French but the operating system is in Spanish.

3. In the properties file, add a name entry with the following attributes:

Attribute	Description
<code>name</code>	Name of the Crystal report to display in the Information Steward web application.
<code>description</code>	Description of the Crystal report to display in the Information Steward user interface.
<code>question</code>	Business question that this report answers.

Attribute	Description
<code>categories</code>	Category(ies) that this report belongs to. Use a semicolon (;) to separate multiple categories. Categories can be predefined, or you can create new categories.
<code>use_is_repo_connection</code>	Source system for the tables on which a Crystal report will run: <ul style="list-style-type: none"> Set to <code>True</code> to connect to the Information Steward repository. Set to <code>False</code> to use the repository connection defined in the report. The default value is <code>True</code> .

The following example shows a name entry in a properties file for a new custom report. The new report will list table names repeated in multiple integrator sources and put the table names in two categories—"Table name reports" for new categories and "Usage reports" for existing categories.

```
name=Table Names in Integrator Sources
description=Lists each table name and the integrator source names that
contain the table.
question=What tables have similar names in different integrator sources?
categories=Table name reports,Usage reports
use_is_repo_connection=true
```

4. Add each Crystal report and associated properties file to the custom reports directory.
5. Specify the custom reports directory in the Central Management Console (CMC):
 - a. Log in to the CMC with a user name that belongs to the Metadata Management Administrator group or to the Administrator group.
 - b. In the navigation list at the top of the CMC home page, select [Applications](#).
 - c. In the [Applications Name](#) list, select [Information Steward](#).
 - d. Select [Configure Application](#) in the navigation pane at left.
 - e. Find Metadata Management Options group.
 - f. Enter a full path for the directory location in [Custom Reports Directory](#).
 - g. Click [Save](#).

10 Cleansing Package Builder Administration

10.1 Changing ownership of a cleansing package

Ownership can be reassigned only by an Information Steward administrator. You can change ownership only for private cleansing packages. Published cleansing packages are either not owned (unlinked) or linked to private cleansing packages and therefore when you change the ownership of a private cleansing package, the new owner automatically can republish to the linked cleansing package.

To reassign ownership of a cleansing package:

1. Log in to the Central Management Console (CMC) as an administrator.
2. At the CMC home page, click [Information Steward](#).
3. In the left pane, expand the [Cleansing Packages](#) node and then click [Private](#).
4. In the right pane, right-click the cleansing package whose owner you want to reassign. In the pop-up menu, choose [► Properties ▾](#) to open the Properties window.

Other ways of accessing the Properties window:

- Choose [► Actions ▸ Properties ▾](#).
 - Click the link in the [Name](#) or [Kind](#) column for the desired cleansing package.
5. In the Properties window, choose a different owner in the [Owner](#) drop-down list, and click [Save](#).

The new owner is now assigned to this cleansing package.

10.2 Deleting a cleansing package

To delete a private or published cleansing package:

1. Log in to the Central Management Console (CMC) as an administrator.
2. At the CMC home page, click [Information Steward](#).
3. In the left pane, expand the [Cleansing Packages](#) node and then click either [Private](#) or [Published](#).
4. In the right pane, right-click the cleansing package that you want to delete. In the pop-up menu, choose [► Manage ▸ Delete ▾](#).
5. In the confirmation window that opens, click OK.
The selected cleansing package is deleted.

10.3 Changing the description of a cleansing package

You can edit the description of published or private cleansing packages.

To edit the description:

1. Log in to the Central Management Console (CMC) as an administrator.
2. At the CMC home page, click *Information Steward*.
3. In the left pane, click *Cleansing Packages* and then click *Private* or *Published*.
4. In the right pane, right-click the cleansing package whose description you want to edit and select *Properties*.

Other ways of accessing the Properties window:

- Choose **► Actions ► Properties ▾**.
 - Click the link in the *Name* or *Kind* column for the desired cleansing package.
5. In the Properties window, edit the description, and click *Save*.

10.4 Unlocking a cleansing package

The status of a cleansing package is displayed in the popup dialog that appears when you hover your mouse pointer over the name of the cleansing package in the *Cleansing Package Tasks* screen and is also indicated by the cleansing package icon.

It may take some time for a cleansing package with a Busy status to complete the operation and change to a Ready state. The state of a cleansing package with a Busy status cannot be changed by an Information Steward administrator. You can either wait for the operation to complete or delete the cleansing package.

When a cleansing package is opened for editing, it enters a locked state so that no other user may edit it. To close a cleansing package, you must either return to the *Cleansing Package Tasks* screen, switch to another cleansing package, or log off from Cleansing Package Builder. If the browser window is closed or the computer is shut down without logging off, the cleansing package may become locked.

Note

When making significant changes to a cleansing package, you may want to save a copy by clicking *Save As* in the *Cleansing Package Tasks* window. If your cleansing package becomes locked, you will have a copy to work on.

A cleansing package may become locked when it is in any of the following states:

- Open for Read Write
- Canceling analysis
- Canceling publish
- Analyzing
- Publishing

- Content upgrading

To unlock a cleansing package:

1. (Data steward) When you encounter a locked cleansing package in the *Cleansing Package Tasks* screen, ask your Information Steward administrator to unlock it.
2. (Information Steward administrator) Change the cleansing package state from Locked to Error.
 - a. Log in to the Central Management Console (CMC).
 - b. Select Information Steward.
 - c. Expand the *Cleansing Package* node and select *Private* or *Published*.
 - d. Right-click the desired cleansing package, and choose *Properties*.
 - e. Change the state to Error and click *Save*.
 - f. Notify the data steward that the cleansing package state is updated to Error.
3. (Data steward) When notified that the cleansing package is unlocked and moved to the Error state, do the following:
 - a. From the *Cleansing Package Tasks* screen, open the cleansing package.
 - b. Verify the condition of the cleansing package and that it displays information as expected.
 - c. Close the cleansing package.

If the cleansing package is returned to a Ready state and its condition was as you expected, you may use it.

If the cleansing package returns to the Error state or the condition was not as expected, the cleansing package is corrupt and should be deleted.

Related Information

[Cleansing package states and statuses \[page 261\]](#)

10.5 Content upgrade for person and firm cleansing packages

Run a content upgrade to apply release changes to your applicable private SAP-supplied person and firm (global) cleansing packages.

Note

We notify you when a content upgrade is available in the “Changes in version <current version>” topic in the *Upgrade Guide*, which accompanies each Information Steward release package.

The content upgrade process can upgrade the following objects, when applicable:

- Standard forms
- Variations

- Reference data
- Rules

If you have created copies of the SAP-supplied person and firm cleansing package and modified them, you can perform a content upgrade on your private copies without losing any modifications.

Cleansing packages that are eligible for upgrade are denoted by a different icon. A cleansing package is not eligible for a content upgrade if it was:

- created using the *Create Person and Firm Cleansing Package* dialog box.
- created as a single domain person and firm cleansing package.
- already upgraded to the latest version.

You can choose to automatically publish a private cleansing package through the content upgrade process. If a published package already exists for the private cleansing package, the software lists the name of the existing cleansing package in the *Content Upgrade* dialog box, and republishes the cleansing package when the upgrade process completes.

10.5.1 Content upgrade states

During content upgrade, the cleansing package icon indicates that the package is locked or busy.

When content upgrade completes, there may be warnings even when the icon indicates the cleansing package is “ready”. An error icon indicates that there were errors during the content upgrade.

Note

When errors occur during content upgrade, the software does not make any content changes to that cleansing package. You must fix the errors and run a content upgrade again on that cleansing package.

To find out if there are warnings for a cleansing package in the ready state, move your mouse pointer over the cleansing package name for a pop-up message. If there is no message, the content upgrade was successful. Otherwise the pop-up message indicates that you should contact your administrator to view the log file for a list of warning messages.

If an error icon appears with the cleansing package, you can access the pop-up message in the same manner as for warnings. The pop-up message instructs you to contact your administrator.

The table below describes the content upgrade states.

State	Description
Locked	<ul style="list-style-type: none"> • Content upgrade in process • Publishing in process (when applicable)
Ready	<ul style="list-style-type: none"> • Content upgrade complete • Content upgrade complete with warnings (when applicable)
Error	Content upgrade failed and the software did not upgrade the cleansing package

Related Information

[Cleansing package states and statuses \[page 261\]](#)

10.5.2 Content upgrade log file

The software writes all errors and warnings for a content upgrade to a log file.

Administrators can find the log file in the platform's logging folder.

The log file name includes the name of the service where Cleansing Package Builder's Publishing Service is installed. By default, the location is in the EIM Adaptive Processing Server. Use the date and time listed in the file explorer to help determine the correct log file.

Example

In Windows, find the log file in `<%BOE_Install%>\SAP BusinessObjects Enterprise XI
<current_version>\logging`

Example

Log file name: `aps_<system_name>.EIMAdaptiveProcessingServer_trace.<number>.glf`

10.6 Cleansing package states and statuses

View the states and statuses of cleansing packages in either SAP Information Steward or the Central Management Console.

The following table describes where and how to view cleansing package states and statuses.

Location	Description
Cleansing Package Tasks	Hover your mouse pointer over the desired cleansing package to display the properties sheet.

Location	Description
Information Steward page of the Central Management Console	<p>You must have Information Steward administrator privileges to perform these steps:</p> <ol style="list-style-type: none"> 1. Log in to the Central Management Console (CMC). 2. Select Information Steward. 3. Expand the <i>Cleansing Package</i> node and select <i>Private</i> or <i>Published</i>. 4. Right-click the desired cleansing package and choose <i>Properties</i>.

[Cleansing package states \[page 262\]](#)

A cleansing package can have one of several states.

[Cleansing package statuses \[page 263\]](#)

A cleansing package may have one of the following statuses: Ready, Busy, Locked, Error.

10.6.1 Cleansing package states

A cleansing package can have one of several states.

You can see a cleansing package state in the properties sheet. View the properties sheet by hovering your mouse pointer over a cleansing package in the main Cleansing Package Builder page. The following table describes the possible states for a cleansing package.

State	Description
Ready	<p>Cleansing package is in good condition and available for editing or viewing.</p> <p>Content upgrade is complete, however there may be errors. Hover your mouse pointer over the package to see if there are any content upgrade errors.</p>
Creating	Cleansing package is being created.
Open for Read Write	Cleansing package is open for editing.
Publishing	Cleansing package is in the process of being published. Wait for publishing to complete.
Locked	Cleansing package is locked, either because a user has opened it for editing, it is undergoing a content upgrade, it is being published, or there is an error in the cleansing package. If there is an error, unlock the cleansing package in the CMC.
Analyzing	Cleansing package is undergoing an auto-analysis, which is performed after a custom cleansing package is initially created.

State	Description
Error	<p>Cleansing package has errors. The Error state can occur in the following situations:</p> <ul style="list-style-type: none"> • The cleansing package is possibly corrupt • The cleansing package was previously locked • There is an error as a result of content upgrade <p>A previously locked cleansing package was locked by an Information Steward administrator in an attempt to make the cleansing package usable again. In the first two situations, a data steward must open the cleansing package, check the condition and close the cleansing package. If the cleansing package is returned to a Ready state, it may be used. If the cleansing package is still in an Error state, it is corrupt and should be deleted.</p> <p>If the error is a result of a failed content upgrade, the software did not upgrade the cleansing package.</p>
Canceling analysis	Cleansing package auto analysis process is canceled (by the data steward). It may take some time for the cleansing package to move to the Ready state.
Canceling publish	Cleansing package publishing process is canceled (by the data steward). It may take some time for the cleansing package to move to the Ready state.

You may also notice an icon appearing next to the cleansing package name. The icon indicates the cleansing package status.

10.6.2 Cleansing package statuses

A cleansing package may have one of the following statuses: Ready, Busy, Locked, Error.

The status of a cleansing package appears in the properties sheet that opens when you hover your mouse pointer over the cleansing package name in the *Cleansing Package Tasks* screen. The status is also indicated by the cleansing package icon. The following table shows the status, associated icon and possible user action:

Status	Possible user action
Ready	Open and edit or view the cleansing package.
Busy	Wait for the process (auto-analysis or publishing) to complete or cancel the process.

Status	Possible user action
Locked	<p>When a cleansing package is opened for editing, its status changes to Locked so that no other user may edit it.</p> <div data-bbox="501 439 1396 660" style="border: 1px solid #ccc; background-color: #f9f9f9; padding: 10px;"> <p>⚠ Caution</p> <p>Ensure that you do not have the cleansing package open in a different browser window, or if the browser closed before the cleansing package was closed, wait at least 20 minutes for Cleansing Package Builder to automatically close the cleansing package and restore it to a Ready state.</p> </div> <p>Contact your Information Steward administrator to unlock the cleansing package from the Central Management Console (CMC). Unlocking a cleansing package changes the cleansing package to an Error state and status.</p>
Error	<p>When an error icon appears it is either because the file is corrupt, or that the administrator has unlocked it. See the Locked status for details.</p>

The state of a cleansing package with a Busy status cannot be changed by an Information Steward administrator. You can either wait for the operation to complete or delete the cleansing package.

A cleansing package with a Locked status can be unlocked by an Information Steward administrator. Unlocking a cleansing package changes its status to Error. Before further use, the condition of the cleansing package must be verified by a data steward.

11 Match Review Administration

11.1 Match Review

The Match Review module in SAP Information Steward enables business users or data stewards to review the results of automated matching and to make any necessary corrections.

Identifying and consolidating duplicate data delivers direct cost savings by eliminating redundant data, and improves business process efficiency. Typically, automated data quality processes are deployed to cleanse data, standardize it, and deduplicate matching records. Match Review provides the ability to manually intervene with automated match review when the automated process cannot determine the duplicate records with enough confidence.

[Match Review workflow \[page 265\]](#)

The Match Review workflow involves processes in SAP Data Services and SAP Information Steward, in addition to tasks performed by administrators, data stewards, reviewers, and approvers.

[Match result table \[page 266\]](#)

The match groups, which are the outputs of an automated matching job, must be stored in a match result table in a staging database.

[Match Review job status table \[page 267\]](#)

A match review job status table contains information about the job process and results.

11.1.1 Match Review workflow

The Match Review workflow involves processes in SAP Data Services and SAP Information Steward, in addition to tasks performed by administrators, data stewards, reviewers, and approvers.

The key tasks to incorporate Match Review in the overall data quality process are as follows:

1. A Data Services job loads the results of an automated matching process to a match results table in a staging database. (This job typically involves address cleansing, data cleansing, standardization, and matching.)
2. The Data Services job groups similar records that are possible duplicates into match groups.
3. A configuration manager configures the Match Review process, and an administrator schedules it.
4. Information Steward monitors the staging database at scheduled intervals and triggers the match review process if new records are available for review.
5. Information Steward groups match groups in a match review task. A match review task contains all new match groups with the same job instance ID that need to be reviewed and (if required) approved.
6. Information Steward adds the task to the worklist of the assigned reviewers and approvers.
7. Reviewers review the match groups and unmatch any that are not duplicates. If Match Review is configured to allow reviewers to reassign the master record in match groups, reviewers can reassign the master record if necessary.

8. If Match Review is configured to require reviewers' actions to be approved, the review results are submitted to approvers who approve or reject the actions.
9. A data steward monitors the overall progress of the match review and ensures that sufficient reviewers and approvers are assigned to finish the task on time.
10. Information Steward posts the match review results by updating the match results table. For auditing purposes, Information Steward stores a complete history of all changes to records and match groups.
11. The data quality process picks up the match review results from the staging database and integrates the data into the target application or database systems.

11.1.2 Match result table

The match groups, which are the outputs of an automated matching job, must be stored in a match result table in a staging database.

An extract, transform, and load (ETL) developer creates and configures an SAP Data Services automated matching job. A member of the Data Review Administrator group defines a connection to the staging database so that SAP Information Steward can access it.

The match result table must have the columns described in the following table.

Match results required columns

Type	Column	Description
Required by match review	Source System	Unique identifier of the system from which the source records originate.
	Source Record ID	Primary key of the record in the source system. The combination of SOURCE_SYSTEM and SOURCE_RECORD_ID uniquely identifies the source record.
	Job Run ID	Unique job run ID of each SAP Data Services job that adds records to the match result table. This field is required to group together match groups into one match review task that appears in the reviewer's worklist. Each delta load from the same job run ID is grouped together as one task.
	Match Review Result	Stores the match review result.
Output from Data Services job	Match Group Number	Alphanumeric string that uniquely identifies a match group.

Type	Column	Description
	Match Group Rank	Indicates whether the record is a master record (M) or subordinate record (S).
	Match Score	Number between 0 and 100 indicating the similarity between the subordinate record and master record for the MTC_GROUP_NUMBER. A high number indicates a high similarity. The master record must have a value of NULL or 0.
Source data fields	Source Column 1	Fields from the various source systems.
	Source Column N	The reviewer uses these fields for comparing records in match review.

Note

Your administrator is responsible for purging the match result table.

11.1.3 Match Review job status table

A match review job status table contains information about the job process and results.

A Data Review administrator defines the match review connection to the staging database and defines the users or user groups and their permissions for this connection. The administrator must be assigned to the Data Review Administrator group in the Central Management Console (CMC) under *Users and Groups*.

The administrator also defines the connection to the staging database in the CMC under [Information Steward > Connections](#). By default, any user from the Data Review Administrator group has full permission for the Connections folder.

The administrator also defines a Match Review job status table in the staging area. The table is updated at the end of the SAP Data Services job to indicate that the job is complete and that the results are ready for review.

The Match Review job status table must have specific columns as described in the following table.

Column	Description
Job Run ID	Each Data Services job that adds records to the match result table has a unique job run ID. This field is required to group together match groups into one match review task that appears in the reviewer's worklist. Each delta load from the same job run ID is grouped together as one task.

Column	Description
Job Status	Status of the job represented by the job run ID. When the Data Services job is complete and the match results are loaded to the staging database, Information Steward expects the job status to be updated to the value specified in the match review configuration. The default value is <i>Pending Match Review</i> . This ensures that Information Steward does not create the match review task while the Data Services job is still in progress and while the match results for the same job run are being added to the staging table.
Job Status Value	Shows the unique column representing the Data Services job status.
Indicates Status Change	Select to monitor the progress of the Match Review. For example, when the job status value changes from "In Progress (I)" to "Completed (C)", you could start the next process in your workflow.

