

IBM InfoSphere Metadata Workbench
Version 8 Release 7

User's Guide



IBM InfoSphere Metadata Workbench
Version 8 Release 7

User's Guide



Note

Before using this information and the product that it supports, read the information in “Notices and trademarks” on page 59.

Contents

Chapter 1. Overview of InfoSphere Metadata Workbench	1	Chapter 5. Creating data lineage, business lineage, and impact analysis reports in the metadata workbench . . .	45
Chapter 2. Accessing the metadata workbench	3	Data lineage, business lineage, and impact analysis report	45
Chapter 3. InfoSphere Metadata Workbench roles	5	Asset types that are included in lineage and analysis reports	47
Chapter 4. Exploring metadata assets . . .	7	Running data lineage reports	48
Tips for navigating the metadata workbench.	7	Running business lineage reports	50
Extended data lineage	8	Contacting IBM	53
Information assets in the metadata workbench.	11	Accessing product documentation.	55
Information asset actions	12	Product accessibility	57
Asset types that are displayed in the metadata workbench	14	Notices and trademarks	59
Finding assets in the metadata workbench	34	Index	63
Querying the metadata repository	35		
Queries.	36		
Creating queries.	38		
Managing queries	41		
Running queries.	42		

Chapter 1. Overview of InfoSphere Metadata Workbench

IT professionals such as business analysts, data analysts, and ETL developers use IBM® InfoSphere® Metadata Workbench to explore and analyze relationships between information assets in the metadata repository.

InfoSphere Metadata Workbench provides IT professionals with a design-time tool for managing and understanding the assets that are generated and used by the IBM InfoSphere Information Server suite, and for extending that analysis to assets and processes that are external to the suite.

By providing lineage reports, InfoSphere Metadata Workbench supports IT professionals who are responsible for compliance and governance initiatives that require lineage information (for example, Sarbanes Oxley or Basel II requirements). By providing an impact analysis that shows the effect of changes to information management environments, InfoSphere Metadata Workbench helps IT professionals to work most efficiently.

InfoSphere Metadata Workbench supports these tasks:

- Explore information assets in the metadata repository by using these features:
 - Details about the relationships of jobs, business intelligence (BI) reports, databases, data files, tables, columns, terms, stewards, servers, extended data sources, and other assets
 - Simple and advanced search and robust querying
 - Integrated cross-suite view of information assets
 - Details of extension mappings that describe external data flow and assets
 - Graphical view of asset relationships
- Analyze dependencies and relationships of key assets and business BI reports by doing these tasks:
 - Trace lineage through jobs and databases to BI reports
 - Perform lineage analysis to understand where data comes from or goes to by using shared table information, job design information, operational metadata from job runs, and extension mappings
 - Perform impact analysis to understand dependencies and the effects of changes to a column or job in IBM InfoSphere Information Server and beyond
 - View operational metadata from job runs
- Manage metadata to obtain in-depth analysis reports by doing these tasks:
 - Create and edit descriptions of information assets
 - Import or create assets that do not originate in InfoSphere Information Server:
 - Applications, stored procedures, and files that are defined as extended data sources
 - ETL processes that are defined as extension mappings
 - Assign terms, stewards, labels, and notes to information assets

Chapter 2. Accessing the metadata workbench

You access IBM InfoSphere Metadata Workbench by using a web browser.

Before you begin

- The supported web browsers are Microsoft Internet Explorer versions 6 and 7, and Mozilla Firefox version 2.
- To view graphical reports, Adobe® Flash Player must be installed. The version must be no earlier than 10.0.22. You can download the player from <http://www.adobe.com/>.
- To access the metadata workbench, you must have the role of Metadata Workbench Administrator or Metadata Workbench User. To view business lineage reports, you must have the role of Business Glossary User. The suite administrator of IBM InfoSphere Information Server assigns suite users to roles.
- Enable JavaScript in the browser.
- Enable HTTP 1.1 in the Microsoft Internet Explorer browser.
- Enable cookies for the site.
- Your screen resolution must be set to 1024 by 768 or greater. Maximize the InfoSphere Metadata Workbench window.
- For detailed system requirements, see <http://www.ibm.com/software/data/infosphere/info-server/overview/requirements.html>.

Procedure

1. Open the web browser and enter the URL: *protocol://host_server:port/workbench*.

The URL to access the metadata workbench is based on HTTP and cluster configurations that have been set up for IBM WebSphere® Application Server.

Protocol is the communication protocol: either Hypertext Transfer Protocol (HTTP) or Hypertext Transfer Protocol Secure (HTTPS).

Host_server and *port* differ depending upon the communication protocol and WebSphere Application Server configuration (clustered or non-clustered):

Table 1. Host and port values for different configurations

WebSphere Application Server cluster configuration	Value of <i>Host_server</i>	Value of <i>Port</i> if HTTP is used	Value of <i>Port</i> if HTTPS is used
If clustering is set up	The host name or IP address and the port of the front-end dispatcher (either the HTTP server or the load balancer). Do not use the host name of a particular cluster member.	HTTP port of the front-end dispatcher (for example, 80).	HTTPS secure port (for example, 443).

Table 1. Host and port values for different configurations (continued)

WebSphere Application Server cluster configuration	Value of <i>Host_server</i>	Value of <i>Port</i> if HTTP is used	Value of <i>Port</i> if HTTPS is used
If clustering is not set up	The host name or IP address of the computer where WebSphere Application Server is installed.	HTTP transport port (configured as <i>WC_defaulthost</i> in WebSphere Application Server). Default: 9080	HTTPS transport secure port (configured as <i>WC_defaulthost_secure</i> in WebSphere Application Server). Default: 9443

2. If HTTPS is enabled, the first time that you access the metadata workbench, a message about a security certificate is displayed if the certificate from the server is not trusted. If you receive such a message, follow the browser prompts to accept the certificate, type your suite user name and password, and click **Login**.

What to do next

The default HTTP session timeout is 30 minutes. As a result, any web session that remains inactive for more than 30 minutes is terminated. You can change the default timeout value in IBM InfoSphere Information Server Web console (**Administration tab > Session Management tab > Global Session Properties**).

Chapter 3. InfoSphere Metadata Workbench roles

The suite administrator assigns roles that define the tasks that users of IBM InfoSphere Metadata Workbench can perform.

IBM InfoSphere Metadata Workbench has the following roles:

Metadata Workbench Administrator

Runs the automated and manual analysis services, publishes queries, and explores metadata models. Performs all tasks that IBM InfoSphere Metadata Workbench users can perform.

The Metadata Workbench administrator must be familiar with the enterprise database metadata and data file metadata that is imported into the repository. The administrator must also be familiar with the metadata that is used in jobs.

Metadata Workbench User

Finds and explores information assets, runs analysis reports, and creates, saves, and runs queries.