Related Information

[Match Review connections \[page 269\]](#)

11.2 Data Review configuration file

Your administrator can update the Data Review configuration file (`data-review.cfg`) for situations such as the following:

- You want different default threshold values for matching and unmatching records automatically.
- A source column name has the same name as `IS_FINISHED` in the Match Review activity table.

Note

Any changes to the `data-review.cfg` file should be made right after installation before any match review tasks have been processed. For the changes to take effect, you must restart the Enterprise Information Management servers.

The configuration file is located in the following directory:

- On Linux, `<BIP_INSTALL_DIR>/java/pjs/services/ICC.DataReviewService/lib/data-review.cfg`
- On Windows, `<BIP_INSTALL_DIR>\SAP BusinessObjects Enterprise XI 4.0\java\pjs\services\ICC.DataReviewService\lib\data-review.cfg`

The following table lists the data review configuration parameters for `data-review.cfg`.

Parameter	Default Value	Description
<code>default.high.threshold</code>	100	Records with a match score above the high threshold are matched automatically. This value should be the same threshold set in the Data Services job that ran the automated matching process and loaded the match result table in the staging database. If it is set too high, you can adjust it here, for example, to filter out records that can be ignored in match review.
<code>default.low.threshold</code>	0	Records with a match score below the low threshold are unmatched automatically. This value should be the same threshold set in the Data Services job that ran the automated matching process and loaded the match result table in the staging database. If it is set too low, you can adjust it here, for example, to filter out records that can be ignored in match review.
<code>br.table.affix</code>	<code>_BR</code>	The string that Information Steward appends to the match review input table name to create the name of the best record lineage table.
<code>br.undefined.primary.string.column.value</code>	<code>_</code>	Value for a non-nullable string column when inserting a best record into the best record lineage table and match review activity table.
<code>br.undefined.primary.numeric.column.value</code>	<code>-1</code>	Value for a non-nullable string column when inserting a best record into the best record lineage table and match review activity table.
<code>mr.activity.table.affix</code>	<code>_ACT</code>	The string that Information Steward appends to the match review input table name to create the name of the match review activity table.
<code>mr.activity.column.prefix</code>	<code>ORIG_</code>	The prefix of the column name that contains the original values in the match review activity table.
<code>is.finished.col.name</code>	<code>IS_FINISHED_</code>	The name of the column in the match review activity table that indicates if the group has finished.

11.3 Match Review connections

Establish a connection to the data review staging database to enable access to the duplicate records.

For Match Review, the duplicate records are the output of an SAP Data Services match transform job. The duplicate records reside in a table in a staging database. Information Steward provides the “For data review” connection purpose for the staging database. Connection parameters vary from one database type to the next.

This section describes the connection parameters to the staging database, which are different for each database type.

Note

For a complete list of supported databases and their versions, see the [Platform Availability Matrix \(PAM\)](#).

[Defining a Match Review connection \[page 270\]](#)

A Match Review connection provides access to the data review staging database, where the duplicate records reside.

[IBM DB2 connection options for Match Review \[page 271\]](#)

Connection options for IBM DB2 *Data Review* staging database.

[Microsoft SQL Server connection options for Match Review \[page 272\]](#)

Connection options for Microsoft SQL Server Data Review staging database.

[Oracle connection options for Match Review \[page 273\]](#)

Connection options for Oracle *Data Review* staging database.

[SAP HANA connection options for Match Review \[page 273\]](#)

Connection options for SAP HANA Data Review staging database.

[SAP ASE connection options for Match Review \[page 274\]](#)

Connection options for SAP ASE (Sybase ASE) Data Review staging database.

11.3.1 Defining a Match Review connection

A Match Review connection provides access to the data review staging database, where the duplicate records reside.

Ensure that you have the proper privileges on the staging database. For information about Match Review user rights, see [User rights in Match Review \[page 86\]](#).

Follow these steps to define a Match Review connection to a staging database.

1. Log in to the Central Management Console (CMC).
2. At the *CMC Home* page, click *Information Steward*.
3. Select the *Connections* node in the Tree panel.
4. Click **► Manage ► New ► Connection ►**.
5. On the *Create Connection* page, enter the following information.

Option	Description
<i>Connection Name</i>	Name for this Match Review connection. <ul style="list-style-type: none">• Maximum length is 64 characters• Can be multi-byte• Case insensitive• Can include underscores and spaces• Cannot include other special characters: ?!@#\$%^&*()-+={}[]:~\ .,`~ You cannot change the name after you save the connection.
<i>Description</i>	Optional. Describes this connection.

6. In the *Connection Type* drop-down list, select the *Database connection* value.
7. In the *Purpose of connection* drop-down list, select *For data review*.

8. In the *Database Type* drop-down list, select the database that contains the match review results.
9. Enter the relevant connection information for the database type.
10. If you want to verify that Information Steward can connect successfully before you save this profile connection, click *Test connection*.
11. Click *Save*.

Note

After you save the connection, you cannot change its name, connection type, purpose and connection parameters which uniquely identify a database.

The newly configured connection appears in the list on the right of the *Information Steward* page.

After you create a connection, select *User Security* from the navigation pane at left to grant permissions to the connection. For the basic steps to grant user permission to the connection, see [Assigning users to specific Data Insight objects \[page 69\]](#).

you must authorize users to it so that they can perform tasks such as create or edit match review configurations, run a match review configuration and force completing a match review task.

11.3.2 IBM DB2 connection options for Match Review

Connection options for IBM DB2 *Data Review* staging database.

DB2 option	Possible values	Description
<i>Server Name</i>	Computer name, fully qualified domain name, or IP address	Name of your DB2 server.
<i>Port Number</i>	Five-digit port number. Default: 50000	Port number for your DB2 server.
<i>Database Name</i>	Refer to the requirements of your database	Name of the staging database to which Match Review connects.
<i>User Name</i>	The value is specific to the database server and language.	User name of the account through which SAP Information Steward accesses the database.
<i>Password</i>	The value is specific to the database server and language.	User password for the database.
<i>Schema</i>	Refer to the requirements of your database.	Name of the database schema.

Related Information

[Defining a Match Review connection \[page 270\]](#)

11.3.3 Microsoft SQL Server connection options for Match Review

Connection options for Microsoft SQL Server Data Review staging database.

Microsoft SQL Server option	Possible values	Description
<i>Server Name</i>	Computer name, fully qualified domain name, or IP address	Name of machine where the SQL Server instance is located.
<i>Port Number</i>	Four-digit port number. Default: 1433	Port number to connect to this Microsoft SQL Server.
<i>Database Name</i>	Refer to the requirements of your database.	Name of the staging database to which Match Review connects.
<i>Enable SSL Encryption</i>	Yes or No	Enable or disable SSL encryption.
<i>Verify Server Certificate</i>	Yes or No	Specifies whether the client certificate should be used to verify the host name.
<i>Host Name</i>		Name used to verify the server's identity.
<i>SSL TrustStore</i>		Absolute location of the SSL TrustStore file.
<i>SSL TrustStore Type</i>	JKS or PKCS12	SSL TrustStore file format.
<i>SSL TrustStore Password</i>		Password to access the private key from the TrustStore file.
<i>Windows Authentication</i>	Yes or No Default: No	Use Windows Authentication to login to Microsoft SQL Server. Set to No to use Microsoft SQL Server Authentication.
<i>User Name</i>	The value is specific to the database server and language.	User name of the account through which Information Steward accesses the database.
<i>Password</i>	The value is specific to the database server and language.	User password for the database.
<i>Schema name</i>	Refer to the requirements of your database.	Name of the database schema.

Related Information

[Defining a Match Review connection \[page 270\]](#)

11.3.4 Oracle connection options for Match Review

Connection options for Oracle *Data Review* staging database.

Oracle option	Possible values	Description
<i>RAC Connection</i>	<i>Yes</i> or <i>No</i> Default: <i>No</i>	Use Oracle RAC to connect to the Match Review staging database.
<i>Server Name</i>	Computer name, fully qualified domain name, or IP address	Name of machine where the Oracle Server instance is located.
<i>Port Number</i>	Four-digit port number. Default: 1521	Port number to connect to this Oracle Server.
<i>SID or Service Name</i>	The value is specific to the database server and language.	Oracle SID or service name for this database
<i>User Name</i>	The value is specific to the database server and language.	User name of the account through which the software accesses the database.
<i>Password</i>	The value is specific to the database server and language.	User password for this database.
<i>Schema name</i>	Refer to the requirements of your database.	Name of the database schema.

Related Information

[Defining a Match Review connection \[page 270\]](#)

11.3.5 SAP HANA connection options for Match Review

Connection options for SAP HANA Data Review staging database.

SAP In-Memory Database option	Possible values	Description
<i>Server Name</i>	Computer name	Name of the computer where the SAP HANA server is located.
<i>Port Number</i>	Five-digit port number. Default: 30015	Port number to connect to this SAP HANA Server.
<i>Enable SSL Authentication</i>	<i>Yes</i> or <i>No</i>	Enables or disables SSL authentication.
<i>Verify Server Certificate</i>	<i>Yes</i> or <i>No</i>	Specifies whether the client certificate should be used to verify the host name.

SAP In-Memory Database option	Possible values	Description
<i>Server Name in Certificate</i>		Specifies the server name used to verify the server's identity.
<i>SSL KeyStore</i>		Specifies the absolute location of the Java keystore file.
<i>SSL KeyStore Type</i>	JKS or PKCS12	Specifies the type of the keystore certificate file.
<i>SSL KeyStore Password</i>		Specifies the password used to access the private key from the keystore file.
<i>SSL TrustStore</i>		Specifies the absolute location of the Java truststore file.
<i>SSL TrustStore Type</i>	JKS or PKCS12	Specifies the Java TrustStore file format.
<i>SSL TrustStore Password</i>		Specifies the password to access the private key from the truststore file.
<i>User name</i>	The value is specific to the database server and language.	User name of the account through which Information Steward accesses the database.
<i>Password</i>	The value is specific to the database server and language.	User password for this database.
<i>Schema Name</i>	Refer to the requirements of your database.	Name of the database schema.

Related Information

[Defining a Match Review connection \[page 270\]](#)

11.3.6 SAP ASE connection options for Match Review

Connection options for SAP ASE (Sybase ASE) Data Review staging database.

SAP ASE option	Possible values	Description
<i>Server Name</i>	Computer name	Name of the computer where the SAP ASE server is located.
<i>Port Number</i>	Five-digit port number. Default: 50000	Port number to connect to this SAP ASE server.
<i>Database Name</i>	Refer to the requirements of your database	Name of the staging database to which Match Review connects.
<i>User name</i>	The value is specific to the database server and language.	User name of the account through which the software accesses the database.

SAP ASE option	Possible values	Description
<i>Password</i>	The value is specific to the database server and language.	User password for this database.
<i>Schema Name</i>	Refer to the requirements of your database.	Name of the database schema.

Related Information

[Defining a Match Review connection \[page 270\]](#)

12 Information Steward Utilities

12.1 Utilities overview

Utilities provide helpful services for maintaining different areas of SAP Information Steward.

Utilities enable you to run maintenance tasks automatically so that your data is continually up-to-date and your systems are running efficiently. For example, the calculate scorecard utility tracks key data domain statistics, and the clean-up failed data utility manages the efficiency of your failed data repository by deleting old information.

Access Information Steward utilities in the Central Management Console (CMC). Set utilities to run on a recurring schedule and make other settings to define what you want the utility to do.

Change a utility default recurrence to any of the following recurrence options:

- Now
- Once
- Hourly
- Daily
- Weekly
- Monthly
- Nth day of month
- 1st Monday of month
- Last day of month
- X day of Nth week of the month
- Calendar

Information Steward utilities and default recurrence setting

Utility	Description	Default recurrence
Calculate Scorecard	Calculates scores of key data domains for selected data quality scorecards.	Daily Eligible for service monitor alerts.

Utility	Description	Default recurrence
Clean up Failed Data	<p>Purges old and obsolete failed data from all of your failed data repository connections on a periodic basis.</p> <p>Purges old failed data based on the history retention setting.</p> <p>Purges obsolete failed data related to tables or views that have been deleted in Data Insight.</p> <p>Set <i>History Retention Period</i> to the number of days to retain failed data repository history.</p>	<p>Daily</p> <p>Default number of days for history retention: 0 days</p> <p>Eligible for service monitor alerts.</p>
	<p>Note</p> <p>The default schedule is 0 (zero) days so that your existing failed data history is not deleted when you upgrade the software. A setting of a negative number (-1) causes the utility to never clean up history. A positive number causes the utility to clean up history every set number of days (for example, set to 5 to run the utility every 5 days).</p>	<p>Note</p> <p>If History Retention Period is set to less than or equal to zero, the software does not send service monitor alerts.</p>
Compute Report Lineage	<p>Computes impact and lineage for only the integrator sources that contain changes since the last time the computation was run.</p> <p>Create another configuration to compute impact and lineage across all integrator sources.</p> <p>Select a <i>Mode</i>:</p> <ul style="list-style-type: none"> • Full: Deletes the relationships before extracting lineage • Optimized: Performs an incremental update 	Now

Utility	Description	Default recurrence
E-mail Notification	<p>For match review, rules, and term approval tasks.</p> <p>Sends a reminder e-mail notification to approvers when a specific task is near or passed the due date.</p>	<p>Daily</p> <p>The system checks daily, but sends one e-mail reminder per due date type (near or passed) per task.</p> <p>Eligible for service monitor alerts.</p>
<div style="background-color: #f0f0f0; padding: 10px; border-left: 2px solid #0070c0;"> <p>Note</p> <p>Users and groups must have e-mail addresses set in their user profile. Administrators must have configured the EIM Adaptive Job Server for e-mails.</p> </div>		
Purge	<p>Increases disk space in the Information Steward repository by deleting:</p> <ul style="list-style-type: none"> • Database logs after integrator sources are deleted. • Profile results, scores, and sample data that exceed configured retention period. • Match review tasks that exceed configured retention period. • Data Cleansing Advisor staged data tables that exceed configured retention period. (Users can restore data cleansing solutions, if necessary, even if staged data has been purged.) • Logically deleted information that includes sample data for profiling and rule tasks. 	<p>Daily</p> <p>Eligible for service monitor alerts.</p>
<div style="background-color: #f0f0f0; padding: 10px; border-left: 2px solid #0070c0;"> <p>Note</p> <p>To delete profile results, scores, and sample data for individual tables or files before retention period expires, see the <i>User Guide</i>.</p> </div>		

Utility	Description	Default recurrence
Update Search Index	<p>Re-creates Metadata Management search indexes. Create additional configurations to re-create search indexes for specific integrator sources.</p> <p>For more information, see Recreate search indexes on Metadata Management [page 280].</p>	Now

[Compute and store lineage information for reporting \[page 279\]](#)

The SAP Information Steward repository provides a lineage staging table named MMT_Alternate_Relationship that consolidates end-to-end impact and lineage information across all integrator sources.

[Recreate search indexes on Metadata Management \[page 280\]](#)

When you run a metadata integrator source, Metadata Management updates the search index with any changed metadata.

Related Information

[Setting up e-mail notification configuration \[page 100\]](#)

[Setting user e-mail address for notifications \[page 57\]](#)

12.1.1 Compute and store lineage information for reporting

The SAP Information Steward repository provides a lineage staging table named MMT_Alternate_Relationship that consolidates end-to-end impact and lineage information across all integrator sources.

The Metadata Management module provides predefined Crystal Reports from the MMT_Alternate_Relationship table. You can also create your own reports from this table.

To view the Crystal Reports, follow these steps:

1. Open [Metadata Management](#) tab.
2. Click [Open](#) and select [Reports](#) from the dropdown list.

For more information about defining custom reports, see the *User Guide*.

Before generating reports that rely on the MMT_Alternate_Relationship lineage staging table, update the lineage information in the lineage staging table. Then, you can either schedule the Compute Report Lineage utility or run the utility on demand to ensure that the reports contain the latest lineage information.

We recommend that you run the Compute Report Lineage utility after any of the following processes:

- Run an Integrator to collect metadata from a source system.
- Change preferences for relationships between objects.

Note

The data in the lineage staging table uses the values in *Impact and Lineage Preferences* and *Object Equivalency Rules* to determine impact and lineage relationships across different integrator sources.

- Establish or modify a user-defined relationship of type Impact or Same As.

Parent topic: [Utilities overview \[page 276\]](#)

Related Information

[Recreate search indexes on Metadata Management \[page 280\]](#)

12.1.2 Recreate search indexes on Metadata Management

When you run a metadata integrator source, Metadata Management updates the search index with any changed metadata.

The search feature of the Metadata Management module of SAP Information Steward allows you to search for an object that might exist in any metadata integrator source.

You might need to recreate the search indexes in situations such as the following:

- The Search Server was disabled and could not create the index while running a metadata integrator source.
- The search index is corrupted.

Parent topic: [Utilities overview \[page 276\]](#)

Related Information

[Compute and store lineage information for reporting \[page 279\]](#)

12.2 Scheduling a utility

Set a recurring schedule for Information Steward utilities in the Central Management Console (CMC).

1. Log in to the Central Management Console (CMC) with a user name that belongs to the Metadata Management Administrator group or the Administrator group.
2. Select *Applications* from the navigation dropdown list at the top.

3. Double-click *Information Steward Application* in the *Application Name* column.
4. Select *Manage Utilities* from the Navigation pane at left.

A list of utilities appears at right with additional information such as last run and date created.

5. Select a utility to schedule from the list of utilities.
6. In the top menu toolbar, click ► *Actions* ► *Schedule* ▾.
7. Expand *Schedule* from the navigation pane at left and select *Recurrence*.
8. In the Recurrence pane, select a recurrence frequency from the *Run object* dropdown list.

The recurrence frequency sets the interval of time in which to run the utility. For example, select to run the utility each hour, once a week, and so on.

Optionally, set the *Number of retries* to a value other than the default of 0, and change the *Retry interval in seconds* from the default value of 1800.

Each recurrence frequency that you choose from the *Run object* dropdown list has additional options that appear in the view after you select the option.

9. Complete the additional options that appear based on the *Run object* value.
10. Click *Schedule*.

The selected utility runs based on the schedule that you just set.

12.3 Rescheduling a utility

Change a utility schedule from the default schedule by rescheduling the utility in the Central Management Console (CMC).