Chapter 4. Exploring metadata assets

Users of the metadata workbench can search and query information assets in the metadata repository in order to view their properties and related assets and to create reports on the flow of data through these assets.

Tips for navigating the metadata workbench

You navigate to view assets in the metadata repository and to create reports about assets and their relationships.

You move through the metadata workbench by clicking hyperlinks and buttons in the application.

Use the following tips:

- Do not use the back and forward features of your Web browser to return to pages that you visit.
- You can start any task from one of the navigation tabs: **Browse**, **Discover**, or **Advanced**.
- To view context information for an asset, move your mouse pointer over the link to the asset in a results list or on an asset information page. The hover help displays assets that contain the object. For example, if your mouse pointer hovers over the link to a job, the names of the project and engine that contain the job are displayed, and you can click either name to open the asset information page for that asset.
- To choose from the list of tasks that you can do with an asset, right-click the asset name in a results list, or click the asset name to display its asset information page.
- To return to the Welcome page, click the **Workbench** tab in the upper left corner of the window.
- To return to any asset information page that you visit, click **Add to Favorites** to add the page to the list of bookmarks or favorites in your Web browser.
- To open an asset information page without navigating away from the page that you were on, right-click the asset name and click **Open in New Window**.

Related tasks

“Finding assets in the metadata workbench” on page 34

You locate particular assets to investigate their properties and relationships and to run reports on them. You can browse, search, and query to locate assets.

“Querying the metadata repository” on page 35

You can use queries to find and report on objects in the metadata repository.

Chapter 5, “Creating data lineage, business lineage, and impact analysis reports in the metadata workbench,” on page 45

Users of the metadata workbench can create reports that analyze the flow of data from data sources, through jobs and stages, and into databases, data files, and business intelligence reports. You can report on the dependencies between assets of certain types. In addition, you can create a business lineage report that displays only the flow of data, without the details of a full data lineage report.

Extended data lineage

You can track the flow of data across your enterprise, even when you use external processes that do not write to disk, or use tools, scripts, or other programs whose metadata is not saved to the metadata repository.

The metadata repository stores information from the tools in the IBM InfoSphere Information Server suite. The metadata workbench uses this information to display the flow of information from your data sources, through IBM InfoSphere DataStage[®] and QualityStage[™] Administrator jobs, and into target data structures.

However, you might want to view data lineage that includes data flows that are not in the metadata repository, such as in the following cases:

- Your enterprise uses another vendor's extract, transform, and load (ETL) tools whose information is not automatically stored in the metadata repository
- Your InfoSphere DataStage job depends on information from stored procedures
- You use Web services that are located elsewhere
- You want to track lineage from mainframe applications and programs, such as COBOL extracts
- You receive flat files or other types of feeds from external sources
- You often run scripts at the operating system level to copy or restructure files before processing them in jobs

In these cases and others, the flow of data through tables and columns can extend beyond the metadata repository. Metadata Workbench Administrators can create extension mappings and extended data sources in the metadata workbench that enable you to track that flow and create data lineage reports from any asset in the flow.

Extension mappings

Extension mappings are source-to-target mappings that represent an external flow of data from one or more sources to one or more targets. The source or target must exist within the metadata repository, either as database or data file metadata or as a component of an extended data source.

By using extension mappings in the metadata workbench, you can create data lineage reports for the following types of flows:

- Data flows that happen completely outside of IBM InfoSphere Information Server
- Data flows that happen both outside and inside InfoSphere Information Server

Extended data sources

Some external processes, including Web services and stored procedures, do not write their data to disk. When you need to report on information about the data structure in these processes to get a clear picture of each step of the transformation process, Metadata Workbench Administrators can capture this information by creating extended data sources and importing them into the metadata workbench.

There are three major types of extended data source assets:

- Application
- Stored procedure definition
- File



Application

Represents a program that performs a specific function directly for the user or, in some cases, for another application.

For example, an application might be a database program, communication program, or SAP program that interacts with a corporate database. You can use an application extended data source to loosely model SAP or other data programs.



Object type

A grouping of methods or a defined data format that characterizes the input and output structures within a single application. For example, an object type could represent a common feature or business process within an application.



Method

A function or procedure that is defined within an object type to perform an operation. Operations pass or receive information as input parameters or output values. For example, a method can be a specific operation or procedure call for reading or writing data through the application and object type. A method might also represent the equivalent of a database table, while the input parameters and output values represent columns of the table.



Input parameter

Input parameters are the most common way to deliver information from a client to a method. Methods require information from the client to perform their intended function. This information can be in the form of presentation options for a report, selection criteria for data to be analyzed, individual columns, or many other possibilities. For example, **MONTH** and **YEAR** could be input parameters for a method that analyzes monthly sales data.



Output value

Methods retrieve and return data to the client or application in the form of output values. Output values can represent the returned

value for the database column or data file field. For example, **JANUARY** and **2000** could be output values for a method that analyzes monthly sales data.



Stored procedure definition

Stored procedures are routines that are available to applications that access database systems, and are stored within the database system. Stored procedures consolidate and centralize complex logic and SQL statements, and might update, append, or retrieve data. For example, stored procedures are used to control transactions as condition handlers or programs, and in some cases are similar to ETL transactions when they update.

The extended data source that represents stored procedures is called a *stored procedure definition* to distinguish it from the stored procedure assets that are saved in the metadata repository by InfoSphere DataStage and QualityStage Administrator.



In parameter

An in parameter carries information that is required for the stored procedures to perform its intended function. For example, variables passed to the stored procedure are in parameters.



Out parameter

An out parameter represents the value or variable returned when a stored procedure is run. For example, a field included in the result set of the stored procedure can be an out parameter.



InOut parameter

You use inOut parameters when a stored procedure requires information from the client to perform its intended function, and then processes and returns a value by using the same parameter. For example, an inOut parameter could be a variable that the stored procedure processes or aggregates and returns to the calling application.



Result column

Result columns represent the returned data values of a stored procedure, when it queries data or processes data in a database.



File

A file represents a storage area for capturing, transferring, or otherwise reading data. Files are often the source of ETL data and can be loaded and moved by using FTP. The extended data source type *file* represents files that cannot be imported into the metadata repository by standard means. Files that can be imported into the metadata repository are called *data files*, and are not extended data sources.

Related concepts

“Data lineage, business lineage, and impact analysis report” on page 45

Data lineage and impact analysis reports show the movement of data within a job or through multiple jobs and show the order of activities within a run of a job. Business lineage reports show a scaled-down view of lineage without the detailed information that is not needed by a business user. Impact analysis reports show the dependencies between assets.

Related reference

“Asset types that are included in lineage and analysis reports” on page 47

Certain asset types can be used to run data lineage, business lineage, or impact analysis reports.