The following are examples of rescheduling a utility:

- You want to obtain rule results for the data quality trend graphs in Data Insight on a monthly basis. Reschedule the Calculate Scorecard utility from the default setting of Daily to Monthly.
- You want to ensure that the Information Steward repository does not get too full, but the default schedule is too frequent. Reschedule the Purge utility from the default setting of Daily to Weekly.

Follow these steps to reschedule a utility:

1. Log in to the CMC with a user name that belongs to the Metadata Management Administrator group or the Administrator group.
2. Select *Applications* from the navigation dropdown list.
3. Double-click *Information Steward Application* in the *Applications Name* list.
4. Click ► *Actions* ► *Manage Utilities* ▾ from the toolbar.
5. Select the utility name from the list of utilities at right.
6. Click ► *Actions* ► *History* ▾ from the toolbar.

The Utility History opens showing all utilities that have run, the status, start, and end time, and so on.

7. Select the utility from the list that you want to reschedule.

The *Schedule Status* should be showing as Recurring.

8. Click [Reschedule](#) at the top.
9. Select [Recurrence](#) from the Navigation pane at left.
10. Select a new frequency from the [Run object](#) dropdown list.
11. Set additional relevant values for the option that you chose.
12. Optional. Select [Instance Title](#) from the Navigation pane at left and enter a different name for the schedule.
13. Click [Schedule](#) at the bottom.
The newly created [Recurring](#) schedule appears in the list in Utility History.
14. Delete the original recurring schedule.

12.4 Running a utility on demand

Select to run a utility immediately from the list of utilities in the Central Management Console (CMC).

You must be an administrator to perform these steps.

1. Log in to the CMC.
2. Select [Applications](#) from the navigation dropdown list at the top.
3. Select [Information Steward Application](#) in the [Applications Name](#) list.
4. Click [Action](#) [Manage Utilities](#) in the menu.
The list of configured utilities displays with the date and time each was last run.
5. Select the name of the applicable utility from the utility list.
6. Select [Action](#) [Run Now](#) from the menu.

Caution

Do not start the Compute Report Lineage utility if an instance is still running. If you start another utility while the utility is currently processing, you may cause the CMC to stall or you may experience a delay. A stall or delay may also happen if you stop the utility and start another instance right away. The stopped utility may take a few minutes to stop processing in the repository after you have stopped it.

7. Refresh the list of utilities in the [Utility Configurations](#) list by clicking the [Refresh](#) icon in the upper right.

12.5 Monitoring utility executions

Monitor utility executions to view the status or the progress of a utility run.

1. Login to the Central Management Console (CMC) with a user name that belongs to one or more of the following administration groups:
 - Data Insight Administrator
 - Metadata Management Administrator
 - Administrator
2. At the top of the CMC Home screen, select [Applications](#) from the navigation list.

3. Select *Information Steward Application* in the *Applications Name* list.
4. Click **▶ Action ▶ Manage Utilities ▶** in the top menu tool bar.
The list of configured utilities displays with the date and time each was last run.
5. Click the *Refresh* icon to update the *Last Run* column on the *Utility Configurations* dialog.
6. To view the status of the utility run, select the utility configuration name and click **▶ Action ▶ History ▶**.
The *Schedule Status* column can contain the following values:

Schedule Status	Description
Failed	The utility did not complete successfully.
Pending	The utility is scheduled to run one time. When it actually runs, there will be another instance with status "Running."
Recurring	The utility is scheduled to recur. When it actually runs, there will be another instance with status "Running."
Running	The utility is currently executing.
Success	The utility completed successfully.

7. To see the progress of a utility instance:
 - a. Select the instance name and click the icon for *View the database log* in the top menu bar of the *Utility History* dialog box.
The *Database Log* dialog box shows the utility messages.
 - b. To find specific messages in the *Database Log* dialog box, enter a string in the text box and click *Filter*.
For example, you might enter **error** to see if there are any errors.
 - c. To close the *Database Log* dialog box, click the *X* in the upper right corner.
8. To save a copy of a utility log:
 - a. Scroll to the right of the *Utility History* dialog box, and click the *Download* link in the *Log File* column in the row of the utility instance you want.
 - b. Click *Save*.
 - c. On the *Save As* dialog box, browse to the directory where you want to save the log and optionally change the default file name.
9. To close the *Utility History* dialog box, click the *X* in the upper right corner.

12.6 Utility service notification alerts

If a critical scheduled utility does not run, does not complete, or runs and fails, the software sends a service notification alert to designated e-mail recipients.

Critical utilities include the following:

- Calculate scorecard
- Clean up failed data
- E-mail notification
- Purge

The service notification alerts can prevent issues stemming from the utility not performing as you expect it to.

❖ Example

- Outdated quality trend information. When the calculate scorecard utility has not run in the last 24 hours, the quality trend information reflects outdated data. You could base a data quality decision on outdated quality trend data.
- Repository runs out of disk space. When the cleanup failed data utility doesn't run as scheduled, the repository can contain old or obsolete failed data information. Your current failed data information may not be stored because the failed data repository is full.

[Service monitor e-mail alert schedule \[page 284\]](#)

The system sends a service monitor e-mail alert based on the utility recurrence settings.

[Activating e-mail alerts for utility service monitor \[page 285\]](#)

Activate e-mail alerts for utilities in the Central Management Console (CMC).

12.6.1 Service monitor e-mail alert schedule

The system sends a service monitor e-mail alert based on the utility recurrence settings.

The alert goes out for the utilities that run in the Application Service of the EIM Adaptive Processing Server. The service monitor checks the utility running status daily. The service monitor checks the recurrence configuration and then checks the status of the last run. The software sends out the following alerts as applicable:

- Last instance run failed on `<date_time>`
- There is no completed instance run in the past `<number_of_days>` since `<date_scheduled>` as of `<date_of_alert>`
- Last instance is still running since `<date_scheduled>` as of `<date_of_alert>`

The following table contains the time schedule for when the service monitor reports problems based on the recurrence setting.

Recurrence setting	Alert
Now	Service monitor sends alert for problems within the last 24 hours.
Once	
Hourly	
Daily	
No recurrence configured	
Weekly	Service monitor sends alert for problems within the last 7 days.
Monthly	Service monitor sends alert for problems within the past 30 days.
Nth day of month	
1st Monday of month	
Last day of month	
X day of Nth weeks of month	

Recurrence setting

Alert

Calendar

Service monitor sends alert for problems within the past 24 hours.

❁ Example

Service monitor e-mail alert example

The software sends the following e-mail when the service monitor detects that the calculate scorecard utility has not run in the last 24 hours.

Action Required: Utility Failed/Not Run

The utilities listed below require your attention:

Calculate Scorecard: There is no completed instance run in the past 24 hours since Mar 1, 2017 8:42:59 PM PST on Mar 2, 2017 8:42:59 PM PST.

📌 Note

The software displays a timestamp that is localized based on the CMC (Central Management Console) preferences settings for *Product Locale*, *Preferred Viewing Locale*, and *Time Zone*. Change these settings in the CMC by selecting **Preferences > CMC Preferences**. You must have the proper permission to change CMC Preferences settings.

12.6.2 Activating e-mail alerts for utility service monitor

Activate e-mail alerts for utilities in the Central Management Console (CMC).

Before you complete the following steps, check that you meet the following prerequisites:

- You have the required permissions to perform the task.
- The recipient users or user groups have their e-mail addresses in the users and groups configuration. For information, see [Setting user e-mail address for notifications \[page 57\]](#)
- Your administrator has configured the EIM Adaptive Job Server destination e-mail with the SMTP domain, server address, and port information. For information, see [Setting up e-mail notification configuration \[page 100\]](#).

1. Log in to the Central Management Console (CMC).

The CMC Home page opens.

2. Select *Applications* from the drop-down list at the top.
3. Double-click *Information Steward Application* from the Organize column.
4. Select *Configure Application* from the navigation pane at left.
5. Scroll to the *Email Settings* group in the right pane.
6. Enter a user or group e-mail address in the *To Email Address* option.

The recipient address that you enter here is applicable only for the utility service monitor e-mail alerts.

7. Click [Save](#).

12.7 Modifying utility configurations

SAP Information Steward provides a default configuration for each utility, some of which you can modify.

Modify the configurations for the following utilities:

- Compute Lineage Report utility: Change the default Mode from *Optimized* to *Full*. Also change the description.
- Update Search Index utility: Change the default Integrator Source from *All Sources* to *Ingegrator Source*. Also change the description.

Note

The Calculate Scorecard, Purge, and Clean up Failed Data utilities do not have configuration parameters.

1. Log in to the Central Management Console (CMC) with a user name that belongs to the Metadata Management Administrator group or the Administrator group.
2. Select [Applications](#) from the navigation list at the top of CMC Home.
3. Double-click [Information Steward Applications](#) in the [Application Name](#) list.
4. Select [Manage Utilities](#) from the navigation list at left.
5. Select the applicable utility.
6. Click [Actions](#) > [Properties](#) in the top menu tool bar.
The [Utilities Configurations](#) window opens.
7. Select the utility to modify and make your changes.
8. Click [Save](#).

12.8 Additional utility configurations

Create an additional utility configuration for Compute Report Lineage and Update Search Index and save it in addition to the default configuration.

The following table lists the default configuration and an example additional configuration for each applicable utility.

Utility	Default configuration	Example for additional configuration
Compute Report Lineage	<i>Mode</i> = Optimized. Recalculates lineage information in the lineage staging table for only integrator sources that have changed since the utility was last run.	Configure an additional Compute Lineage Report utility with <i>Mode</i> set to <i>Full</i> . Full mode recalculates lineage information across all integrator sources.
Update Search Index	<i>Integrator Source</i> = All Sources. Rebuilds search indexes in Metadata Management for all integrator sources.	Configure an additional Update Search Index utility with <i>Integrator Source</i> set to only one of the integrator sources. Rebuilds indexes for the metadata collected for only that integrator source.

Note

You cannot create an additional utility configuration for the following utilities:

- Calculate Scorecard
- Clean up Failed Data
- Email Notification
- Purge

[Creating an additional utility configuration \[page 287\]](#)

Create a new utility configuration for either Compute Report Lineage or Update Search Index and save it in addition to the default configuration.

12.8.1 Creating an additional utility configuration

Create a new utility configuration for either Compute Report Lineage or Update Search Index and save it in addition to the default configuration.

1. Log in to the Central Management Console (CMC) as an administrator.
2. Select *Applications* from the navigation dropdown list at the top of CMC Home.
3. Double-click *Information Steward Application* to open the *Information Steward Settings* view.
4. Select *Manage Utilities* from the navigation pane at left.
5. Select **Manage > New Utility Configuration** in the top toolbar.
6. Select the applicable utility from the *Utility Type* dropdown list.

Options are Report Lineage Utility or Search Index Utility.

7. Required. Type a name for the utility in *Name*.
8. Optional. Type a description for the utility in *Description*.
9. For Compute Report Lineage utility. Required. Change the default setting for *Mode* from **Full** to **Optimized**.

Mode	Description
Optimized	Performs an incremental update across all integrator sources.
Full	Deletes the relationships and recalculates the entire impact and lineage information across all integrator sources.

10. For Update Search Index utility. Optional. Change the default setting for *Integrator Source* from **All Sources** to an existing source that is listed in the dropdown list.

- **All Sources**: Rebuilds indexes for the metadata collected for all of your sources.
- `<source_name>`: rebuilds indexes for the metadata collected for only the selected integrator source.

11. Click *Save*.

After you select **Manage** **Refresh**, the new utility appears in the *Utility Configurations* list of utilities in alphabetical order.

12.9 Deleting default schedule for Calculate Scorecard and Purge utilities

If you create a new schedule for *Calculate Scorecard* or *Purge* utilities and you want to override the default recurring schedule, delete the old recurring instance.

1. Log in to the Central Management Console (CMC) with a user name that belongs to the Metadata Management Administrator group or the Administrator group.
2. Select *Applications* from the navigation dropdown list.
3. Double-click *Information Steward Application* in the *Application Name* column.
4. Select *Manage Utilities* from the Navigation pane at right.

A list of utilities appears with additional information such as last run and date created.

5. Select the applicable utility, either *Calculate Scorecard* or *Purge*, from the utility list.
6. Click **Actions** **History**.
7. Select the recurring schedule to delete from the list under *Schedule Name*.
8. Click the *X* default icon in the icon toolbar to delete the default schedule.

13 Performance and Scalability Considerations

13.1 Resource intensive functions

The main Information Steward functions that influence performance and sizing are as follows.

Data profiling

Information Steward can perform basic and advanced data profiling operations to collect information about data attributes like minimum and maximum values, pattern distribution, data dependency, uniqueness, address profiling, and so on. These operations require intense, complex computations and are affected by the amount of data that is processed.

Validation rule processing

Information Steward validates source data against business rules to monitor data quality and generate scores. Validation is a computation-heavy process that affects performance based on the number of records and the number of rules.

The complexity of rules also affects performance. For example, if you have lookup functions in rule processing, the processing takes more time and disk space. Typically, lookup tables are small, but if they are big, the tables can adversely affect performance.

Cleansing Package Builder

After a custom cleansing package is initially created, the Auto-analysis service for Cleansing Package Builder is processed. Auto-analysis service analyzes the sample data that was provided during the steps in the wizard to create a custom cleansing package. Some of the analyzing that is performed is grouping similar records, which generates parsing rules based on analysis of combinations of parsed values found in the data. With a large number of records and parsed values, this becomes very computation and memory intensive.

Metadata integrators

Metadata integrators collect metadata about different objects in the source systems and store it in the Information Steward repository. If a large amount of metadata is being collected, it can affect performance and the size of the repository.

Metadata browsing and view data

You can browse metadata for tables and columns and also view source data. If the table has a large number of columns, the [View Data](#) operation is affected because a lot of data must be accessed, transferred, and displayed in the user interface.

Lineage and impact analysis

After the metadata is collected, Information Steward can do end-to-end lineage and impact calculations. Typically, these are periodically run as utilities to capture the latest information. If the amount of metadata is large, this calculation can affect performance.

Data Insight joins

Processing tables in join conditions can slow processing speed based on the table size. Increase processing speed by ranking (ordering) the tables in a join and setting caching options.

Related Information

[Factors that influence performance and sizing \[page 292\]](#)

[Scalability and performance considerations \[page 296\]](#)

[Best practices for performance and scalability \[page 315\]](#)

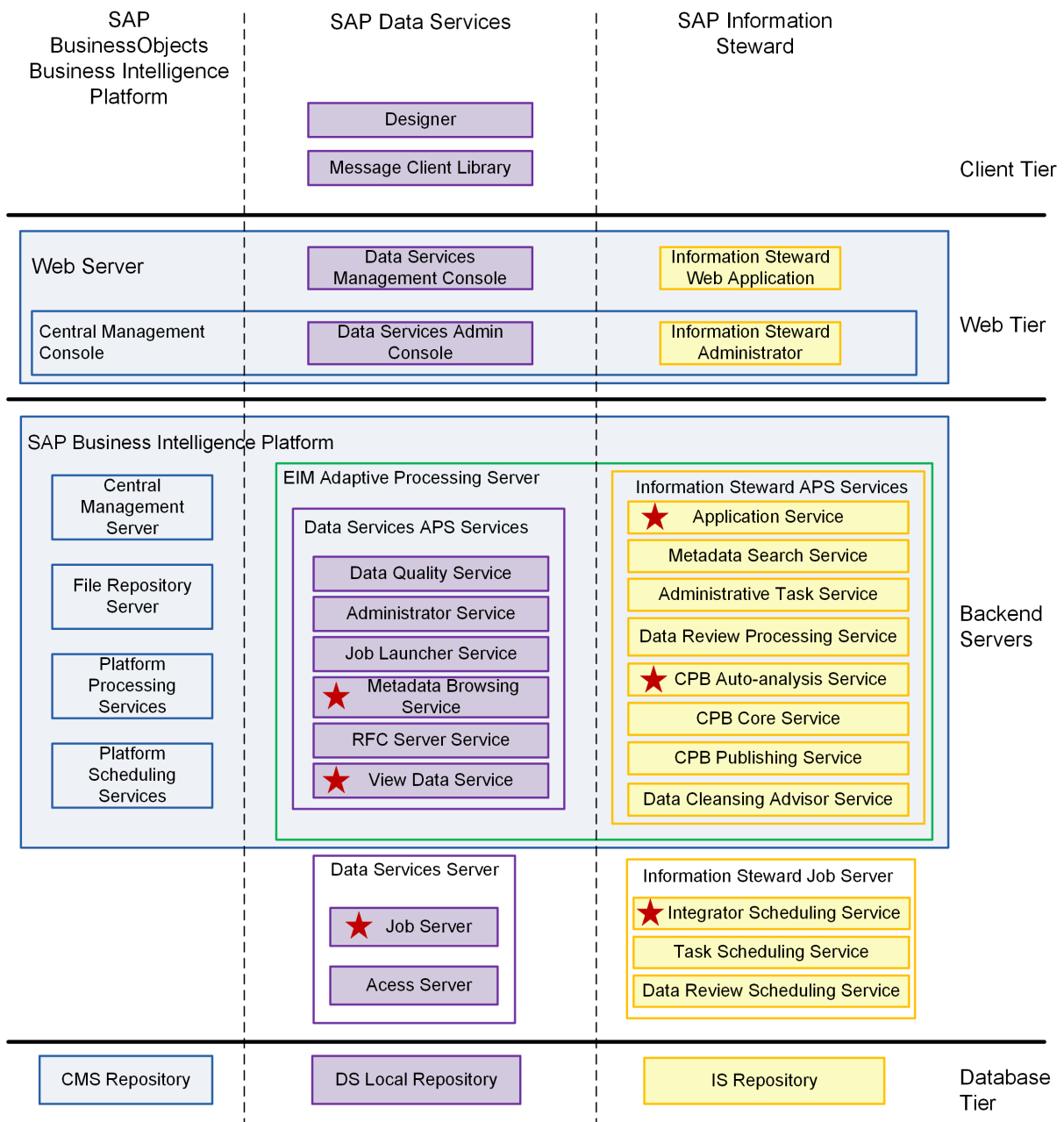
13.2 Architecture

The following diagram shows the architectural components for SAP Business Intelligence platform, SAP Data Services, and SAP Information Steward. The resource-intensive servers and services are indicated with a red asterisk:

- Data Services Job Server processes data profiling and rule validation tasks in Data Insight.

- Cleansing Package Builder Auto-analysis Service analyzes sample data and generates parsing rules for Cleansing Package Builder.
- Information Steward Integrator Scheduling Service processes Metadata Integrators.
- Application Service performs lineage and impact analysis for Metadata Management and provides the Information Steward web application access to the Information Steward repository.
- Data Services Metadata Browsing Service obtains metadata from Data Insight connections.
- Data Services View Data Service obtains the source data from Data Insight connections.
- Web Application Server handles requests from users on the web applications for Information Steward.

For more details about these servers and services, see [Services \[page 12\]](#).



13.3 Factors that influence performance and sizing

Use Information Steward to work on large amounts of data and metadata. The following factors affect the performance and sizing. These factors affect the required processing power (CPUs), RAM, and hard disk space (for temporary files during processing) and the size of the Information Steward repository.

Related Information

[Data characteristics and profiling type \[page 293\]](#)

[Amount of data \[page 292\]](#)

[Data characteristics and profiling type \[page 293\]](#)

[Number of metadata sources \[page 295\]](#)

[Number of concurrent users \[page 293\]](#)

13.3.1 Data Insight

Learn about sizing data, concurrent users and profiling types.

13.3.1.1 Amount of data

The amount of the data is calculated using the number of records and the number of columns. In general, the more data that is processed, the more time and resources it requires. This is true for profiling and rules validation operations. The data may come from one or more sources, multiple tables within a source, views, or files.

This factor affects the required CPU, RAM, and hard disk space and the Information Steward repository size. The larger the record size, the more resources are required for efficient processing.

If the data being processed has many columns, it requires more time and resources. If you are doing column profiling on many columns or if the columns being processed have long textual data, it affects performance. In short, if the record length is more, more resources are required.

Related Information

[Factors that influence performance and sizing \[page 292\]](#)

13.3.1.2 Data characteristics and profiling type

For distribution profiling such as Value, Pattern, or Word distribution, if the data has many distinct values, patterns, or words, it requires more resources and time to process.

This factor affects the required CPUs, RAM, and Information Steward repository size. The more distinct the data, the more resources required.

Related Information

[Factors that influence performance and sizing \[page 292\]](#)

13.3.1.3 Number of concurrent users

Information Steward is a multi-user web-based application. As the number of concurrent users increases, nearly all aspects of the application are affected. More users may mean more profiling tasks, more scorecard views and rule execution, and so on. The key word is concurrent.

If all users run tasks concurrently, it affects the required CPUs, RAM, and hard disk, and the Information Steward repository size. If most of the users are just viewing the scorecard, then the performance depends more on the web application server where the Information Steward repository is created.

Related Information

[Factors that influence performance and sizing \[page 292\]](#)

13.3.1.4 SAP tables and performance

Improve performance when processing views that contain SAP tables by following these best practices.

Performance is sometimes affected when you process views that contain SAP tables. SAP tables are usually large and joining two large SAP tables in a view can further affect performance. Follow these best practices for better performance when processing views that contain SAP tables:

- When possible, only join SAP tables from the same source.
- Use only SAP ABAP-supported functions in mapping, filters, or join condition expressions.
- Reduce the amount of data the software has to analyze by utilizing all of the tools designed to create views that are specific to your needs.
- If your view includes one or more large SAP tables, specify the *ABAP* data transfer method in the view definition and ensure that the SAP application connection parameter *ABAP execution option* is set to *Generate and Execute*.

- Use the *Set optimal order* option in the *Join Conditions* tab, and set the larger source tables to the first positions.

Related Information

[SAP tables and ABAP-supported functions \[page 294\]](#)

[Execute options for ABAP program generation for SAP tables \[page 155\]](#)

13.3.1.5 SAP tables and ABAP-supported functions

For better performance, use ABAP-supported functions in your views that contain SAP tables.

There are functions that are supported for use with SAP tables and ABAP programs. You can use any available functions when forming your expressions, but non-supported ABAP functions are processed differently than the ABAP-supported functions, so they will likely take longer to process.

ABAP-supported functions:

- Absolute (abs)
- Decode (decode)
- Left trim (ltrim)
- Left trim ext (ltrim_blanks_ext)
- Replace null (nvl)
- Right trim (rtrim)
- Right trim ext (rtrim_blanks_ext)
- String length (length)
- Substring (substr)
- System date (sysdate)
- System time (systemtime)
- To lower case (lower)
- To upper case (upper)

Related Information

[SAP tables and performance \[page 293\]](#)

13.3.2 Metadata Management

Learn about sizing the number of metadata sources and concurrent users

13.3.2.1 Number of metadata sources

Metadata integrators collect metadata from various sources. As the number of metadata sources increases, it takes more resources and time. This factor affects the required CPUs, RAM, and Information Steward repository size.

Related Information

[Factors that influence performance and sizing \[page 292\]](#)

13.3.2.2 Number of concurrent users

If multiple users view impact and lineage information concurrently, the response time is affected.

Related Information

[Factors that influence performance and sizing \[page 292\]](#)

13.3.3 Cleansing Package Builder

Learn about sizing sample data, number of concurrent users, and types of cleansing packages.

13.3.3.1 Amount of sample data

For Cleansing Package Builder, the amount of sample data used to create the cleansing package affects the performance of the Auto-analysis service. This in turn can affect response time for the user interface.

Related Information

[Factors that influence performance and sizing \[page 292\]](#)

13.3.3.2 Data characteristics

When you create custom cleansing packages, if there are many parsed values per row, it requires more resources and time to analyze and create parsing rules. For Cleansing Package Builder, the larger the data set, the more CPU processing power is required. The more parsed values per row, the more RAM that is required.

Related Information

[Factors that influence performance and sizing \[page 292\]](#)

13.3.3.3 Number of concurrent users

If there many users create custom cleansing packages concurrently, the more CPUs and RAM required.

Related Information

[Factors that influence performance and sizing \[page 292\]](#)

13.4 Scalability and performance considerations

Information Steward uses SAP Business Intelligence platform and Data Services platform for most of the heavy computational work. It inherits the service-oriented architecture provided by these platforms to support a reliable, flexible, highly available, and high performance environment.

Here are some features and recommendations for using the platforms for performance. These are not mutually exclusive, nor are they sufficient by themselves in all cases. You should use a combination of the following ways to improve throughput, reliability, and availability of the deployment.

Related Information

[Resource intensive functions \[page 289\]](#)

[Factors that influence performance and sizing \[page 292\]](#)

[Best practices for performance and scalability \[page 315\]](#)

[Information Steward web application \[page 300\]](#)

[Scheduling tasks \[page 300\]](#)

[Queuing tasks \[page 301\]](#)

[Degree of parallelism \[page 301\]](#)
[Grid computing \[page 304\]](#)
[Multi-threaded file read \[page 307\]](#)
[Data Insight result set optimization \[page 308\]](#)
[Performance settings for input data \[page 308\]](#)
[Settings to control repository size \[page 309\]](#)
[Settings for Metadata Management \[page 313\]](#)
[Settings for Cleansing Package Builder \[page 312\]](#)

13.4.1 Scalability levels in deployment

A typical Information Steward environment can be scaled at the following levels:

1. Web tier level: Deploy multiple instances of Information Steward web application for high availability and large numbers of concurrent users.
2. Business Intelligence platform and Information Steward services level: Deploy multiple Business Intelligence platform services and/or Information Steward services in a distributed environment for load balancing and scalability. For example, you can deploy multiple metadata integrators on different servers for load balancing. You can also have multiple Information Steward task servers on different servers for high availability.
3. Data Services Job Server level: A Data Services Job Server group used for a given Information Steward deployment can have one or more Data Services Job Servers added to it. As the need for Data Services processing increases (for Data Insight operations), you can scale up by adding more Data Services Job Servers to the Job Servers group.

Related Information

[Scalability and performance considerations \[page 296\]](#)
[Information Steward web application \[page 300\]](#)

13.4.2 Distributed processing

Business Intelligence platform provides a distributed scalable architecture. This means that services that are needed for a specific functionality can be distributed across machines in the given landscape. As long as the services are in the same Business Intelligence platform environment, it doesn't matter which machine they are on; they just need to be in the same CMS cluster. The Information Steward web application and the Information Steward repositories can be on different machines.

Information Steward uses some Business Intelligence platform services, and also has its own services that can be distributed across machines for better throughput. The general principle is that if one of the services needs many resources, then it should be on a separate machine. Similarly, if you add capacity to existing hardware, it can be used for more than one service.

The Data Insight module of Information Steward uses the Data Services Job Server, which supports distributed processing.

This section offers some recommendations on different combinations of Information Steward services that can be combined or decoupled.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Data Insight related services \[page 298\]](#)

[Metadata Management related services \[page 298\]](#)

[Cleansing Package Builder related services \[page 299\]](#)

[Information Steward repository \[page 299\]](#)

[Information Steward web application \[page 300\]](#)

[Grid computing \[page 304\]](#)

13.4.2.1 Data Insight related services

The most important part of the processing for Data Insight is done by the Data Services Job Server. You can install Data Services Job Server on multiple machines and make them part of the single job server group that is used for Information Steward. The profiling and rules tasks are distributed by the Information Steward Job Server to the Data Services Job Server group. The actual tasks are executed by a specific Data Services Job Server based on the resource availability on that server. So, if one server is busy, the task can be processed by another server. In this way multiple profiling and rule tasks can be executed simultaneously.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Information Steward web application \[page 300\]](#)

13.4.2.2 Metadata Management related services

There are two important processes that can affect performance:

- Application Service
- Metadata integrators

Generally, these two processes should be run on separate servers. If they are on the same server, they should run at different times. If there are many metadata integrators and they collect a lot of metadata from different sources, each one of them could be deployed on a separate server.

→ Tip

Schedule intensive processes, such as metadata integrators, during non-commercial hours.

For information about how to deploy integrators on separate servers, see “Scenario 3: Scaling for Metadata Management” in the *Master Guide*.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Information Steward web application \[page 300\]](#)

13.4.2.3 Cleansing Package Builder related services

The Auto-analysis service is a resource intensive service in Cleansing Package Builder. The CPU and RAM requirements depend on the number of parsed values found in the data. Also, if multiple concurrent users create large cleansing packages, it is recommended that you dedicate one server to the Auto-analysis service and allocate enough memory to the Java process.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Information Steward web application \[page 300\]](#)

[Common runtime parameters for Information Steward \[page 187\]](#)

13.4.2.4 Information Steward repository

The Information Steward repository stores all of the metadata collected, profiling and rule results, and sample data. The repository should be on a separate database server. To avoid resource contention, the database server should not be the same database server that contains the source data. The database server that does not contain the source data may or may not be on the same database server that hosts the Business Intelligence platform repository.

The Information Steward repository should be on the same subnetwork as the Data Services Job Server that processes large amounts of data and the metadata integrator that processes the largest amount of metadata.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Information Steward web application \[page 300\]](#)

13.4.2.5 Information Steward web application

Typically all web applications are installed on a separate server with other web applications. No other services or repositories are typically installed with the web application, so that response time for Information Steward user interface users is not affected by services that process data.

Related Information

[Scalability and performance considerations \[page 296\]](#)

13.4.3 Scheduling tasks

Information Steward runs many tasks that are resource intensive. Business Intelligence platform provides the ability to schedule them. This ability can be used to distribute the tasks so that the same resources can be utilized for multiple purposes. The following can be scheduled:

- Profiling tasks
- Rule tasks
- Match review tasks
- Metadata integrators
- Calculate Scorecard utility
- Compute Report Lineage utility
- Purge utility
- Clean up Failed Data utility
- Update search index utility

If you schedule these tasks so that they run at different times and when few users access the system, you can achieve good performance with limited resources. This time slicing is highly recommended for profiling, rules tasks, and metadata integrators. For example, if you have users that process profiling tasks on demand during business hours, then the metadata integrators and rules task should be scheduled during non-business hours.

If there are large profiling jobs, they should be scheduled during non-business hours and ideally on a dedicated powerful server.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Scheduling a task in the CMC \[page 164\]](#)

[Scheduling a utility \[page 280\]](#)

13.4.4 Queuing tasks

Information Steward can queue when there are many Data Insight tasks that are requested to run at the same time. This depends on the user configuration for the *Average Concurrent Tasks* option. Based on this setting and the number of Data Services Job Servers in the group, Information Steward calculates the total number of tasks allowed to run simultaneously in a given landscape. Only that many tasks are sent to the Data Services Job Server group for processing. The remaining tasks are queued. As soon as one of the running tasks finishes, the next task in the queue is processed.

Using this setting, you can control how many Data Insight -related processes are running so that the resources can be utilized and scheduled for other processes running on the system.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Information Steward configuration settings \[page 44\]](#)

13.4.5 Degree of parallelism

For Data Insight functionality, Information Steward uses the Data Services engine. The Data Services engine supports parallel processing in multiple ways, one of which is Degree of Parallelism (DOP). The basic idea is to split a single Data Services job into multiple processing units and utilize available processing power (CPUs) on a server to work on those processing units in parallel. The distribution of work is different for profiling versus rule processing.

Note

DOP is only used for Data Insight functionality for column profiling and rule processing. Metadata Management and Cleansing Package Builder do not use DOP.

Note

In general, do not set the DOP more than the number of available CPUs. To fine tune the performance, set the DOP value based on the number of concurrent tasks and available hard disk and RAM resources. Gradually increase the value of DOP to reach an optimal setting. For more information, see the *SAP Data Services Performance Optimization Guide*.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Column profiling \[page 302\]](#)

[Rule processing \[page 303\]](#)

[Hard disk requirements \[page 303\]](#)

[When to use degree of parallelism \[page 303\]](#)

13.4.5.1 Column profiling

For column profiling, the task is distributed proportionately for the different number of columns. The formula is: $\text{Number of execution units} = \text{Number of Columns} / \text{DOP}$.

For example, if there is a column profiling task for 100 million rows with 20 columns, with DOP = 4, the task will be broken down in execution units that process 5 columns each for all 100 million rows. The data is "partitioned" for 5 columns each for each execution unit to process.

Using multiple CPUs simultaneously causes the throughput to increase proportionately.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Degree of parallelism \[page 301\]](#)

13.4.5.2 Advanced profiling

Advanced profiling tasks (dependency, redundancy, and uniqueness) require complex sorting operations, so it is important to optimize the degree of parallelism settings. The degree of parallelism setting is used to execute sorting operations in parallel sorting operations to increase throughput.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Degree of parallelism \[page 301\]](#)

13.4.5.3 Rule processing

For rule processing, the number of execution units is proportional to the number of rules (as opposed to the number of columns for profiling). The formula is: Number of execution units = Number of rules / DOP.

For example, if there is a rule execution task for 100 million rows with 20 rules, with DOP = 4, the task will be broken down in execution units that process 5 rules each for all 100 million rows.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Degree of parallelism \[page 301\]](#)

13.4.5.4 Hard disk requirements

Adding CPUs and increasing DOP does not guarantee improved throughput. When many processes run in parallel, they also share other hardware resources such as RAM and disk space.

When the Data Services engine processes data, it creates temporary work files in the Pageable Cache Directory. Naturally, if many processes are running simultaneously, all of them create temporary files in the same location. Because this directory is accessed by all of the processes simultaneously, there is a potential for disk contention. In most environments, depending on the disk capacity and speed, you will reach a point where increasing DOP will not improve performance proportionately.

Therefore, it is important to enhance all aspects of the hardware at the same time: the number of CPUs, hard disk capacity, speed, and RAM. You should have a very efficient disk access to go along with the increased number of DOP. Make sure the pageable cache directory is set accordingly.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Degree of parallelism \[page 301\]](#)

[Grid computing \[page 304\]](#)

13.4.5.5 When to use degree of parallelism

Higher DOP should be employed in the following conditions:

- When you have only a few very powerful machines with a lot of processing power, RAM, and fast disk access.

- When you have a large amount of data for profiling or rule tasks.

ⓘ Note

If you run many profile tasks simultaneously with DOP > 1, each of them could be split in multiple execution units. For example, 4 tasks with a DOP of 4 could result in 16 execution units. Now there are 16 processes competing for resources (CPU, RAM, and hard disk) on the same machine. So it is always a good idea to schedule jobs efficiently or to use multiple Data Services Job servers.

ⓘ Note

DOP is a global setting that affects the entire landscape.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Degree of parallelism \[page 301\]](#)

[Grid computing \[page 304\]](#)

13.4.6 Grid computing

You can perform grid computing using the Data Services Job Server group. This group is a logical group of multiple Data Services Job Server. When you install Information Steward, you can assign a single Data Services Job Server group for that Information Steward instance. There are two ways you can utilize the Job Server group with Distribution level setting: distribution level table and distribution level sub-table.

A single profiling or rule task can work on one or more “tables”. The term “table” is used in a general sense of the number of records with rows and columns. In reality, this can come from an RDBMS table, a flat file, an SAP application, and so on.

ⓘ Note

DOP and distribution level are global settings and affect the entire landscape.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Distribution level table \[page 305\]](#)

[Distribution level: Sub table \[page 305\]](#)

[When to use grid computing \[page 306\]](#)

13.4.6.1 Distribution level table

When you set the distribution level to Distribution level table, each table of the task is executed on separate Data Services Job Server in the group. Setting DOP is effective on the independent machines, and one task on that particular server could be further distributed.

For example, if you have 8 tables in a task and it is submitted to a Data Services Job Server group with 8 Data Services Job Servers, then each Data Services Job Server processes one table. If the DOP is set to 4, then each Data Services Job Server tries to parallelize the task into 4 execution units, one for each particular table. There is no interdependency between the different job servers; they share no resources.

When a Data Services Job Server group receives a task that involves multiple tables and the distribution level is set to Table, it uses an intelligent algorithm that chooses the Data Services servers based on the available resources. If a particular server is busy, then the task is submitted to a relatively less busy Data Services server. These calculations are based on the number of CPUs, RAM, and so on. If you have two Data Services servers, one with many resources and another with low resources, it is quite possible that the bigger server gets a proportionally higher number of tasks to execute.

You can also choose servers based purely in a "round robin" fashion, in which case the task is submitted to the next available Data Services server. For more information, see the *SAP Data Services Administrator Guide*.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Grid computing \[page 304\]](#)

13.4.6.2 Distribution level: Sub table

Note

This setting is only effective for column profiling. Use this setting with caution, as it may have a negative impact on the performance of other types of profiling, such as advanced profiling.