Information assets in the metadata workbench

The objects stored in the metadata repository of InfoSphere Information Server are referred to as information assets in the metadata workbench. Each asset is an instance of an asset type.

Users and administrators of the metadata workbench find and display particular information assets, investigate the properties and relationships, and run reports on the assets, such as data lineage and impact analysis reports.

Each asset has its own asset information page that displays the following information:

- Properties of the asset
- Related objects, including terms, stewards and objects that contain the asset
- Information that is generated by running automated services and performing manual linking actions
- Notes about the asset
- Actions that you can perform on the asset, including editing its description, running reports, and assigning terms

When your mouse pointer hovers over a link to an information asset in a results page or in an asset information page, the context of that asset is displayed, typically the objects that contain the asset. For example, if you hover over the link to a job, the names of the project and engine that contain the job are displayed, and you can click either name to open the asset information page for that asset.

Related concepts

“Asset types that are displayed in the metadata workbench” on page 14
Asset information pages show properties, relationships, and available actions that are appropriate to the selected type of asset.

Related tasks

Chapter 5, “Creating data lineage, business lineage, and impact analysis reports in the metadata workbench,” on page 45

Users of the metadata workbench can create reports that analyze the flow of data from data sources, through jobs and stages, and into databases, data files, and business intelligence reports. You can report on the dependencies between assets of certain types. In addition, you can create a business lineage report that displays only the flow of data, without the details of a full data lineage report.

“Querying the metadata repository” on page 35

You can use queries to find and report on objects in the metadata repository.

“Finding assets in the metadata workbench” on page 34

You locate particular assets to investigate their properties and relationships and to run reports on them. You can browse, search, and query to locate assets.

Information asset actions

You can edit an asset, run reports on an asset, and take other actions that are specific to the asset from the asset information page and from the right-click menu for an asset.

The available actions for an asset are displayed in these places:

- The right side of the asset information page
- The menu when you right-click an asset name in a list

Depending on the type of asset, some or all of the following actions are displayed:

Add Note

Create a note for the selected asset. Not all assets can have notes. An asset can have multiple notes.

You can remove the note assignment in IBM InfoSphere Business Glossary.

Add to Favorites

Add the URL of the asset information page for the selected asset to the Favorites list or Bookmarks list of your web browser.

Assign to Label

Assign a label to the asset.

Labels are simple, short descriptors for assets that have something in common that is meaningful to your enterprise. Labels can be used in search and filtering. You can think of labels as keywords or tags that you apply to assets.

Labels are created, edited, and deleted in InfoSphere Business Glossary.

Assign to Steward

Assign a steward to manage the selected asset.

Stewards that you assign are also displayed in InfoSphere Business Glossary, IBM InfoSphere FastTrack, and IBM InfoSphere Information Analyzer.

In the metadata workbench, you can assign only one steward to an asset. If a steward is already assigned to an asset, the steward that you select from the list replaces the original steward.

If a steward was assigned to the asset in IBM InfoSphere FastTrack or in InfoSphere Information Analyzer, more than one steward might be displayed. If an asset has more than one steward assigned to it and you assign a steward to the asset in the metadata workbench, then the previously assigned stewards are replaced by the single steward.

You can remove the steward assignment in InfoSphere Business Glossary.

Assign to Term

Assign a term to the selected asset. An asset can have multiple terms assigned to it. Terms are created in InfoSphere Business Glossary and IBM InfoSphere FastTrack. You can remove the term assignment in InfoSphere Business Glossary.

Business Lineage

Create a report that displays a business view of data lineage.

For this option to be displayed, the administrator of IBM InfoSphere Metadata Workbench must have configured the asset to be included in business lineage reports. In addition, you must have at least the Business Glossary User role.

The following information is displayed, depending on the type of asset:

- The flow of data to or from the selected asset, through columns, and into business intelligence (BI) reports.
- The flow of data to or from the selected asset through database tables, database views or data file structures, and into BI reports.
- The context, steward, and assigned terms of each asset in the report.

You cannot drill down into an asset in the business lineage report to get more information.

Copy Name

Copy the name of the asset to the clipboard.

Copy Shortcut

Copy the URL of the asset information page for the selected asset to the clipboard.

Data Lineage

Create a report that displays any of the following information, depending on the type of asset:

- The flow of data to or from the selected asset, through columns and stages, through one or more jobs, and into business intelligence (BI) reports.
- The flow of data to or from the selected asset through one or more jobs, through database tables, database views or data file structures, and into BI reports.

Delete Delete the asset from the metadata repository.

Edit Edit the description of the selected asset. You can add images to the asset information page for some types of assets. You can assign custom attributes to some types of assets.

Edit Aliases (Business Name)

Assign an alias name for the asset that gives a business meaning to the asset. For example, a BI report whose name is Source10 can be assigned the alias business name Risk_Level_3.

Edit Note

Edit every field of the note.

Exclude from (or Include in) Business Lineage

Remove the asset from business lineage reports if the asset is currently included. Include the asset in business lineage reports if it is currently excluded.

For this option to be displayed, you must have the Metadata Workbench Administrator role.

Graph View

Display a graphical model view of the asset, its related assets, and the relationships to them.

Model View

Display a model view of the asset that shows the class within the relationships of logical data model to physical data model to implemented data resources.

Impact Analysis

Create a report that displays the assets that depend on the presence of the selected asset and the assets whose presence the selected asset depends on.

Open Details

Open the asset information page in the current window. Any changes that you made in the current window but did not save are lost.

Open Details in New Window

Open the asset information page in a new tab in the current window. The current window remains open.

Remove Note

Delete the note from the asset. The name of the deleted note is displayed with a strikethrough mark. When the page is refreshed, the deleted note is not displayed.

You can also display context for a specific asset by moving your mouse pointer over the link to the asset. The hover help typically displays assets that contain the selected asset. For example, if you hover over the link to a job, the names of the project and engine that contain the job are displayed, and you can click either name to open the asset information page for that asset.

Asset types that are displayed in the metadata workbench

Asset information pages show properties, relationships, and available actions that are appropriate to the selected type of asset.

Table 2. Asset types that are displayed in the metadata workbench

Asset type	Definition
 Annotation	A comment that is created by developers of IBM InfoSphere DataStage and QualityStage jobs.