When you set the *Task distribution level* option to *Sub table*, Data Services distributes a single table task across multiple machines in the network based on the DOP setting. The basic idea is to split a single task into multiple independent execution units and send them to different Data Services Job Servers for execution. You can think of this as DOP, but, instead of multiple CPUs of a single machine, this is across multiple machines.

For example, you have a column profiling task with 100 million rows and 40 columns, the distribution level is set to Sub-Table, the DOP is 8, and there are 8 Data Services Job Servers in the group. The task is split into 8 execution units for 5 columns each and sent to the 8 Data Services Job Servers, which then can execute it in parallel. There is no sharing of CPU or RAM. But all of the Data Services Job Servers share the same pageable cache directory. You must ensure this directory location is shared and available to all Data Services Job Servers. This location should have a very efficient disk and the network on which this setup is done should be very fast so that it does not become a Job Servers bottleneck; otherwise, the gains of parallel processing will be negated.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Grid computing \[page 304\]](#)

13.4.6.3 When to use grid computing

Configuring grid computing is similar to degree of parallelism, but with the additional aspects of distribution level.

- Distribution level table: Use when you have many concurrent profiling and rule tasks that work on large amount of data. Set the distribution level to Table, so that individual tasks are sent to different servers.
- Distribution level sub-table: Use when you have a very few column profiling tasks on very large amount of data. In this case, use distribution level "Sub-table". It is important that these machines share an efficient hard disk and that they are connected by a fast network.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Grid computing \[page 304\]](#)

13.4.7 Using SAP applications as a source

Information Steward allows direct integration with the SAP Business Warehouse and the SAP ERP Central Component (ECC) system. One of the main benefits is the ability to connect directly to the production SAP systems to perform data profiling and data quality analysis based on the actual, most timely data, instead of connecting to a data warehouse, which is loaded infrequently. To fully utilize this advantage without risking the performance and user experience on the production ECC system, consider these requirements.

To use utilize the back-end resources in the most efficient and sustainable way, the connection user defined for the interaction between Information Steward and the SAP back-end system should be linked to background processing. This is the recommended setup for the connection user.

The option to use a dialog user for the connection between Information Steward and the SAP back-end system should be considered carefully and should only be considered for smaller datasets not exceeding 50,000 records from a medium width table. In this case, the dialog mode processing will block a dialog process for the processing time and its resources. Using this approach on larger data sets would require changing the heap size for the maximum private memory for a dialog process and therefore has impact on the overall memory required by the ECC system. In addition, the extraction of larger sets of data will most likely exceed the recommended maximum work process runtime for dialog process on the ABAP Web Application server significantly. Depending on the amount of records and number of columns of the table the data is retrieved from, the extraction process runtime will vary. Due to the significant resource required and the

extended runtime and allocation of a dialog process the usage of dialog user for the communication is not recommended.

You can connect to the SAP ECC system and retrieve data for profiling and data quality analysis using one of the following methods.

- Transfer the data directly via the RFC connection established.
In this case the data is stored in an internal table first and then transferred via the dialog mode RFC connection.
- Run in background mode processing scenario 1.
The extracted data is written into a file to a shared directory for both the ABAP Web Application and the Information Steward. After the extraction is completed, Information Steward picks up the file for further processing.
For this first background mode processing scenario, specify the SAP Applications connection parameters *Working directory on SAP server* and *Application Shared Directory*, and set *Data transfer method* to Shared directory.
- Run in background mode processing scenario 2.
In this case the data is written to directory on the back-end server, which is made available to Information Steward via an FTP server connection. Similar to the second scenario, the data files are then picked up by Information Steward from the directory specified, in this case via the FTP connection path.
For this second background mode processing scenario, specify the *Working directory on SAP server* and the *FTP* parameters when you define the SAP Applications connection, and set *Data transfer method* to FTP.

The data transfer via option two and three have slightly longer turnaround times, but consumes fewer resources on the back-end system.

It is recommended that you perform optimized requests for data extraction from the SAP system when you define Information Steward profiling tasks. This optimization should be considered in all scenarios. Therefore, specify only a subset of the columns in your profile task and use filters to focus on a specific data set. The more optimized the request, the faster the extraction and overall process execution. In this case, Information Steward only requests the data for extraction as defined by the filter. The filter criteria is passed to the SAP system.

Related Information

[SAP Applications connection parameters \[page 151\]](#)

[SAP application connection best practices \[page 318\]](#)

[General best practices \[page 315\]](#)

[Profiling and rules \[page 316\]](#)

[Organizing projects \[page 318\]](#)

13.4.8 Multi-threaded file read

When reading flat files, the Data Services engine can route the tasks in multiple threads to achieve higher throughput. With multi-threaded file processing, the Data Services engine reads large chunks of data and

processes them simultaneously. In this way, the CPUs are not waiting for the file input/output operations to finish.

You can set the number of file processing threads to the number of CPUs available. Use this setting when you have a large file to run profile or rule tasks on.

Related Information

[Scalability and performance considerations \[page 296\]](#)

13.4.9 Data Insight result set optimization

This optimization is applicable to all of the scheduled profile and rule tasks. Information Steward provides the ability to optimize redundant tasks for Data Insight. If the profile or rule results set for a particular table is already available, then it is not processed. This is controlled by the *Optimization Period* setting. Data in certain tables does not change very often. The results set is refreshed at a specified frequency. It is not required to execute that again within that time period. If the data is not going to change in a profiling task, there is no need to process the task again. Similarly, if the scorecards are calculated only on a nightly basis, there is no need to recalculate the score.

Suppose the *Optimization Period* is set to 24 hours. A rule task was executed for Table1 and the results set is already stored. If that same task is tried again within 24 hours, it is not processed again. Imagine another case where a single rule task involves Table1 and Table2. In this case, rules are not executed on Table1, but they are processed for Table2 because this table does not have a results set available.

If you want to always obtain the latest results due to the changing nature of the data, set the *Optimization Period* to the expected period of change in data.

Note

Any profiling or rule tasks that are run on demand do not use this optimization and all of the data is processed.

Related Information

[Scalability and performance considerations \[page 296\]](#)

13.4.10 Performance settings for input data

For Data Insight functionality, Information Steward provides the following settings to improve performance.

These settings are available for both profile and rule tasks when you create a task. Defaults are set in *Configure Applications*.

Max Input Size and Input Sampling Rate

When you create a rule or profiling task, specify the maximum rows and the rate at which you want to process them. Because processing time and resource requirements are proportional to the number of records being processed, these settings are very important. Set these numbers only to what is required for the task. For example, if you know that there are 100 million rows in a table and you want to obtain a sense of your data profile quickly, set the *Max Input Size* to 1 million rows and the *Sampling Rate* to 100. Every 100th record will be processed up to a maximum of 1 million records.

Filter condition

You can also control exactly which data is processed using the filter condition. Because the number of records affects performance, set the filter condition and process only the amount of data required.

Suppose you have 10 million records for all countries/regions and there are 1 million records for the U.S. If you are interested in profiling data for the U.S. only, you should set the filter for country/region to US. In this way, only 1 million records are processed.

To further improve performance (if applicable), you can combine the filter condition with *Max Input Size* and *Sampling Rate*.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Performance settings for input data \[page 308\]](#)

[Information Steward configuration settings \[page 44\]](#)

13.4.11 Settings to control repository size

For Data Insight functionality, Information Steward provides the following settings to control the size of the Information Steward repository. This depends on the number of records as well, because the more data, the bigger the potential results set. However, the repository size can be controlled for even very large amounts of data.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Profile performance settings \[page 310\]](#)

[Rule processing \[page 311\]](#)

[Metadata Management \[page 311\]](#)

13.4.11.1 Profile performance settings

When you configure Data Insight, set options that control profile settings, including maximum sample size and maximum distinct patterns.

These profile settings affect the size of the SAP Information Steward repository. Before you set values for the profile settings, consider what your data domain experts need to understand the data. Try to set the value as low as possible. Not only does a lower value use less repository space, but a lower value takes less time for Data Insight to render results.

Option	Description
<i>Max sample data size</i>	The sample size for each profile attribute.
<i>Number of distinct values</i>	The number of distinct values to store for the value distribution result. Note Data Insight presents any distinct values over the setting in the Other category in results.
<i>Number of patterns</i>	The number of patterns to store for the pattern distribution result. Note Data Insight presents any patterns that exceed the number of patterns set here in the Other category in results.
<i>Number of words</i>	The number of words to store for the word distribution result.
<i>Results retention period</i>	The number of days before the profiling results are deleted. The longer you keep the results, the bigger the repository. For more information, see the Purge utility.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Settings to control repository size \[page 309\]](#)

[Information Steward settings option descriptions \[page 45\]](#)

13.4.11.2 Rule processing

The size of the Information Steward repository is controlled by the size of the rule data and the length of time the data is saved to produce trend data.

Option	Description
Max sample data size	The number of failed records to save for each rule. The higher the number, the more records that are available to view as sample data. This results in a larger repository size, and the response time for viewing sample failed data is affected.
Score retention period	The number of days before the scores are deleted. The longer you keep the score data, the larger the repository size.

Related Information

- [Scalability and performance considerations \[page 296\]](#)
- [Settings to control repository size \[page 309\]](#)
- [Information Steward settings option descriptions \[page 45\]](#)
- [Utilities overview \[page 276\]](#)

13.4.11.3 Metadata Management

The size of the Information Steward repository is also controlled by the amount of metadata that is collected and retained. Optimize the amount by selectively choosing the components that you are interested in for different metadata integrators.

Related Information

- [Scalability and performance considerations \[page 296\]](#)
- [Settings to control repository size \[page 309\]](#)

13.4.11.4 Match review processing

The size of the Information Steward repository is controlled by the size or the match review data and the length of time the data is saved to produce trend data.

Option	Description
Data Review tasks retention period	The number of days before the Match Review tasks are deleted. The longer you keep the data, the larger the repository size. For more information, see the Purge utility in Utilities overview [page 276] .

13.4.11.5 Data Cleansing Advisor processing

The size of the Data Cleansing Advisor local repository is controlled by the size of the Data Cleansing Advisor staged data and the length of time the data is saved.

Option	Description
Data Cleansing Advisor retention period	<p>The number of days before the Data Cleansing Advisor staged data is deleted. The longer you keep the data, the larger the repository size.</p> <p>The default storage period is 120 days. If you set the value to -1, the staged data is never purged.</p> <p>For more information, see the Purge utility in Utilities overview [page 276].</p>

13.4.12 Settings for Cleansing Package Builder

To make Cleansing Package Builder run better, set the options for runtime parameters and scheduling when to publish cleansing packages.

13.4.12.1 JVM runtime parameters

If there are many parsed values (more than 20) per row in the data being used for Cleansing Package Builder, then the Auto-Analysis service requires adjustment to the JVM runtime parameter that controls memory allocation for EIMAPS. If the memory cap for Java is left at 1GB, even though the system has 16GB, the service will run out of memory. The best practice is to allocate 2 to 3GB of memory to the Java services via the -Xmx setting.

Related Information

[Scalability and performance considerations \[page 296\]](#)

13.4.12.2 Publishing cleansing packages

Schedule cleansing package processing during non-business hours. Depending on the size of the cleansing package, publishing can be a time consuming task. SAP-supplied cleansing packages for Name-Title-Firm are typically very large. When resources are available, have a dedicated server for publishing, if you publish frequently.

Related Information

[Scalability and performance considerations \[page 296\]](#)

13.4.13 Settings for Metadata Management

To make Metadata Management run better, set the options for runtime parameters and memory allowance for utilities and impact and lineage diagrams.

13.4.13.1 Runtime parameter for metadata integrators

Here is a list of parameters that affect performance:

- JVM arguments for metadata integrators that collect a large amount of metadata should be adjusted for higher memory allocation. It is recommended that you update the JVM parameters on the integrator parameters page to **-Xms1024m -Xmx4096m**.
- Run-time parameters for the maximum number of concurrent processes to collect metadata should be set to the number of CPUs that can be dedicated for metadata collection. Typically metadata integrators are installed on independent servers, so you can set it to the number of CPUs on the server. These parameters for parallel processing may be different for different metadata integrators.
- Each metadata integrator provides some method of performance improvement specifically for the type of metadata that it collects. For example, with SAP Enterprise Metadata Integrator, you can reduce processing time by selectively choosing different components.
- For the first time the integrator is run, the run-time parameter *Update Option* for metadata integrators should be set to *Delete existing objects before starting object collection*. For subsequent runs, change it to *Update existing objects and add newly selected objects*. For example, for SAP BusinessObjects Enterprise Metadata Integrator, you can first collect metadata only for Web Intelligence documents (by selecting only that component as specified in the previous paragraph). For the first run, set the option to *Delete existing*. In the next run, you may collect all of the Crystal Reports metadata. For subsequent runs, the parameter should be set to *Update existing*.

Related Information

[Scalability and performance considerations \[page 296\]](#)

[Common runtime parameters for Information Steward \[page 187\]](#)

13.4.13.2 Memory setting for impact and lineage diagrams

If multiple users simultaneously view large impact and lineage diagrams on Metadata Management, you should increase the memory that your web application server will use.

To increase the JVM heap size and perm size of the Tomcat web application server:

1. Open a command-line window and access the *Apache Tomcat Properties* window.

```
"C:\Program Files (x86)\SAP BusinessObjects\Tomcat6\bin\tomcat6w.exe" //ES//  
BOEXI40Tomcat
```

2. Select the *Java* tab.
3. In *Maximum memory pool*, replace the default value (1024 MB) with a larger value. For example, type the value **4096**.
4. Click *OK*.
5. Restart the Tomcat web application server.
 - a. Click **▶ Start ▶ Administrative Tools ▶ Services ▶**.
 - b. Select your Apache Tomcat service in the list of services, and click *Restart* in the left pane.

13.4.13.3 Utility configurations

The Metadata Management utilities Compute Lineage Report and Update Search index have some configuration parameters. Run *Compute Report Lineage* in *Optimized* mode so that it is updated incrementally.

Related Information

[Scalability and performance considerations \[page 296\]](#)

13.4.14 Settings for Metapedia

Metapedia is based on the concept of how terms are associated with metadata objects and Data Insight objects, and organized into categories.

A Metapedia term is a word or phrase that defines a business concept in your organization.

When creating a term, we recommend that you give it a unique name. From an information governance perspective, it's important that each term have only one definition that applies to the whole organization.

There may be situations in which multiple terms must share the same name. To allow terms to share the same name, set the *Allow Duplicated Term Names* option to *Yes*. This option is set to *No* by default.

Note

When creating a term in Metapedia, the information you enter into the *Technical name* field must be unique. If a technical name already exists in the Information Steward repository, an error displays and you won't be able to save the term.

1. From CMC Home select *Applications*.
2. Double-click *Information Steward Applications*.
3. Select *Configure Application*.
4. Scroll down to *Metapedia Options*.
5. Select *Yes* for *Allow Duplicated Term Names*. (Default is *No*.)

13.5 Best practices for performance and scalability

Learn about the best practices for each module in Information Steward.

13.5.1 General best practices

Follow these best practice guidelines:

- When using a distributed environment, enable and run only the servers that are necessary. For more information, see the topic "Enabling and disabling servers" in the *Business Intelligence Platform Administrator Guide*.
- Use dedicated servers for resource intensive servers like the Data Services Job Server, Metadata Integrators, and the Cleansing Package Builder Auto-Analysis service.
- Consider these best practices guidelines when deploying the Application Service:
 - The Application Service can be combined on the same computer with the Search Service, Data Review Service, and Administrative Task Service.
 - The Application Service should be on a separate computer than the web application server to obtain higher throughput.
 - The Application Service can be on its own computer, or it can be combined with any Metadata Integrator. The rationale for this combination is that Metadata Integrators usually run at night or other non-business hours, and the Application Service runs during normal business hours when users are performing tasks on the Information Steward web application (such as viewing impact and lineage relationships, adding tables to profile, or creating Metapedia terms).
- Install the Information Steward Web Application on a separate server. The Business Intelligence platform Web Tier must be installed on the same computer as the Information Steward Web Application. If you do not have Tomcat or Bobcat, you need to manually deploy the Information Steward Web Application.

- If you have many concurrent users, you can use multiple Information Steward web applications with Load Balancer.
For more information, see the topic “Fail-over and load balancing” in the *SAP BusinessObjects Business Intelligence platform Web Application Deployment Guide*.
- To obtain a higher throughput, the Information Steward repository should be on a separate computer but in the same sub-network as the Information Steward Web applications, Enterprise Information Management Adaptive Processing Server, Information Steward Job Server, and Data Services Job Server.
- Make sure that the database server for the Information Steward repository is tuned and has enough resources.
- Allocate enough memory and hard disk space to individual servers as needed.
- Follow good scheduling practices to make sure that resource intensive tasks do not overlap each other. Schedule them to run during non-business hours so that on-demand request performance is not affected.

Also see: Installation Guide: Deploying web applications with WDeploy

Related Information

[Data Insight best practices \[page 316\]](#)

[Metadata Management best practices \[page 319\]](#)

[Cleansing Package Builder best practices \[page 320\]](#)

13.5.2 Data Insight best practices

Follow best practices to reduce processing resources when Data Insight views consists of joins of tables and files from multiple sources.

[Profiling and rules \[page 316\]](#)

[SAP application connection best practices \[page 318\]](#)

[Organizing projects \[page 318\]](#)

[Joins in Data Insight views \[page 319\]](#)

To help reduce processing resources when Data Insight views consist of joins of tables and files from multiple sources, follow these best practices.

13.5.2.1 Profiling and rules

Profiling and rule tasks can consume a large amount of processing resources. If you expect that your SAP Information Steward Data Insight profiling and rule tasks will consume a large amount of processing resources, there are some best practices to implement as applicable that may help reduce the amount of processing resources.

To improve the execution of Data Insight profiling tasks, take advantage of the SAP Data Services Job Server groups and parallel execution. Put the Data Services Job Server on multiple computers that are separate from the web application server.

To perform the following tasks, you must access the Data Services Server Manager on each computer.

- Add a Data Services job server and associate it with the Information Steward repository. For more information, see [Data Insight can use Data Services Job Servers \[page 102\]](#).
- Specify the path of the pageable cache that will be shared by all job servers in the *Pageable cache directory* option.