Table 2. Asset types that are displayed in the metadata workbench (continued)

Asset type	Definition
 Application	An extended data source asset that represents a program designed to perform a specific function directly for the user or for another application. Applications are the general collection of methods and parameters for reading or writing data.
 BI collection	A structure that organizes data within a business intelligence (BI) model. BI collections are the data sources of BI reports.
 BI collection member	The basic abstraction of a data value that is projected from a database column.
 BI hierarchy	An organizational structure that defines an ordering or relationship of data within a BI collection.
 BI level	A logical step within a BI hierarchy.
 BI model collection	A grouping of BI collections that are relevant to a BI application.
 BI report field	A field in a BI report that is typically sourced from a database column. Some BI report fields, including page numbers and section headers, are not data fields.
 "BI report" on page 21	A report that is sourced from a database and is imported into the metadata repository of IBM InfoSphere Information Server.
 Category	A type of directory or folder that contains and references terms and that organizes the InfoSphere Business Glossary in a hierarchy. A category can also contain other categories.
 Column analysis summary	An IBM InfoSphere Information Analyzer process that describes the condition of data at the field level.
 Column definition	A column-level data definition that stores data values within a InfoSphere DataStage and QualityStage table definition.
 Column mapping	A row in an IBM InfoSphere FastTrack mapping specification that describes a transformation from one or more source columns and terms to one or more target columns and terms.
 Connector	A software component that provides access from InfoSphere DataStage and QualityStage to an external source of data.
 Custom attribute	A user-created attribute that stores additional information about glossary terms and categories in the metadata repository.
 Data column	A supertype that includes the different types of physical data columns that are contained within the metadata repository: database columns, data file fields, BI report members, and report fields.
 Data element	A user-defined data type.

Table 2. Asset types that are displayed in the metadata workbench (continued)

Asset type	Definition
 "Data file" on page 23	A software file that stores data in the form of data file structures.
 Data file field	A field within a data file structure. A data file field is equivalent to a database column and is the smallest data unit that is used to store the data values of an object.
 Data file structure	A collection of data file fields in a data file. A data file structure is the file equivalent of a database table.
 "Database" on page 26	A relational storage collection that is organized by schemas and procedures. A database stores data that is represented by tables.
 Database column	A column in a database table.
 Database connection	A connection for accessing a database or file, for example, an ODBC or Oracle connection.
 Data rule	The data rules of an InfoSphere Information Analyzer project that are assigned to a term.
 Extension mapping	An extended data source asset that represents an external flow of data from one or more sources to one or more targets.
 Extension mapping document	An extended data source asset that is a document with extension mappings.
 File	An extended data source asset that represents a storage area for capturing, transferring, or reading data. A file is typically loaded and moved by using an FTP process. A file is often the source of ETL transactions.
 Folder	A user-defined tree structure for storing the contents of InfoSphere DataStage and QualityStage projects.
 Foreign key	A non-unique identifier that defines a relationship between two database tables. A foreign key in one table typically matches the primary key in the related table.
 Foreign key definition	A foreign key relationship between pairs of table definitions.
 Host	A computer that hosts databases or data files.
 Host (Engine)	A computer that hosts the engine components of IBM InfoSphere Information Server products. The engine runs parallel, server, and sequencer jobs to extract, transform, load, and standardize data. The engine computer can also host databases and data files.
 IMS™ database	The root object that defines an IMS database and its organization.
 IMS field	A field of an IMS segment.

Table 2. Asset types that are displayed in the metadata workbench (continued)

Asset type	Definition
 IMS segment	An IMS segment type, its position within the IMS hierarchy, and its relationships to other segments.
 In parameter	An extended data source asset that delivers information from a client to a stored procedure definition.
 IBM InfoSphere Information Server report	A report that is created and saved in the console or the Web console.
 Information service	A single operation or a collection of operations that expose results from processing by information providers.
 Information services application	A container for a set of services in IBM InfoSphere Information Services Director. All services within a single application are deployed or undeployed together.
 Information services operation	A container for the business logic of an information service. The operation describes the actual task that is performed by the information provider. Examples of operations include jobs, IBM InfoSphere Federation Server queries, or the invocation of a stored procedure in an IBM DB2® database.
 Information services project	A collaborative environment in IBM InfoSphere Information Services Director that contains applications, services, and operations.
 InOut parameter	An extended data source asset that represents a parameter that combines the input parameter and the output parameter.
 Input parameter	An extended data source asset that delivers information from a client.
 "Job" on page 30	An InfoSphere DataStage and QualityStage job design specification. There are several types of jobs: <ul style="list-style-type: none">  Mainframe job  Parallel job  Sequence job  Server job
 Job run	A collection of the activities that are generated when a compiled job is run.
 Job run activity	A single activity of a job run.
 Job run event	The outcome of a job run or job run activity. Events indicate the set of resources that are affected by a job run, for example, the number of rows that are read from a particular table. Event types include read, write, and fail. Fail events have the following icon:  .

Table 2. Asset types that are displayed in the metadata workbench (continued)

Asset type	Definition
 Link	The path that links and defines the flow of data between two stages in a job.
 Local container	A grouping of job content and logic, such as stages and links, that can be reused within the same job.
 Logical data model	A set of related entities and their business associations that is defined in an entity-relationship model. A logical data model can be implemented by a physical data model or by a database schema.
 Machine profile	The paths and parameters for accessing a mainframe computer. Machine profiles are created in IBM InfoSphere DataStage and QualityStage Designer.
 Mapping project	A container that organizes mapping specifications and associated data resources in IBM InfoSphere FastTrack.
 Mapping specification	A container for a set of mappings in IBM InfoSphere FastTrack. The mapping specification describes how data is extracted, transformed, or loaded from one data source to another.
 Mapping specification generation	A set of mappings in IBM InfoSphere FastTrack that define an InfoSphere DataStage and QualityStage job.
 Method	An extended data source asset that represents a function or a procedure that performs an operation. A method can contain input parameters and output values. A method either sends information by using an input parameter asset or receives information by using an output value asset.
 Notes®	Notes about an asset that are created by the user.
 Object type	An extended data source asset that represents a grouping of methods or a defined data format that characterizes the input and output structures within a single application. For example, an object type could represent a common feature or business process within an application.
 Out parameter	An extended data source asset that returns data to the stored procedure definition asset.
 Output value	An extended data source asset that returns data to the client or to an application asset. An output value is the returned value for the database column or data field data.
 Parameter	The runtime value or the default design value for a parameter that is used in a job, stored procedure, or stage type.
 Parameter set	A group of job parameters that are used together and can be reused.

Table 2. Asset types that are displayed in the metadata workbench (continued)

Asset type	Definition
 Physical data model	A design schema for information assets that defines the physical structures and relationships of data within a subject domain or application. Physical data models are independent of implementation or platform details. They can implement a logical data model and can be implemented by a database schema or by a data file.
 Policy	Documentation of the constraints and maintenance procedures that apply to a particular object. A policy documents and captures additional information about business rules and processes.
 Primary key	A unique identifier of a database table that can also be used to define relationships between tables.
 Result column	An extended data source asset that represents the data that is returned from a database query.
 Routine	A built-in or user-defined routine that is called in a derivation or constraint, or that is called before or after a job or stage.
 "Schema" on page 28	A named collection of related database tables and integrity constraints. A schema defines all or a subset of the data that is in a database.
 Data server	A supertype of host that includes data servers and engines.
 Shared container	A grouping of job content and logic, such as stages and links, that can be used by multiple jobs.
 "Stage" on page 31	A component instance that performs a unit of work within an InfoSphere DataStage and QualityStage job or container. There are separate icons for each type of stage.
 Stage column	A flow variable or column that is used to denote data flow items within a link or stage.
 Stage type	A component type that provides the implementation and structure of a stage. Each stage is associated with a stage type.
 Stage variable	A type of stage that is defined by IBM InfoSphere DataStage and that typically has an action, such as concatenate or a calculation, associated with it. The output of a stage variable is mapped to a stage column. Stage variables can be reused in the stage and can be used for lookup data.
 Standardization component	A component file in a standardization rule set.
 Standardization rule set	A series of customizable files that define how to process input data for the Standardize and Investigate stages in IBM InfoSphere QualityStage.

Table 2. Asset types that are displayed in the metadata workbench (continued)

Asset type	Definition
 "Steward" on page 33	A user or group who is designated as responsible for one or more information assets in the metadata repository. Stewards are created and managed in IBM InfoSphere Business Glossary.
 Stored procedure	A procedure that is stored in the database to encode behavioral aspects of data manipulation.
 Stored procedure definition	An extended data source asset that represents a procedure that is stored in the database to encode behavioral aspects of data manipulation, for example, assertions, constraints, and triggers. A stored procedure definition can also produce data in tabular form.
 Table analysis summary	An InfoSphere Information Analyzer process that consists of primary key analysis and the assessment of multicolumn primary keys and potential duplicate values.
 Table definition	A table-level data definition that structures data values within an InfoSphere DataStage and QualityStage project. Table definitions contain column definitions.
 "Term" on page 23	A word or phrase that classifies one or more information assets that are in the metadata repository. Each term has a parent category. Terms and categories are created and managed in IBM InfoSphere Business Glossary.
 Term History	The changes that have been made to the description or other properties of a term since the term was first defined. Term history is displayed in IBM InfoSphere Business Glossary.
 Transforms Function	A built-in or user-defined macro expression that is used in a derivation or constraint in InfoSphere DataStage and QualityStage.
 "Transformation project" on page 32	The root of the InfoSphere DataStage and QualityStage folder tree. Projects hold collections of objects such as jobs, stages, and table definitions.
 User group	A group of users of IBM InfoSphere Information Server. User groups are designated in the Administration tab of the Web console.
 View	A dynamic or virtual database table whose data is computed or collated.
 Warehouse mapping document	An extended data source asset that is a document with warehouse mappings from IBM InfoSphere Warehouse. A warehouse mapping represents an external flow of data from one or more source databases to one or more target databases.

Related concepts

“Information assets in the metadata workbench” on page 11

The objects stored in the metadata repository of InfoSphere Information Server are referred to as information assets in the metadata workbench. Each asset is an instance of an asset type.

Related tasks

Chapter 5, “Creating data lineage, business lineage, and impact analysis reports in the metadata workbench,” on page 45

Users of the metadata workbench can create reports that analyze the flow of data from data sources, through jobs and stages, and into databases, data files, and business intelligence reports. You can report on the dependencies between assets of certain types. In addition, you can create a business lineage report that displays only the flow of data, without the details of a full data lineage report.

“Querying the metadata repository” on page 35

You can use queries to find and report on objects in the metadata repository.

“Finding assets in the metadata workbench” on page 34

You locate particular assets to investigate their properties and relationships and to run reports on them. You can browse, search, and query to locate assets.

“Creating queries” on page 38

Users of the metadata workbench can create simple and complex queries to find assets in the metadata repository. Queries are based on the attributes and relationships of a selected asset type.

“Managing queries” on page 41

Metadata Workbench Administrators can edit, publish, import, and export queries. Metadata Workbench Users can import queries, edit the queries that they create, and modify published queries to save as new queries.

BI report

A business intelligence (BI) report is the metadata structure of a business intelligence report that is sourced from a database.

BI reports are instances of the ReportDef class of the Business Intelligence model.

Use a bridge to import BI reports into the metadata repository.

You can perform the following actions in the metadata workbench:

- Run impact analysis reports and lineage reports that trace the flow of information through jobs, stages, and databases into BI reports.
- Assign a term or a steward to a BI report.
- Add an image or edit the description in the asset information page of the BI report.

The asset information page for a BI report lists the properties of the BI report and displays related assets of the following types:

BI Report

Displays the general properties, the business name or the alias name of the BI report, whether the BI report is included in business lineage reports, the label of the report that describes its business meaning, the terms that are assigned to the report, the steward assigned to the report, and the database tables that are sources for the report.

Report fields are the database columns of the report. BI report collection is the group of database columns and user-defined columns of the database table that is used to build the report.

Double-click the name of the BI report collection, database table, or report field to display its details.

Notes Displays notes about the asset that were created in the metadata workbench or in other products in IBM InfoSphere Information Server.

Right-click the name of the note to do the following actions to the note:

- Display it in a new window
- Edit or delete it, if you created it
- Bookmark the URL of the asset information page that the note belongs to

Modification Details

Displays the name of the user who created or last modified the asset, and the date and time of creation and last modification of the asset.

Category

A category is a directory or folder that contains a set of glossary terms. Typically, the terms contained in each category are related in some way that is meaningful to your organization. You can use IBM InfoSphere Business Glossary to define categories.

A category is an instance of the Category class in the Common Model.

To edit the category or to assign a steward to the category, use InfoSphere Business Glossary.

The asset information page for a category contains the following information:

Category

Displays the general properties of the category, including name, short and long descriptions, and steward.

Terms Displays a list of all the terms contained in this category and a list of terms that this category refers to. Click a term to display the details of that term.

Attributes

Displays the custom attributes associated with the category. Attributes store information about terms and categories that do not fit into the standard attributes and relationships of the business glossary.

Subcategories

Displays the parent category and subcategories of the current category.

Information Server Report

Displays reports about the asset that are published by the reporting services in IBM InfoSphere Information Server. Click the name of the report to display it.

Notes Displays notes about the asset that were created in the metadata workbench or in other products in IBM InfoSphere Information Server.

Right-click the name of the note to do the following actions to the note:

- Display it in a new window
- Edit or delete it, if you created it
- Bookmark the URL of the asset information page that the note belongs to

Modification Details

Displays the name of the user who created or last modified the asset, and the date and time of creation and last modification of the asset.

Term

You use a term to classify, define, and group assets according to the needs of the enterprise. A term is sometimes referred to as a glossary term or a business term.

A term is an instance of the BusinessTerm class in the Common Model.