There are several more best practices to help reduce processing resources.

- For a predictable distribution of tasks when using multiple Data Services Job Servers, ensure that the hardware and software configurations are homogeneous. This means that they should all have similar CPU and RAM capacity.
- Irrespective of using DOP and/or multiple Data Services Job Servers, set the pageable cache directory on high speed and high capacity disk.
- If you are processing flat files, store them on a high speed disk so that read performance is good.
- Process only data that must be processed. Use settings such as *Max input size*, *Input sampling rate*, and *Filter conditions* appropriately.
- When working with Data Insight views, use the correct join and filter conditions so that you are pulling in only required rows. Also use the join ranking and caching options for faster processing.
- Choose only columns that you want to profile or the rules that you want to calculate for the score. Selecting all columns may lead to redundant processing.
- Word distribution profiling is done only on a few columns. Do not choose this for all columns. Otherwise, performance and the size of the Information Steward repository is affected.
- If you have lookup functions in rule processing, it takes more time and disk space. Typically, lookup tables are small, but if they are big, lookup tables can adversely effect performance. Ensure that the tables on which lookup is performed are small. As an alternative to lookup, you can use the SQL function.
- Choose the DOP settings and the distribution level carefully. Remember that these are global settings and affect the entire landscape.
- When doing column profiling on large amounts of data with many small capacity Data Services Job Servers, use the distribution level *Sub table* only. This setting can have an adverse effect on other types of tasks. You may want to change it back to *Table* level after that column profiling task is done.
- Store the reference data required for address profiling on a high speed and high capacity disk.
- Make sure that the database server that contains source data is tuned and has enough resources.
- Schedule the Purge utility to run during non-business hours. If column profiling and rules are executed many times a day, try to schedule the Purge utility to run more than once, so that it can increase free disk space in the repository.

Also see: User Guide: Creating a rule

Related Information

[General best practices \[page 315\]](#)

[Using SAP applications as a source \[page 306\]](#)

[Joins in Data Insight views \[page 319\]](#)

[Organizing projects \[page 318\]](#)

13.5.2.2 SAP application connection best practices

1. Choose the data retrieval method (synchronous or asynchronous) based on the data set size and performance requirements of your SAP system.
2. For smaller data sets, use the synchronous method.
3. For larger data sets, use the asynchronous method, where the data from SAP systems is written into a file that Information Steward uses.

It is recommended to use background processing on the SAP back-end, which can be controlled by the user type of the connection the user defined.

Dialog processing should be considered carefully. In this case, adjust the run-time parameter such as heap size for the maximum private memory and maximum work process runtime.

Related Information

[SAP Applications connection parameters \[page 151\]](#)

13.5.2.3 Organizing projects

How scorecards and projects are organized depends on how business users want to view the scorecards. But this organization can affect the response time for users. If a project contains many scorecards, details of all the scorecards must be retrieved for viewing. So it is good idea to create multiple projects according to the area of interest and have a limited number of scorecards in those projects.

For example, you could create projects for different geographical locations. Within each project, you could have different scorecards, such as Customer, Vendor, and so on. Or you could create projects based on Customer, Vendor, and so on, and then have scorecards based on geography.

The organization also helps you decide what data sources you want to use and the filter conditions involved for profiling and rule execution.

Another benefit of proper organization is that you can control the user security per project and restrict access to only specific users.

To avoid future problems, different Data Insight user groups should work together at the beginning of the project to decide these aspects.

Related Information

[General best practices \[page 315\]](#)

[Profiling and rules \[page 316\]](#)

[Using SAP applications as a source \[page 306\]](#)

13.5.2.4 Joins in Data Insight views

To help reduce processing resources when Data Insight views consist of joins of tables and files from multiple sources, follow these best practices.

- Use appropriate filter conditions so that you are processing only the required rows.
- Use the join ranking and caching options for faster processing.
- Generally, put SAP tables first.
- Move join filter conditions from the ON clause to the WHERE clause.
- If a filter applies to all tables, specify the filter instead of relying on the join to filter the data.
- Edit the view and change the join order on the *Join Conditions* tab, *Set Optimal Order* feature, by setting the biggest table as the first item in the list, and the smallest table as the last item in the list.
- Be careful of filtering by functions. Some functions are not pushed down (the database might not support them), and they behave unexpectedly. If possible, rely on simple constant filters.
- If using a Left Outer join, and it is possible to filter most of the right table, use a view on top of views. In other words, create a view of the left table, create a view of the right table with the filter specified, and then join the two views. This way, the filter is applied before the outer join, and it will reduce the amount of data for the join operation. It is important to reduce the data set to thousands of rows instead of millions.

13.5.3 Metadata Management best practices

Follow these best practices guidelines when configuring Metadata Management.

- Metadata integrators for SAP Enterprise, Data Services, and SAP Business Warehouse should be installed on their own dedicated servers if they require large processing time or they run in overlapping time periods with other metadata integrators or Data Insight tasks.
- Any Metadata Integrator can be combined with the Application Service. The rationale for this combination is that Metadata Integrators usually run at night or other non-business hours, and the Application Service runs during normal business hours when users are performing tasks on the Information Steward web application (such as viewing impact and lineage relationships, adding tables to profile, or creating Metapedia terms).
- Another guideline to consider for Metadata Search Service:
 - Can be on its own computer, or it can be combined with any Metadata Integrator. The rationale for this combination is that Metadata Integrators usually run at night or other non-business hours, and the Search Server runs during normal business hours when users are searching on the Metadata Management tab of Information Steward.
- The File Repository Servers should be installed on a server with a high speed and high capacity disk.
- Adjust runtime parameters correctly.

Related Information

[Changing runtime parameters for integrator sources \[page 247\]](#)

[General best practices \[page 315\]](#)

13.5.4 Cleansing Package Builder best practices

- Obtain higher throughput by putting Cleansing Package Builder Auto-Analysis service should be on a dedicated server.
- Use sample data that represents various patterns in your whole data set. Large amounts of sample data with too many repeating patterns leads to redundant processing overhead.
- Runtime parameters should be set correctly for Auto-Analysis service. Specifically, memory requirements are very important. If enough memory is not made available to the process, it will run out of memory. If the memory cap for Java is left at 1GB, even though the system has 16GB, the service will run out of memory. The best practice is to allocate 2 to 3GB of memory to the Java services via the -Xmx setting.

Related Information

[General best practices \[page 315\]](#)

[Common runtime parameters for Information Steward \[page 187\]](#)

13.5.5 Match Review best practices

Follow these best practices guidelines when defining a match review configuration.

- Select only the columns that you want to review.
- Create filters to focus on specific data.
- Ensure that the database management system has an index created on the filter column.

Also see: User Guide: Match Review, Creating or editing a match review configuration

Related Information

[General best practices \[page 315\]](#)

13.6 SAP tables and ABAP-supported functions

For better performance, use ABAP-supported functions in your views that contain SAP tables.

There are functions that are supported for use with SAP tables and ABAP programs. You can use any available functions when forming your expressions, but non-supported ABAP functions are processed differently than the ABAP-supported functions, so they will likely take longer to process.

ABAP-supported functions:

- Absolute (abs)
- Decode (decode)
- Left trim (ltrim)
- Left trim ext (ltrim_blanks_ext)
- Replace null (nvl)
- Right trim (rtrim)
- Right trim ext (rtrim_blanks_ext)
- String length (length)
- Substring (substr)
- System date (sysdate)
- System time (sysptime)
- To lower case (lower)
- To upper case (upper)

Related Information

[SAP tables and performance \[page 293\]](#)

14 Migration

14.1 Migration overview

Migration as it relates to SAP Information Steward is the process of moving application configurations between landscapes.

The application life cycle often involves multiple landscapes, with each landscape used for a different phase or set of activities. Although the exact practices vary by company, typically the life cycle moves between non-production (development and test) and production phases.

You can use SAP Information Steward in both phases. Each phase could involve different computers and require different security settings. For example, the initial test may require only limited sample data and low security, while final testing may require a full emulation of the production environment including strict security. The software provides mechanisms for moving objects between landscapes.

Related Information

[Migrating between system landscapes \[page 326\]](#)

[Develop and test environments \[page 322\]](#)

[Production environment \[page 323\]](#)

[Lifecycle phases of a cleansing package \[page 324\]](#)

14.1.1 Develop and test environments

In develop and test environments, you define and test the following objects for each module of SAP Information Steward:

- **Data Insight**
Define profile tasks, rules, and scorecards that instruct Information Steward in your data quality requirements. The software stores the rule definitions so that you can reuse them or modify them as your system evolves.
- **Metadata Management**
Define integrator sources, integrator source groups, and integrator source instances that collect metadata to determine the relationships of data in one source to data in another source.
- **Metapedia**
Implement a business glossary of terms related to your business data and organize the terms hierarchically into categories. Metadata objects can be associated with terms and Data Insight objects such as rules and scorecards.
- **Cleansing Package Builder**

Create and refine person and firm or custom cleansing packages. Cleansing packages contain the parsing rules and other information that defines how the Data Cleanse transform in SAP Data Services should parse and standardize name, firm, operational, or product data.

- **Match Review**
Define and run match review configurations to create match review tasks. The tasks display match groups created by the Match transform in Data Services. Reviewers review the match groups and make any necessary corrections.

After you define the objects, use SAP Information Steward to test the execution of your application. At this point, you can test for errors and trace the flow of execution without exposing production data to any risk. If you discover errors during this phase, you can correct them and retest the application.

The software provides feedback through trace, error, and monitor logs during this phase.

Related Information

[Production environment \[page 323\]](#)

[Migration overview \[page 322\]](#)

[Lifecycle phases of a cleansing package \[page 324\]](#)

14.1.2 Production environment

After migrating and configuring your Information Steward environment to a production environment, consider taking the following actions:

- Define schedules to run Data Insight profile or rule tasks.
- Define schedules to run metadata integrators.
- Define schedules to run match review configurations.
- Re-create associations between Metapedia terms and metadata objects or Data Insight rules and scorecards.
- Before you use any cleansing package in Data Services, republish the cleansing package in Cleansing Package Builder and give an updated copy of the Data Services ATL file to your Data Services administrator.
- As appropriate for your needs, create or modify schedules for Information Steward utilities.

During production

After you move the software into production, monitor it in the CMC for performance and results in the following ways:

- Monitor your Data Insight tasks and Metadata Integrator runs and the time it takes for them to complete. The trace and monitoring logs provide information about each task and run. You can customize the log details. However, the more information you request in the logs, the longer the task or integrator runs. Balance run time against the information necessary to analyze performance.

- Check the accuracy of your data.
Evaluate results from production runs and when necessary, return to the development or test landscapes to optimize performance and refine your target requirements.

To enhance or correct your jobs:

1. Make changes in your test environment.
2. Repeat the object testing.
3. Move changed objects back into production.

For more information about exporting an ATL file for SAP Data Services, see the Cleansing Package Builder section of the *User Guide*.

Related Information

[Migrating between system landscapes \[page 326\]](#)

[Scheduling a task in the CMC \[page 164\]](#)

[Defining a schedule to run a Metadata Integrator \[page 228\]](#)

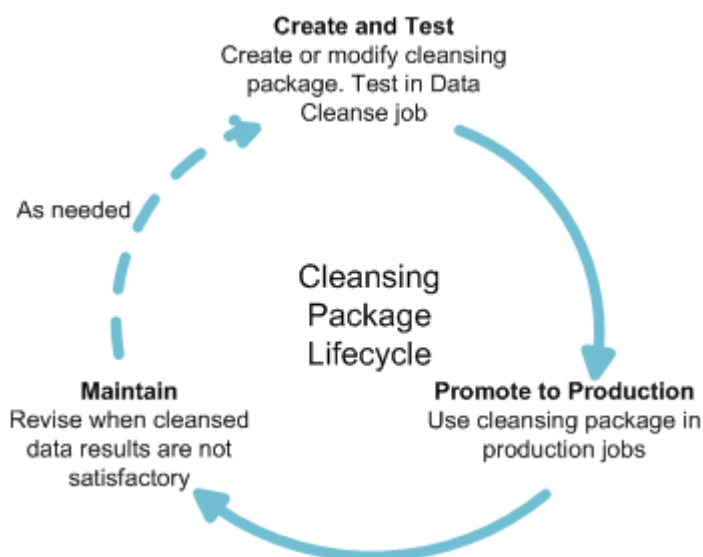
[Develop and test environments \[page 322\]](#)

[Lifecycle phases of a cleansing package \[page 324\]](#)

14.1.3 Lifecycle phases of a cleansing package

The lifecycle of a cleansing package encompasses the Cleansing Package Builder module of Information Steward and the Data Cleanse transform of Data Services. The lifecycle typically has three phases:

- Create in Cleansing Package Builder in a non-production environment and test using a Data Cleanse job.
- Promote to a production environment and use in Data Cleanse jobs.
- Maintain and revise when new data is added or threshold is no longer met.



As a cleansing package progresses between phases you can move it between landscapes using the Promotion Management tool in the Central Management Console (CMC).

Develop and test phase

Typically, Information Steward and the Data Services test environment both access the same Central Management System (CMS). In this case a published cleansing package and any changes to it are available in both Information Steward and Data Services.

Cleansing packages are created and modified in Cleansing Package Builder. During the initial creation phase a private custom cleansing package is created based on a sample of your data, or a custom or person firm cleansing package that can be based on another cleansing package. After the cleansing package meets the initial “done” criteria, it is published. The published cleansing package is available for use in a Data Cleanse job.

The Data Cleanse job is run in a Data Services test environment using a larger data sample than was used in the initial creation of the cleansing package. Based on the results of the Data Cleanse job, in Cleansing Package Builder the private cleansing package may be refined or modified and then republished using the same name. The Data Cleanse job would then be run again to check the new results.

Production phase

Once the cleansing package has been refined and the results of the Data Cleanse job produce the required cleansed data, the published cleansing package is promoted to a production environment.

To promote a cleansing package, use the Promotion Management tool in the CMC. Set up a job, and then add cleansing packages to it. In the [Add Objects](#) window, the cleansing packages available for promotion are found in `Data Quality > Cleansing Packages > Published`.

Note

Cleansing packages are referred to as “Objects” or “InfoObjects” in the CMC.

Note

You must use the same mechanism to remove a cleansing package from a system that you used to add it. If you used Promotion Management to promote a cleansing package, you must use the Promotion Management rollback option to remove it (this rolls back to the previous version). If you used the Data Services installation wizard to install the cleansing package, you must use the Data Services installation wizard to uninstall it.

For more information about Promotion Management, see the *Business Intelligence platform Administrator Guide*.

Maintenance phase

In the production environment, the cleansed data produced by Data Cleanse jobs is monitored. If new data is added or the threshold is no longer met, then the cleansing package is modified and tested in the test environment until the desired results are obtained. After that, the cleansing package is once again promoted to the production environment.

14.2 Migrating between system landscapes

You can migrate your Information Steward content between system landscapes, such as between development, test, and production environments.

Prerequisites:

When migrating SAP Information Steward between landscapes, ensure that same version of each product is installed and configured in all landscapes:

- SAP Business Intelligence (BI) platform
or
SAP Information platform services (IPS)

Note

It is recommended that you install IPS to provide flexibility to upgrade Data Services and Information Steward independently from the BI platform.

- SAP Data Services
- SAP Information Steward

Although each landscape must contain the prerequisite software, the landscapes themselves do not need to be exactly identical. For example, in your source landscape you might have an Oracle database for your Information Steward repository and a Websphere web application server. Your target landscape might have a SQL Server database for your Information Steward repository and a Tomcat web application server.

You can migrate many objects at the same time using the Promotion Management tool in the Central Management Console (CMC). Otherwise, you can migrate smaller objects through the Information Steward import and export feature.

To migrate content between system landscapes:

1. Ensure that the prerequisite software has been installed and configured.
2. Create an inventory of objects in your source Information Steward environment (the environment from which you want to migrate).

Take note of objects you want to migrate or recreate in your target environment, such as:

- Projects or objects within a project
You can export all project content together or choose individual objects within a project, such as specific rules, views, tasks, and so on.
- Connections
- Schedules

- Cleansing packages
- Match Review configurations
- Custom attributes
- Metapedia terms and categories
- Users and their group membership information

The objects in your inventory are unique to your environment and needs.

3. In your source Information Steward environment, export the following items:
 - custom attribute definitions
 - projects (an exported project includes rules, rule bindings, views, file formats, key data domains, tasks, and tables)
 - Metapedia terms and categories
 - Match Review configurations
4. In your target landscape (the landscape to which you want to migrate your Information Steward environment), create projects which correspond to those in your source landscape.
The contents of a project can be migrated, but the project definition cannot.
5. In your target Information Steward environment, import the files you exported in step 3.

Note

You must import custom attribute definitions before you import project content or Metapedia content.

6. Promote your published cleansing packages from the CMS in your source landscape to the CMS in your target landscape using the promotion management tool in the Central Management Console (CMC).

The cleansing packages available for promotion are found in `Data Quality > Cleansing Packages > Published`.

For more information about the promotion management tool, see the *SAP Business Intelligence platform Administrator Guide*.
7. As needed, recreate the objects in your inventory that have not been migrated. Define Data Insight connections, configure Metadata Integrators, and so on.
8. If your target is IPS, but your source has BI launch pad users that want to access Information Steward features such as viewing the lineage of a document (Crystal Report or Web Intelligence document) or viewing the Metapedia definition, configure the BI launch pad from the Information Steward application in the CMC. For more information see [Configuring BI Launch Pad for Information Steward \[page 254\]](#).

Related Information

[Object portability \[page 335\]](#)

[Configuring BI Launch Pad for Information Steward \[page 254\]](#)

[Exporting and importing Match Review configurations \(task\) \[page 339\]](#)

14.2.1 Objects supported for migration

Use the Promotion Management tool in the Central Management Console (CMC) to move Information Steward objects from an existing Information Steward environment to the new Information Steward environment.

For instructions see [Migrating CMS content using Promotion Management \[page 329\]](#).

All the information associated with each object, including user security, is retained. Access more information about the Promotion Management tool in the CMC Help file ([Help > Help](#)).