Terms are contained within categories, which make up the structure of the glossary. In the metadata workbench, IBM InfoSphere FastTrack, and IBM InfoSphere Information Analyzer, you can assign terms to other assets. Terms can be assigned to multiple assets, and assets can be assigned to multiple terms.

You can edit terms in IBM InfoSphere Business Glossary.

In the metadata workbench, you can assign terms on the information page of the asset, or right-click the name of the asset and then click **Assign Term**.

The asset information page for a term includes the following categories of information:

Term Displays the properties of the term, the parent category, the steward, and synonym terms.

Attributes

Lists the custom attributes of the term and their value. Double-click the name of the custom attribute to display its details.

Related IT Assets

Displays assets that the term is assigned to. Double-click the name of the asset to display its details.

Displays unpublished IBM InfoSphere FastTrack column mapping specifications that use the term.

Related Terms

Displays the terms that relate to or are related to by the selected term.

Notes Displays notes about the asset that were created in the metadata workbench or in other products in IBM InfoSphere Information Server.

Right-click the name of the note to do the following actions to the note:

- Display it in a new window
- Edit or delete it, if you created it
- Bookmark the URL of the asset information page that the note belongs to

Modification Details

Displays the name of the user who created or last modified the asset, and the date and time of creation and last modification of the asset.

Data file

A data file is a file-system storage medium for data. Data files contain one or more data file structures, which are the file equivalents of database tables.

A data file is an instance of the DataFile class in the Common Model.

You can edit the description and alias (business name), assign a term or steward, include or exclude from business lineage reports, and run lineage and impact analysis reports on the asset.

The asset information page for a data file includes the following categories of information:

File Displays the properties of the data file, the server name and directory path where the data file is located, the business name or the alias name of the data file, whether the file is included in business lineage reports, the label of the file that describes its business meaning, and the data file structures that the data file contains.

File Design Usage

Displays jobs that read from or write to the data file, based on job design information that is interpreted by the automated services.

File Operational Usage

Displays jobs that read from or write to the data file at run time, based on operational metadata that is interpreted by the automated services.

File User-Defined Usage

Displays jobs that read from or write to the data file, based on the data file design information that is interpreted by the automated services.

Notes Displays notes about the asset that were created in the metadata workbench or in other products in IBM InfoSphere Information Server.

Right-click the name of the note to do the following actions to the note:

- Display it in a new window
- Edit or delete it, if you created it
- Bookmark the URL of the asset information page that the note belongs to

Policy Displays policies that are associated with the asset. Policies are rule sets that are created in IBM InfoSphere Information Analyzer.

Modification Details

Displays the name of the user who created or last modified the asset, and the date and time of creation and last modification of the asset.

Data file structure

A data file structure is created when you import a sequential file. It contains information about the structure of the imported file.

A data file structure is an instance of the DataCollection class in the Common Model.

You can edit the description and alias (business name), assign a term or steward, and run lineage and impact analysis reports on the asset.

The asset information page for a data file structure contains the following information:

File Structure

Displays the general properties of the data file structure, including the name of the structure, name of the product that created the structure, the business name or the alias name of the data file structure, whether the data file structure is included in business lineage reports, the label of the file

that describes its business meaning, short and long description, name of terms that are assigned to the structure, fields in the structure, and steward.

Right-click the name of a field to display details about that field.

File Structure Design Information

Displays stages that write to and read from the data file structure, based on job design information that is interpreted by the automated services.

File Structure Operational Information

Displays jobs that read from or write to the data file structure at run time, based on operational metadata that is interpreted by the automated services.

File Structure User-Defined Information

Displays stages that read from or write to the data file structure, based on the results of manual linking actions that are performed by the Metadata Workbench Administrator. Also displays any extension mapping documents with mappings between source and target assets.

Indexes and Analysis

Displays the primary key that is defined for the data file structure and the foreign keys that the data file structure refers to. Foreign keys are fields in a different table. These keys relate tables to each other. The name of the column and of the table where the foreign key is sourced are also displayed.

Displays the IBM InfoSphere Information Analyzer analysis summary report, if one exists, of the table. Double-click the name of the report to display the report contents.

Policy Displays policies that are associated with the asset. Policies are rule sets that are created in IBM InfoSphere Information Analyzer.

Notes Displays notes about the asset that were created in the metadata workbench or in other products in IBM InfoSphere Information Server.

Right-click the name of the note to do the following actions to the note:

- Display it in a new window
- Edit or delete it, if you created it
- Bookmark the URL of the asset information page that the note belongs to

Modification Details

Displays the name of the user who created or last modified the asset, and the date and time of creation and last modification of the asset.

Host

A host is the hardware that hosts a database, a data file, or a project whose metadata is imported into or referenced by IBM InfoSphere Information Server. A host can have both databases and engines.

Hosts are instances of the HostSystem class in the Common Model.

The asset information page for a host includes the following information:

Host Displays the properties of the host, the network node, and the databases and data files that the host sources.

Notes Displays notes about the host that are created in IBM InfoSphere Business Glossary or in IBM InfoSphere Information Analyzer.

Right-click the name of the note to do the following actions to the note:

- Display it in a new window
- Edit or delete it, if you created it
- Bookmark the URL of the asset information page that the note belongs to

Database

A database is a relational database or catalog that stores data that is defined by database tables and schemas.

A database is an instance of the Database class in the Common Model.

You can run impact analysis and lineage reports. You can assign a term and a steward to a database.

The asset information page for a database includes the following information:

Database

Displays the properties of the database, the server that hosts the database, the business name or the alias name of the database, whether the database is included in business lineage reports, the label of the database that describes its business meaning, and the schemas.

Note: The name that you give in **Alias (Business Name)** is not the alias that is used by IBM InfoSphere DataStage jobs.

Database Design Information

Displays jobs that read from or write to the database, based on job design information that is interpreted by the automated services.

Database Operational Information

Displays jobs that read from or write to the database at run time, based on operational metadata that is interpreted by the automated services.

Database User-Defined Information

Displays jobs that read from or write to the database, based on the results of manual linking actions that are performed by the Metadata Workbench Administrator.

BI Report Information

Displays business intelligence (BI) reports and BI report models that are sourced from the database tables. Right-click the name of the report or the report model to display a list of additional tasks that you can do.

Notes Displays notes about the asset that were created in the metadata workbench or in other products in IBM InfoSphere Information Server.

Right-click the name of the note to do the following actions to the note:

- Display it in a new window
- Edit or delete it, if you created it
- Bookmark the URL of the asset information page that the note belongs to

Policy Displays policies that are associated with the asset. Policies are rule sets that are created in IBM InfoSphere Information Analyzer.

Alias Displays alternative names that InfoSphere DataStage jobs use to reference the database. The alias name is defined by the manual linking action, Database Alias.