To access the Promotion Management tool:

1. From the *CMC Home* page, select *Promotion Management* from the *Manage* list. The *Promotion Jobs* tab opens.
2. Select *New Job* and complete the options as applicable.
3. Click *Create*. The new job is created, and the *Add Objects from the system* dialog box opens.

You can move the following objects from the *Add Objects from the system* dialog box. For each object, select the objects and click *Add*, or *Add & Close*.

Object	Path in Promotion Management
Profiling connections	Select ▶ Information Steward ▶ Connections ▶ from the file tree at left.
Data Insight projects	Select ▶ Information Steward ▶ Enterprise Profiling ▶ Projects ▶ from the file tree at left.
Profiling tasks and rule tasks	<ol style="list-style-type: none"> 1. Select ▶ Information Steward ▶ Enterprise Profiling ▶ Projects ▶ from the file tree at left. 2. Click <i>Manage Dependencies</i>.
Metadata Management Integrator Sources	Select ▶ Information Steward ▶ Metadata Management ▶ Integrator Sources ▶ from the file tree at left.
Metadata Management Source Groups	Select ▶ Information Steward ▶ Metadata Management ▶ Source Groups ▶ from the file tree at left.
User-created utilities and scheduling information for Information Steward utilities	Select ▶ Information Steward ▶ Metadata Management ▶ Utilities ▶ from the file tree at left.
Metapedia user security settings	Select ▶ Information Steward ▶ Metadata Management ▶ Metapedia ▶ from the file tree at left.
Published cleansing packages	Select ▶ Data Quality ▶ Cleansing Packages ▶ Published ▶ from the file tree at left.
Private cleansing packages	Select ▶ Data Quality ▶ Cleansing Packages ▶ Private ▶ from the file tree at left.
Information Steward security (users and user groups)	<ol style="list-style-type: none"> 1. Select the applicable user type (such as <i>User</i> or <i>User Groups</i>) from the file tree at left. 2. Choose the applicable user groups.

14.2.2 Migrating CMS content using Promotion Management

Use the Promotion Management tool in the Central Management Console (CMC) to move objects from an existing Data Services or Information Steward 4.x environment to the latest version environment.

Note

You cannot use Promotion Management to move Data Insight objects from one environment to another, such as from a test environment to a production environment. Instead, use the import—export process. For information about using the import—export process, see the *SAP Information Steward User Guide*.

Example

After you perform setup and testing on a development system, export projects and objects from the development system and import them into the production system.

Promotion Management retains all the information associated with the migrated objects, including user security. For complete information about Promotion Management, see the *SAP BusinessObjects Business Intelligence platform Administrator Guide* or *Information platform services Administrator Guide*.

1. Log into the CMC on the target server, the new machine on which you installed IPS.
2. Under the *CMC Home* dropdown list, select *Promotion Management*.
3. The *Promotion Jobs* tab opens.
4. Connect to the source system:
 - a. Choose **Settings > Manage Systems**.
 - The *Manage Systems* dialog box opens.
 - b. Click *Add* and add the connection details for the source server in the *Add System* dialog box.

The source system is the system where you have the previous versions of BI, Data Services, and Information Steward installed.
 - c. Click *Mark as Origin* to select it and click *OK*.

Note

Keep the *Manage Systems* dialog box open for the next step.

5. Connect to the destination system:
 - a. Click *Add* and add the connection details for the source server in the *Add System* dialog box.

The source server is the new machine on which you installed IPS and the new version of Data Services and Information Steward.
 - b. Click *OK*, and then click *Close*.

Note

Keep logged into the CMC for the next step.

6. Create a job:
 - a. Choose *New Job*, and then enter a name for the migration job in the *New Job* tab.
 - b. Select the location from the *Save Job in** dropdown list.

- c. Select [Login to a New CMS](#) from the *Source** dropdown list.
- d. Select the source system, enter the user name and password, and then click [Login](#).
- e. Optional. Select an option from the Destination system dropdown list:
 - [Login to a New CMS](#): Places the objects on the target system when you promote the job.
 - [Output to LCMBAIR File](#): Exports the objects to a BAIR file and then imports the file on the destination system.
- f. Click [Create](#).

The *Add Objects from the system* dialog box opens.

7. Expand the nodes in the column at left to display objects at right. Select one or more objects listed in the table, and then click [Add](#).

In the following table, view the information in the **Path in Promotion Management** column to find the listed object. For example, for [Data Services > Repositories](#), expand the Data Services node and select Repositories to see the Repositories objects.

Product	Object	Path in Promotion Management
Data Services	Repositories	Data Services > Repositories >
Data Services	RFC Configurations	Data Services > RFC Configurations >
Data Services	Published cleansing packages	Data Quality > Cleansing Packages > Published >
Data Services	Data Services security (users and user groups)	Users or User Groups. For more information see the <i>SAP BusinessObjects Business Intelligence platform Administrator Guide</i>
Information Steward	Profiling connections	Information Steward > Connections >
Information Steward	Data Insight projects	Information Steward > Enterprise Profiling Projects >
Information Steward	Data Review	Information Steward > Data Review >
Information Steward	Profiling tasks and rule tasks	<ol style="list-style-type: none"> 1. Select a project object from Information Steward > Enterprise Profiling Projects > 2. Click Manage Dependencies. 3. Select the profiling tasks and rule tasks to upgrade. 4. Click Apply & Close.
Information Steward	Metadata Management Integrator Sources	Information Steward > Metadata Management > Integrator Sources >

Product	Object	Path in Promotion Management
Information Steward	Metadata Management Source Groups	▶ Information Steward ▶ Metadata Management ▶ Source Groups ▶
Information Steward	User-created utilities and scheduling information for Information Steward utilities	▶ Information Steward ▶ Metadata Management ▶ Utilities ▶
Information Steward	Metapedia user security settings	▶ Metadata Management ▶ Metapedia folder ▶
Information Steward	Published cleansing packages	▶ Data Quality ▶ Cleansing Packages ▶ Published ▶
	<div style="border: 1px solid #ccc; padding: 5px; background-color: #f9f9f9;"> <p>Note</p> <p>Skip cleansing packages if you migrated them as a part of the Data Services migration.</p> </div>	
Information Steward	Information Steward security (users and user groups)	Users or User Groups. For more information see the <i>SAP BusinessObjects Business Intelligence platform Administrator Guide</i>

Note

When selecting the repositories to migrate, review the list of existing repositories and carry forward only the valid ones. If there are any obsolete or invalid repositories, deselect those from the promotion job.

Note

Promotion jobs can take a long time to complete when there is a large cleansing package. For this reason, you might want to create a separate migration job for cleansing packages. If you have multiple cleansing packages, create several separate migration jobs.

- After all objects are added, click [Close](#).

Note

You can add objects later using the [Add Objects](#) icon.

- Optional. Click the [Manage Dependencies](#) icon and select any dependent objects to migrate from the [Manage Dependencies](#) dialog box.

- Click [Apply & Close](#).

- Click the [Promote](#) icon.

The [Promote](#) dialog box opens.

- Optional. Promote security settings by following these substeps:

- Choose [Security Settings](#) at left.
- Select [Promote Security](#), [Promote Object Security](#), and [Promote User Security](#).
- Click [Save](#).

- d. Choose *Test Promote*.
- e. Click the *Test Promote* button.

Note

If there are many objects to be promoted, this may take some time to complete.

- f. Verify that the test promotion is successful.
13. If you created several migration jobs, select them and then click the *Promote* icon to complete the migration of the CMS content.

Note

Depending on the size of the projects and the contents, it might take several minutes to complete the promotion job. If the content is too big, consider breaking the contents into multiple migration jobs.

14. Click *History* to verify the job status. Click the *Refresh* icon as necessary until the job is completed.
15. Log into the CMC on the destination system to verify that Promotion Management migrated all objects (repositories, users and user groups, if any).

Note

With regard to EIM Adaptive Processing Service (APS) configuration, the IPS Data Services System Landscape Directory (SLD) configuration parameters are set to default values. If you previously modified any of the Service configuration parameters, log into the CMC to change the parameters to your custom settings.

Note

With regard to Data Services Application settings, the IPS SLD settings are set to default values. If you previously modified any of these settings, log into the CMC and change the options to your custom settings ([▶ CMC Home](#) > [Applications](#) > [Data Services Application](#) ▶).

For additional information, see *Master Guide*: “Separate Information platform services and BI platform”.

14.2.3 Migrating the repository to a different database type

You can migrate your existing repository to a new database server.

1. Back up your current Information Steward repository database.
2. Using tools and/or processes provided by your database vendor, move your current Information Steward repository database from one database server to another database server.
3. Use the ISRepositoryUtility tool to synchronize the migrated database with the current Central Management Server (CMS) system database:
 - a. Recover the repository
 - b. Complete the repository recovery

Note

You will need to upgrade your repository using the upgrade option of ISRepositoryUtility if you're upgrading from an older version of Information Steward.

Related Information

[Recovering the repository \[page 114\]](#)

[Completing the repository recovery \[page 117\]](#)

[Repository utility overview \[page 112\]](#)

14.3 SAP tables and ABAP program migration considerations

When the SAP application in a production environment has to preload the ABAP program, you must perform extra migration steps.

When you set connections to your SAP table sources in the Central Management Console (CMC), you specify the source for the ABAP program in the *ABAP execution option* parameter:

- *Execute preloaded*
- *Generate and execute*

The option that you select determines the source for ABAP programs when you migrate Data Insight views that contain SAP tables to a new landscape.

Execute preloaded

Information Steward doesn't generate the ABAP programs; the SAP application preloads the ABAP programs.

With *Execute preloaded*, migrate from test to production as follows:

1. From your test environment: Select an export option in *Export ABAP programs* and export the applicable ABAP programs.
2. Import the exported ABAP programs to your production environment using the specific SAP application's import process.
3. Perform Information Steward migration steps to migrate your projects and objects to your production environment.

Results: The Information Steward application in the production environment automatically uses the ABAP program that you imported from the test environment.

Generate and execute

Information Steward generates the ABAP programs in either the test environment or the production environment.

With *Generate and execute*, migrate projects and other objects from your test to production environments as usual. There's no need for you to export the ABAP programs from your test environment.

If you specify the *ABAP* data transfer method in the Data Insight view, the view generates the ABAP scripts that read data from SAP tables. For reading large data sets, the *ABAP* data transfer method provides better performance.

Keep the following information in mind when you select *Generate and execute*:

- When you migrate from one SAP system to another SAP system, you must export the ABAP script from the originating system and import it into the target system.
- If the Data Insight view uses the default data transfer method as defined in the CMC, you can preload the ABAP programs so there's no need to export the ABAP programs from the test environment.

Even though it seems less complicated to choose the *Generate and execute* setting, there are reasons to choose the *Execute preloaded* option.

❖ Example

You've completed your testing in the Information Steward test environment, and you're ready to migrate to a production environment. You have the *ABAP execution option* in your test environment set to *Generate and Execute*. Therefore, Information Steward has been generating the ABAP programs.

The security policy in your production environment doesn't allow applications like Information Steward to generate ABAP programs. Therefore, you have to switch your setting for *ABAP execution option* to *Execute preloaded* so the SAP application generates the ABAP programs. Also see "Work with SAP tables in views" in the *User Guide*.

For more information about migration, importing, and exporting projects and objects, see the *User Guide*.

Related Information

[Defining a Data Insight connection to an application \[page 149\]](#)

[SAP Applications connection parameters \[page 151\]](#)

[Migrating between system landscapes \[page 326\]](#)

14.3.1 ABAP program naming conventions

There are special naming conventions for ABAP programs.

When you export an ABAP program from your test environment, Information Steward provides the option to rename it. Renaming an ABAP program file prevents possible conflicts with any existing programs in the environment. If you decide to rename the ABAP program, use the naming conventions listed below:

- must be under 30 characters
- must be alphanumeric characters (A to Z, 0 to 9)
- must start with Z, Y, or / (slash) character
- must be unique (to avoid conflicts with an ABAP program from another project and view)

📘 Note

The slash character is only allowed in the start position and cannot be used anywhere else in the name.

14.4 Object portability

You can move many Information Steward objects between landscapes.

You can move certain objects from the following list of Information Steward products:

- Data Insight
- Metadata Management
- Metapedia
- Cleansing Package Builder
- Match review
- General objects

For objects migrated by exporting from the source, then importing to the target, only the current version of each object is migrated. History is not preserved. Additionally, user security is not migrated with this mechanism.

See also:

- User Guide: Exporting user-defined relationships
- User Guide: Importing user-defined relationships
- User Guide: Export and import of term and categories with Excel

Related Information

[Migrating between system landscapes \[page 326\]](#)

[Data Insight object portability \[page 335\]](#)

[Metadata Management object portability \[page 336\]](#)

[Cleansing Package Builder object portability \[page 337\]](#)

[Match Review object portability \[page 337\]](#)

[Information Steward general object portability \[page 338\]](#)

14.4.1 Data Insight object portability

You can move each Data Insight object between landscapes as long as you use the correct mechanism based on the object type.

The following table lists each Data Insight object that you can move between landscapes and the applicable mechanism to use for migration.

Object	Mechanism to move
Projects (workspace)	Recreate in new landscape.

Object	Mechanism to move
Profiling connections	Recreate in new landscape.
Failed data connections	Recreate in new landscape.
Tables	Export from source, then import to target.
ABAP Programs	Export from source, then import to target SAP application.
Rules and bindings	Export from source, then import to target.
Scorecards	Export from source, then import to target.
Tasks (rule and profiling)	Export from source, then import to target.
Views	Export from source, then import to target.
File formats	Export from source, then import to target.
Annotations	Can't be migrated.
Custom attribute definitions and values	Export from source, then import to target.

For information about exporting and importing projects and objects, see the *User Guide*.

Related Information

[Import and export custom attribute definitions \[page 339\]](#)

14.4.2 Metadata Management object portability

You can move specific Metadata Management objects between landscapes as long as you use the correct mechanism based on the object type.

The following table contains the Metadata Management objects that you can move between landscapes, and the mechanism to use.

Object	Mechanism to move
Integrator sources	Recreate in new landscape.
Source groups	Recreate in new landscape.
Metadata objects	Recreate in new landscape.
User-defined relationships between objects	Export from source, then import to target.
User-defined preferences (impact and lineage preferences and object equivalency rules)	Recreate in new landscape.
Annotations	Cannot be migrated.
Custom attribute definitions	Export from source, then import to target.

Related Information

[Configure metadata integrator sources \[page 189\]](#)

[Creating source groups \[page 253\]](#)

[Running a Metadata Integrator \[page 227\]](#)

[Import and export custom attribute definitions \[page 339\]](#)

14.4.3 Cleansing Package Builder object portability

Cleansing Package Builder objects that you can move between landscapes:

Object	Mechanism to move	Comments
Published cleansing packages	Promotion Management in the CMC.	The cleansing packages available for promotion are found in ▶ Data Quality ▶ Cleansing Packages ▶ Published .
Data Services ATL files	Recreate in new landscape.	To create an ATL file in the new landscape, import a cleansing package as a private cleansing package, and then republish it. To retrieve the ATL file, click ▶ More ▶ Get Data Services ATL .

See also:

- SAP Business Intelligence platform Administrator Guide
- User Guide: Exporting an ATL file for Data Services

14.4.4 Match Review object portability

You can move Match Review objects between landscapes using various mechanisms.

Match Review objects that you can move between landscapes:

Object	Mechanism to move	Comments
Configurations	Export from source, then import to target.	For details, see the Related Information.
Tasks	Cannot be migrated.	New tasks are created automatically when a Match Review configuration is run.
Match Review connections	Recreate in new landscape.	

Object	Mechanism to move	Comments
My Worklist	Cannot be migrated.	
Annotations	Cannot be migrated.	
Reviews and review history	Cannot be migrated.	

For more information about creating and running match review configurations, see the Match Review section of the *User Guide*.

Related Information

[Defining a Match Review connection \[page 270\]](#)

[Exporting and importing Match Review configurations \(task\) \[page 339\]](#)

14.4.5 Information Steward general object portability

Information Steward general objects that you can move between landscapes:

Object	Mechanism to move	Comments
Data Services Job Servers	Recreate in new landscape.	Used to run Data Insight and Match Review tasks.
Custom attributes	Export from source, then import to target.	
User security	Either promote using Promotion Management or create in new landscape.	Promotion Management will move your current users, user groups, and permissions. If your new environment has different users, it may be more efficient to create new user security then to promote existing information.
Schedules (tasks, integrators, utilities)	Create in new landscape.	
Information Steward utilities	Configure in new landscape	

Related Information

[Import and export custom attribute definitions \[page 339\]](#)

[Creating users for Information Steward \[page 54\]](#)

[Adding users and user groups to Information Steward groups \[page 56\]](#)

[Scheduling a task in the CMC \[page 164\]](#)

[Defining a schedule to run a Metadata Integrator \[page 228\]](#)

[Utilities overview \[page 276\]](#)

14.5 Exporting and importing objects

14.5.1 Import and export custom attribute definitions

To share custom attributes between systems, export their definitions and then import the definitions to another instance of SAP Information Steward.

⚠ Caution

To retain the custom attributes associated with your rules or views, export the custom attributes *separate* from the related rule or view.

After you upgrade or install a new version, import the custom attribute definitions to the new environment *before* you import the rules or views.

Information Steward stores exported custom attributes in the following files:

- `CustomAttributeDefinitions.xml`: Contains the exported custom attributes.
- `CustomAttribute_export_<datetime>.zip`: Contains each `CustomAttributeDefinitions.xml` file from the specific `<datetime>`.

14.5.2 Exporting and importing Match Review configurations (task)

You can export Match Review configurations and then import them into another instance of Information Steward to share them between systems.

Exported Match Review configurations are stored in an `.xml` file (`CustomAttributeDefinitions.xml`), within a `.zip` file (named `MRConfiguration_date_time.zip` by default).