Modification Details

Displays the name of the user who created or last modified the asset, and the date and time of creation and last modification of the asset.

Database table

A database table is the structure that represents and stores columns within a schema. The metadata workbench displays information about schemas that have been imported into IBM InfoSphere Information Server.

A database table is an instance of the DataCollection class in the Common Model.

You can edit the description and alias (business name) of the table, assign a term or steward to be associated with the table, or run impact analysis and lineage reports.

The asset information page for a database table contains the following information:

Database Table

Displays the general properties of the table: name, the business name or the alias name of the database table, whether the database table is included in business lineage reports, the label of the database table that describes its business meaning, tool that created the database table, short and long description, term assigned to the table, and steward.

Displays the name of the database that contains the table, the name of the schema, the name of any identical tables, and the name of any views that refer to the database table.

When you run manual linking actions and two schemas are identified as identical, then the database tables and database columns that are contained by the schemas are also marked as identical when their names match.

Database Table Design Information

Displays the stages that write to and read from this table, based on the job design information that is interpreted when automated services are run.

Database Table Operational Information

Displays the stages that write to and read from this table, based on the values of parameters at run time. This operational metadata is interpreted by automated services.

Database Table User-Defined Information

Displays stages that write to and read from this table, based on the results of manual linking actions.

BI Report Information

Displays the names of the business intelligence (BI) reports and the report collections that use information from this table.

Indexes and Analysis

Displays the primary key that is defined for the database table and the foreign keys that the database table refers to. Foreign keys are fields in a different table. These keys relate tables to each other. The name of the column and of the table where the foreign key is sourced are also displayed.

Displays the IBM InfoSphere Information Analyzer analysis summary report, if one exists, of the database table. Double-click the name of the report to display the report contents.

Notes Displays notes about the asset that were created in the metadata workbench or in other products in IBM InfoSphere Information Server.

Right-click the name of the note to do the following actions to the note:

- Display it in a new window
- Edit or delete it, if you created it
- Bookmark the URL of the asset information page that the note belongs to

Policy Displays policies that are associated with the asset. Policies are rule sets that are created in IBM InfoSphere Information Analyzer.

Mapping Specifications

Displays source and target mapping specifications from IBM InfoSphere FastTrack that refer to or use the database table.

Modification Details

Displays the name of the user who created or last modified the asset, and the date and time of creation and last modification of the asset.

Schema

A schema is composed of database tables and can include all or part of the data in the database. The layout of a database outlines the way data is organized into tables.

A schema is an instance of the DataSchema class in the Common Model.

You can edit the descriptions and alias (business name) of the schema, assign an image to it, assign a term or steward to the schema, whether the schema is included in business lineage reports, the label of the schema that describes its business meaning, or run impact analysis and lineage reports.

The asset information page for a schema contains the following information:

Schema

Displays the general properties of the schema, including name, the business name or the alias name of the schema, whether the schema is included in business lineage reports, owner, short and long description, steward, host server, database name, database tables, the label of the schema that describes its business meaning, and stored procedures.

The property Owner is the owner of the database who is defined during the database installation and who does administrative actions in the database. The property Steward is the steward of the schema in the metadata repository and is assigned by the metadata workbench administrator.

Notes Displays notes about the asset that were created in the metadata workbench or in other products in IBM InfoSphere Information Server.

Right-click the name of the note to do the following actions to the note:

- Display it in a new window
- Edit or delete it, if you created it
- Bookmark the URL of the asset information page that the note belongs to

Policy Displays policies that are associated with the asset. Policies are rule sets that are created in IBM InfoSphere Information Analyzer.

Schema User-Defined Information

Displays a list of extension mapping documents that have mappings between source and target assets.

Modification Details

Displays the name of the user who created or last modified the asset, and the date and time of creation and last modification of the asset.

View

A view is a virtual database table and can be imported into the metadata repository.

A view is an instance of the DataCollection class in the Common Model.

You can edit the description and alias (business name) of the view. You can also assign a term, note, or steward to the view.

The asset information page for a view contains the following information:

View Displays the general properties of the view: the view name, the business name or the alias name of the view, whether the view is included in business lineage reports, product that imported the view, the label of the view that describes its business meaning, the database that the view is derived from, a description of the view, the SQL expression that created the view, and the steward who manages the view.

Also displays the database and schema that contain the view, and the columns that are defined in the view.

View Design Information

Displays the stages that write to and read from this database view, based on the design parameters of the job that are interpreted by the automated services.

You must have run the automated services for this information to be displayed.

View Operational Information

Displays the stages that write to and read from this database view at run time, based on operational metadata that is interpreted by the automated services.

You must have run the automated services for this information to be displayed.

View User-Defined Information

Displays the stages that write to and read from this database view, based on the results of manual linking actions that are performed by the Metadata Workbench Administrator.

BI Report Information

Displays the names of the business intelligence (BI) reports that contain data that is obtained from this view.

Indexes and Analysis

Displays the primary key that is defined for the view and the foreign keys that the view refers to. Foreign keys are fields in a different view. These keys relate views to each other. The names of the column and of the view where the foreign key is sourced are also displayed.

Displays the IBM InfoSphere Information Analyzer analysis summary report, if one exists, of the view. Double-click the name of the report to display the report contents.

Specifications

Displays the IBM InfoSphere FastTrack source and target mapping specifications that refer to and use the view.

Notes Displays notes about the asset that were created in the metadata workbench or in other products in IBM InfoSphere Information Server.

Right-click the name of the note to do the following actions to the note:

- Display it in a new window
- Edit or delete it, if you created it
- Bookmark the URL of the asset information page that the note belongs to

Policy Displays policies that are associated with the asset. Policies are rule sets that are created in IBM InfoSphere Information Analyzer.

Modification Details

Displays the name of the user who created or last modified the asset, and the date and time of creation and last modification of the asset.

Job

A job is a design specification that is created in IBM InfoSphere DataStage and QualityStage Designer to extract, transform, or load data. The job can be a DataStage job or a QualityStage job.

A job is an instance of the DSJob class in the Transformation model.

The job can be a parallel, server, mainframe, or sequence job.

You can edit the description of the job, run impact analysis and lineage reports, and you can assign a term and steward.

The asset information page for a job includes the following information:

Image Displays the job as it is displayed in the Designer client.

Job Displays the properties of the job, the project and folder that it is in, and the stages and containers that it includes.

Job Design Information

Displays data items that the job reads from or writes to. Displays the previous and next jobs based on job design information that is interpreted by the automated services. Displays job design parameters and whether runtime column propagation is enabled.

Job Operational Information

Displays the previous and next jobs based on the values of parameters at run time, based on operational metadata that is interpreted by the automated services.

Job User-Defined Information

Displays the data items that a job reads from or writes to, based on the results of manual linking actions that are performed by the Metadata Workbench Administrator.