To import Match Review configurations:

1. Log in to Information Steward with a user name that belongs to one of the following groups:
 - Data Review Configuration Manager
 - Data Review Administrator
2. In the *Manage* list, choose *Match Review Configurations*.
3. Click the *Import* icon.
4. In the *Import Match Review Configurations(s)* window, click *Browse* to open a window where you will choose the file
5. In the *Select file to upload* window, select the `.zip` file that you want to import, and click *Open*.
6. Back on the *Import Match Review Configurations(s)* window, if you want a different name for the configuration on the target system, change the name in *Configuration Name*.
7. If you want a different name for the target connection, select one in the drop-down list.
8. Click *Import*.
9. Select the name of the configuration you want to export and click the Export icon

15 Supportability

15.1 Information Steward logs

For each profiling and rule task and Metadata Integrator run, SAP Information Steward writes information in the following logs:

- Database Log - Use the database log as an audit trail. This log is in the Information Steward Repository. You can view this log while the Metadata Integrator or Data Insight profile or rule task is running. The default logging level for the database log is `Information` which writes informational messages, such as number of reports processed, as well as any warning and error messages. It is recommended that you keep the logging level for the database log at a high level so that it does not occupy a large amount of disk space.
- File Log - Use the file log to provide more information about a Metadata Integrator or Data Insight profile or rule task run. The Metadata Integrator creates this log in the Business Objects installation directory and copies it to the File Repository Server. You can download this log file after the Metadata Integrator run completed. The default logging level for the file log is `Configuration` which writes static configuration messages, as well as informational, warning, and error messages. You can change the logging level for the file log if you want more detailed information. If your logs are occupying a large amount of space, you can change the maximum number of instances or days to keep logs.

15.1.1 Metadata Management logs

15.1.1.1 Log levels

Logging levels and descriptions

Each log level logs all messages at that level or higher. Therefore, for example, the default logging level, [Information](#), logs informational, warning, and error messages. Similarly, if you change the log level to [Integrator trace](#), Information Steward logs trace, configuration, informational, warning, and error messages.

Log levels

Log level	Description
Off	Turns off logging any messages.
Error	Logs messages that indicate a serious failure.
Warning	Logs messages that indicate a potential problem.
Information	Logs informational messages.
Configuration	Logs static configuration messages.

Log level	Description
Integrator trace	Logs integrator tracing information.
SQL trace	Logs SQL tracing information.
System trace	Logs highly detailed tracing information.
All	Logs all messages.

15.1.1.2 Changing Metadata Management log levels

To change the Metadata Management log levels, you must have the Schedule right on the integrator source.

1. On the Information Steward page in the Central Management Console (CMC), expand the *Metadata Management* node, and expand the *Integrator Sources* node to display all configured integrator sources.
2. Select the integrator source for which you want to change the logging level by clicking anywhere on the row except its type.

Note

If you click the integrator type, you display the version and customer support information for the integrator.

3. Select *Action > Schedule* in the top menu tool bar.
4. Click the *Parameters* node in the tree on the left.
5. From the drop-down list, select the logging level that you want for *Database Log Level* or *File Log Level*.
6. Click *Schedule*.

Future runs of the recurring schedule for this integrator source will use the logging level you specified.

15.1.1.3 Viewing integrator source logs

To view Metadata Management integrator source logs:

1. From the Central Management Console (CMC) click Information Steward.
The *Information Steward* page opens with the *Information Steward* node selected in the Tree panel.
2. Expand the *Metadata Management* node and select the *Integrator Sources* node.
A list of configured integrator sources appears in the right panel with the date and time each was last run.
3. Select the integrator source and click *Action > History* in the top menu tool bar.
The *Integrator History* pane displays each schedule in the right panel.
4. Select the schedule name and click the icon for *View the database log*.
The *Database log* shows the task messages which are a subset of the message in the log file.
5. To find specific messages in the *Database log* window, enter a string in the text box and click *Filter*.
For example, you might enter **error** to see if there are any errors.

Note

For information about troubleshooting integrator sources, see [Troubleshooting \[page 250\]](#)

6. To close the *Database log* window, click the *X* in the upper right corner.

15.1.2 Data Insight logs

15.1.2.1 Changing Data Insight log levels

To change the Data Insight log levels, do the following:

1. On the Information Steward page in the Central Management Console (CMC), expand the *Data Insight* node, and expand the *Projects* node.
2. Select the project source, and select the project for which you want to change the log level by clicking anywhere on the row except its type.
3. Select *Action > Schedule* in the top menu tool bar.
4. Click the *Parameters* node in the tree on the left.
5. From the drop-down list, select the log level that you want for *Database Log Level* or *File Log Level*.
6. Click *Schedule*.

Future runs of the recurring schedule for this Data Insight profile or rule task will use the logging level you specified.

15.1.2.2 Viewing Data Insight task logs

To view Data Insight task logs:

1. From the Central Management Console (CMC) click Information Steward.
The *Information Steward* page opens with the *Information Steward* node selected in the Tree panel.
2. In the Tree panel, expand the *Data Insight* node.
3. Expand the *Projects* node.
4. Select the name of your project in the Tree panel.

A list of tasks appears in the right panel with the date and time each was last run.

5. Select the task and click *Action > History* in the top menu tool bar.

The *Data Insight Task history* pane displays each instance the task was executed.

6. Select the instance name and click the icon for *View the database log*.
The *Database log* shows the task messages which are a subset of the message in the log file.
7. To find specific messages in the *Database log* window, enter a string in the text box and click *Filter*.
For example, you might enter **error** to see if there are any errors.

Note

For information about troubleshooting Cleansing Package Builder, see “Troubleshooting” in the *User Guide*.

8. To close the *Database log* window, click the *X* in the upper right corner.

15.1.2.3 Web Services log

If an error occurs when using the Information Steward Web Services, change the log level to show more details.

1. Log in to the Central Management Console (CMC) as administrator.
2. At the top of the CMC Home screen, select *Applications* from the navigation list.
3. Select *Information Steward Application* in the *Applications Name* list.
4. Click **► Action ► Configure Web Services ►** in the top menu tool bar.
5. From the *Log Level* drop-down list, select one of the following values:

Log level	Description
Debug	Log debugging, informational, warning, and error messages. To avoid security risks associated with clear text passwords and logins stored in this logfile, use a lower log level setting.
Information	Log informational, warning, and error messages
Warning	Log messages that indicate a potential problem
Error	Log messages that indicate a serious failure

6. Run your Information Steward Web Services client and view the Web Services log in the Information Steward log file, which is located in this directory:

```
<SAP install path>\SAP BusinessObjects Enterprise XI 4.x\logging
```

15.1.3 Metadata Browsing Service and View Data Service logs

Metadata Browsing Service and View Data Service are used by SAP Information Steward to connect and view data in profiling sources. Communication occurs between the following components:

- SAP Information Steward Web Application
- SAP solutions for enterprise information management (EIM) Adaptive Processing server
 - SAP Information Steward
 - SAP Data Services View Data Service
 - Data Services Metadata Browsing Service
- SAP Data Services backend engine

Log locations

The following log files are in the platform log directory where the Central Management Server (CMS) is installed: %BOE Install%\SAP BusinessObjects\SAP BusinessObjects Enterprise XI <version>\logging

Requests and responses between the Information Steward Web Application and the Information Steward APS Services (within the SAP solutions for EIM Adaptive Processing Server) are contained in the `InformationSteward.Explorer.log`.

Requests and responses that occur within the SAP solutions for EIM Adaptive Processing Server either between the Data Services APS Services and Information Steward APS Services or between the Data Services APS services and the Data Services backend engine are contained in the following directories:

- ICC.MetadataService_<hostname>_<timestamp>_<version>.log
- ICC.Viewdataservice_<hostname>_<timestamp>_<version>.log

Requests and responses for communicating that occurs within the Data Services backend engine between the Data Services backend engine and the Data Services APS services within the EIM Adaptive Processing Server are contained in the following logs:

%DS_COMMON_DIR%\log\MetadataService

%DS_COMMON_DIR%\log\ViewdataService

Look for the most recent log file associated with job execution:

<hostname>_<serviceprovider>_<timestamp>_trace<version>.txt

Log levels

For Information Steward and Data Services logs, you can configure the level of information collected as shown in the following table:

Log level	Information Steward	Data Services
None	No information is collected	No information is collected
Info	No information is collected (same as None)	All traces
Finer	All traces, requests, and responses	All traces, requests, and responses

Related Information

[Configuring Metadata Browsing Service and View Data Service \[page 95\]](#)

[Metadata Browsing Service configuration parameters \[page 96\]](#)

[View Data Services configuration parameters \[page 98\]](#)

15.1.4 Viewing Information Steward logs

The log files are in the platform Log directory where CMS is installed, <%BOE Install%>\SAP BusinessObjects Enterprise XI 4.0\logging. For example, C:\Program Files (x86)\SAP BusinessObjects\SAP BusinessObjects Enterprise XI 4.0\logging

You can view the following log files:

- InformationSteward.Administrator.log
- InformationSteward.AdminService.log

- `InformationSteward.ApplicationService.log`
- `InformationSteward.DataReviewService.log`
- `InformationSteward.DataReviewTaskSchedulingService.log`
- `InformationSteward.Explorer.log`
- `InformationSteward.IntegratorService.log`
- `InformationSteward.MetadataService.log`
- `InformationSteward.SchedulingService.log`
- `InformationSteward.SearchService.log`
- `InformationSteward.TaskSchedulingService.log`
- `InformationSteward.ViewdataService.log`
- `EIMAdaptiveProcessingServer.glf` (Cleansing Package Builder information)
- `EIMAdaptiveProcessingServer_dc TRACE.glf` (Data Cleansing Advisor information)

On machines where only Web Application is installed, the log files (`InformationSteward.Administrator.log` and `InformationSteward.Explorer.log`) are stored in the Web Application temp directory. For example, `C:\Program Files (x86)\SAP BusinessObjects\Tomcat6\temp\ICC`

15.2 Monitoring with CA Wily Introscope

15.2.1 CA Wily Introscope Prerequisites

To use CA Wily Introscope for integration with Information Steward, you must have the following prerequisites:

- Installation of Introscope Enterprise Manager.
- During the installation of SAP Business Intelligence platform or SAP Information platform services, do the following:
 - On the *Interscope Integration* page, select *Integration connectivity to Introscope Enterprise Manager. You will be prompted for information.*
 - On the *Connection Information to existing Introscope Enterprise Manager* page, enter the following information:

Option	Description
Enterprise Manager Host	Host name of your Introscope Enterprise Manager.
Enterprise Manager Port	Port number of your Introscope Enterprise Manager.

Note

You can also configure the Introscope agent by selecting **CMC > Servers > Server node > Placeholders** in the CMC. The Introscope Enterprise Manager host and port are also configured here for the Introscope agent to communicate with the monitoring application.

For more information:

- *Business Intelligence Platform Administrator Guide* : “ Integrating the monitoring application with SAP Solution Manager”
- *SAP Business Intelligence Platform Administrator Guide*: “Integrating the monitoring application with SAP Solution Manager”

15.2.2 Workflows

The following workflows are supported for instrumentation:

- Browse metadata
- Add tables to project
- View data
- View profiling results
- View profiling sample data
- Export/import rules/rule binding
- Import Metapedia terms
- Metadata Management browsing
- Test rules
- View score card
- View lineage

You can use CA Wily Introscope as part of SAP Solution Manager for measuring Information Steward performance instrumentation. When installing the platform, the Introscope agent is provided for your deployment

Introscope agents collect performance metrics from SAP Business Intelligence platform Java back-end servers. Agents also collect information from the surrounding computing environment. The agents then report these metrics to the Enterprise Manager.

16 Metadata Management export utility

16.1 Metadata Management export utility

To export an entire Metadata Management configuration or a subset of objects to an XML file, use the export utility.

The SAP Information Steward XML export utility is named `MMSObjectExporter`. The XML exported by this utility follows the XML Schema Definition file, `ObjectExportSchema.xsd`, located at `<Install_dir>\MM\xml\`.

The `MMSObjectExporter` utility does not export the following information:

- MITI-related objects
- The relationship between objects across integrator configurations.

The `MMSObjectExporter` utility is located in `<Install_dir>\MM\bin`. You specify the output by using required and optional command-line arguments.

To view instructions and a list of commands to use with the utility, open the `MMSObjectExporter.bat` file with a text editor.

❁ Example

The following code shows the list of commands and instructions in the `MMSObjectExporter.bat` file:

```
REM This utility exports the Metadata Manager repository objects selected
REM by its argument list to an XML file with content defined by
REM ExportSchema.xsd.
REM
REM Argument list:
REM
REM For the following required arguments, do NOT include the name of the
REM argument
REM on the command line, just the argument values in the specified order.
REM configuration - the configuration id or configuration name of the
integrator run to be exported.
REM filename - the name of the XML file, including path
REM NOTE: The value for this argument must be enclosed in
quotation
REM marks if the name includes a blank character.
REM
REM For the following optional arguments, include the name of the argument on
the
REM command line followed by a space and the argument value.
REM mainObject - the object type to be exported (default is hierarchy)
REM includeNulls - include properties with null values? (boolean, default
is TRUE)
REM includeCustomAttributes - include object Custom Attributes? (boolean,
default is TRUE)
REM includeParents - include parents of objects? (boolean, default is
TRUE)
REM includeChildren - include children of objects? (boolean, default is
TRUE)
REM includeRelationships - include object relationships? (boolean,
default is TRUE)
```

```

REM    validate - validate the exported XML file? (boolean, default is TRUE)
REM    xsdUrl - specify the URL to use for the location of the XSD in the
schemaLocation
REM    attribute of the mmExport element. Defaults to
ObjectExportSchema.xsd
REM    in InformationSteward\MM\xml
REM
REM The mainObject argument has an optional name list to limit the object
selection
REM to a specific set of objects of that type. Objects related to the
object(s)
REM specified by mainObject are also included in the xml. mainObject
examples:
REM    1. mainObject Universe
REM    2. mainObject "Universe=MyUniverse"
REM    3. mainObject "Universe=MyUniverse,YourUniverse"
REM
REM Example 1 selects all Universes in the configuration. Example 2 selects
only
REM the Universe named MyUniverse, and Example 3 selects only the two
Universes
REM named MyUniverse and YourUniverse.
REM
REM If the mainObject argument is not specified, then the top level object
types
REM will be found from the integrator's <object-hierarchy> element.
REM
REM If there are no objects found in the repository that meet the mainObject
requirements,
REM then the utility exits with an error message. If some but not all of the
objects
REM are found, then the XML file is generated, but a warning message is
written
REM to the log about the objects that were not found.

```

16.2 Exporting objects to XML

Export Metadata Management objects to XML using the MMOBJECTEXPORTER utility.

1. View the list of commands and arguments listed at the top of the MMOBJECTEXPORTER utility in [Metadata Management export utility \[page 347\]](#).
2. Open a command prompt, go to the location of the utility, and type the name of the utility with the applicable arguments to export the file to export.

❖ Example

```
mmobjectexporter CMS_source "c:\temp\first export.xml" boeUser Jane
boePassword MyPw mainObject Universe
```

The following table explains the entries in the example command line.

Entry	Description
mmobjectexporter	Invokes the utility.



Entry	Description
CMS_source	Uses the configuration argument and represents the name of the configuration for export to XML. Alternatively, you can specify the configuration ID.
"c:\temp\first exported.xml"	Uses the filename argument and represents the path and name of the XML file to create. The argument is in quotation marks because the file name contains a space.
boeUser Jane	Uses the boeUser argument to specify the name of the CMS user.
boePassword MyPw	Uses the boePassword argument to specify the password for the CMS user.
mainObject Universe	Uses the mainObject argument to select all Universes in the configuration. Alternatively, the argument could be more specific. For example, using the syntax <code>mainObject "Universe=MyUniverse"</code> selects only the Universe named MyUniverse. Using the syntax <code>mainObject "Universe=MyUniverse, YourUniverse"</code> selects the two Universes named MyUniverse and YourUniverse . The parents and children of these objects would also be included, because the arguments for includeParents and includeChildren are not invoked, and so they default to TRUE .

Important Disclaimers and Legal Information

Hyperlinks

Some links are classified by an icon and/or a mouseover text. These links provide additional information.

About the icons:

- Links with the icon : You are entering a Web site that is not hosted by SAP. By using such links, you agree (unless expressly stated otherwise in your agreements with SAP) to this:
 - The content of the linked-to site is not SAP documentation. You may not infer any product claims against SAP based on this information.
 - SAP does not agree or disagree with the content on the linked-to site, nor does SAP warrant the availability and correctness. SAP shall not be liable for any damages caused by the use of such content unless damages have been caused by SAP's gross negligence or willful misconduct.
- Links with the icon : You are leaving the documentation for that particular SAP product or service and are entering an SAP-hosted Web site. By using such links, you agree that (unless expressly stated otherwise in your agreements with SAP) you may not infer any product claims against SAP based on this information.

Videos Hosted on External Platforms

Some videos may point to third-party video hosting platforms. SAP cannot guarantee the future availability of videos stored on these platforms. Furthermore, any advertisements or other content hosted on these platforms (for example, suggested videos or by navigating to other videos hosted on the same site), are not within the control or responsibility of SAP.

Beta and Other Experimental Features

Experimental features are not part of the officially delivered scope that SAP guarantees for future releases. This means that experimental features may be changed by SAP at any time for any reason without notice. Experimental features are not for productive use. You may not demonstrate, test, examine, evaluate or otherwise use the experimental features in a live operating environment or with data that has not been sufficiently backed up.

The purpose of experimental features is to get feedback early on, allowing customers and partners to influence the future product accordingly. By providing your feedback (e.g. in the SAP Community), you accept that intellectual property rights of the contributions or derivative works shall remain the exclusive property of SAP.

Example Code

Any software coding and/or code snippets are examples. They are not for productive use. The example code is only intended to better explain and visualize the syntax and phrasing rules. SAP does not warrant the correctness and completeness of the example code. SAP shall not be liable for errors or damages caused by the use of example code unless damages have been caused by SAP's gross negligence or willful misconduct.

Bias-Free Language

SAP supports a culture of diversity and inclusion. Whenever possible, we use unbiased language in our documentation to refer to people of all cultures, ethnicities, genders, and abilities.

© 2024 SAP SE or an SAP affiliate company. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of SAP SE or an SAP affiliate company. The information contained herein may be changed without prior notice.

Some software products marketed by SAP SE and its distributors contain proprietary software components of other software vendors. National product specifications may vary.

These materials are provided by SAP SE or an SAP affiliate company for informational purposes only, without representation or warranty of any kind, and SAP or its affiliated companies shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP or SAP affiliate company products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

SAP and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP SE (or an SAP affiliate company) in Germany and other countries. All other product and service names mentioned are the trademarks of their respective companies.

Please see <https://www.sap.com/about/legal/trademark.html> for additional trademark information and notices.