Sequence

Displays the job that sequenced the selected job.

Annotations

Displays annotations that are added to the job in the Designer client.

Notes Displays notes about the asset that were created in the metadata workbench or in other products in IBM InfoSphere Information Server.

Right-click the name of the note to do the following actions to the note:

- Display it in a new window
- Edit or delete it, if you created it
- Bookmark the URL of the asset information page that the note belongs to

Mapping Specification

Displays the mapping specifications from IBM InfoSphere FastTrack that generates this DataStage job.

Information Services Usage

Displays related operations of IBM InfoSphere Information Services Director and whether a Web service is enabled.

Modification Details

Displays the name of the user who created or last modified the asset, and the date and time of creation and last modification of the asset.

Stage

A stage is an individual step in an IBM InfoSphere DataStage job. Each stage defines a specific action or activity within the job.

A stage is an instance of the DSSStage class in the Transformation model.

You can edit the description of the stage or assign a term to be associated with the stage. The other properties of a stage are defined in InfoSphere DataStage.

The asset information page for a stage contains the following information:

Stage Displays the general properties of the stage, including name, description, term, stage type, and job.

Displays the names of links, which contain the stage columns that are input and output for the stage. Click the twistie next to the name of the link, and then click the name of the stage column to display its details.

Stage Design Information

Displays the next and previous stages that are accessed, and the name of the database tables or files that are written to and read from for this stage. The information is based on design parameters that are interpreted by the automated services.

Stage Operational Information

Displays the next and previous stages that are accessed, and the name of the database tables written to and read from. The information is based on operational metadata that is interpreted by the automated services. Double-click the name of the stage to display its details.

Stage User-Defined Information

Displays the next and previous stages that are accessed, and the name of the database tables written to and read from. This information is based on the results of manual linking actions that are performed by the Metadata Workbench Administrator. Double-click the name of the stage to display its details.

Also displays extension mapping documents that contain mappings between source and target assets.

Parameters

Displays a list of parameters and their values for the stage. Parameter values contain connection information to implemented data resources.

Notes Displays notes about the asset that were created in the metadata workbench or in other products in IBM InfoSphere Information Server.

Right-click the name of the note to do the following actions to the note:

- Display it in a new window
- Edit or delete it, if you created it
- Bookmark the URL of the asset information page that the note belongs to

Modification Details

Displays the name of the user who created or last modified the asset, and the date and time of creation and last modification of the asset.

Transformation project

A transformation project is a project that is created in IBM InfoSphere DataStage and QualityStage Designer.

A transformation project is an instance of the DSProject class in the Transformation model.

The asset information page for a transformation project includes the following information:

Project

Displays the name of the server on which the engine tier of IBM InfoSphere Information Server is installed and the folders or directories that contain the elements of the project. Double-click the name of the engine or the folder to display its details.

Includes Containers

Displays the shared containers in this project.

Includes Jobs

Displays the jobs in this project. A job consists of stages that are linked together to describe the flow of data from a data source to a data target. Double-click the name of the job to display its details.

Includes IMS Databases

Displays the Information Management System (IMS) database, if any, that is used in the project.

Includes Machine Profiles

Displays the mainframe machine profiles, if any, that are used when IBM InfoSphere DataStage uploads generated code to a mainframe.

Includes Routines

Displays the routines that use a COBOL function from a library that is external to InfoSphere DataStage. Double-click the name of the routine to display information about the routine and its source code.

Includes Parameter Sets

Displays the set of parameters, or processing variables, for jobs.

Includes Stage Types

Displays the types of stages that are in the project. Double-click the name of the stage type to display its stages and the jobs that use the stage type.

Includes Table Definitions

Displays table definitions, a set of related columns definitions that are stored in the metadata repository and that can be loaded into stages. Double-click the name of the table definition to display its details.

Includes Transforms

Displays the list of built-in and custom transforms in the project. A transform changes data from one type to a different type.

Includes Standardization Rule Sets

Displays the standardization rule sets for specific countries. A rule set determines how fields in input records are parsed.

Modification Details

Displays the name of the user who created or last modified the asset, and the date and time of creation and last modification of the asset.

Engine

An engine is the host on which the engine tier of IBM InfoSphere Information Server is installed. It can also be a server that hosts databases whose metadata is imported or referenced by IBM InfoSphere Information Server.

Engines are instances of the HostSystem class in the Common Model.

You can edit the short and long descriptions of an engine, attach an image, or assign a steward.

The asset information page for an engine includes the following information:

Engine

Lists the network node, data connection, projects that the engine hosts, data connectors that implement stages that are used in a project, and the databases and data files that the engine hosts if it is also a server. Double-click the name of connector, project, or file to display its details.

Engine Notes

Displays notes about the engine that are created in the metadata workbench or in other products in IBM InfoSphere Information Server.

Right-click the name of the note to do the following actions to the note:

- Display it in a new window
- Edit or delete it, if you created it
- Bookmark the URL of the asset information page that the note belongs to

Modification Details

Displays the name of the user who created or last modified the asset, and the date and time of creation and last modification of the asset.

Steward

A steward is a user or group that is responsible for assets in the metadata repository. The steward serves as the contact for information about those assets.

A steward is an instance of the Principal class in the Common Model.

Users and groups in IBM InfoSphere Information Server can be designated as stewards in IBM InfoSphere Business Glossary.

A steward can be assigned to categories and to terms by using InfoSphere Business Glossary. You can use the metadata workbench to assign a steward to other types of assets.

When you edit a steward in the metadata workbench, you can add an image to the asset information page, and you can edit the contact information for the steward.

The asset information page for a steward displays the following information:

User Displays contact details of the steward. ID is the login user name of the steward in InfoSphere Information Server.

Manages Information Assets

Displays assets that the steward is responsible for. Double-click the name of the asset to display its details.

Notes Displays notes about the steward that are created in the metadata workbench or in other products in InfoSphere Information Server.

Right-click the name of the note to do the following actions to the note:

- Display it in a new window
- Edit or delete it, if you created it
- Bookmark the URL of the asset information page that the note belongs to

Modification Details

Displays the name of the user who created or last modified the steward, and the date and time corresponding to creation and last modification.

Finding assets in the metadata workbench

You locate particular assets to investigate their properties and relationships and to run reports on them. You can browse, search, and query to locate assets.

About this task

When you locate an asset, you can click it to display its asset information page, or right-click the asset name and choose a task.

You can also display assets, find assets, and run queries from the Welcome page.

Procedure

Choose a method from the following table.

To do this action	Do this action
Display all assets of a particular type	On the Discover tab, click the name of an asset type, or choose an asset type from the Additional Types list and then click Display .
Browse the projects and jobs that are contained by InfoSphere Information Server engines	On the Browse tab, click Engines .

To do this action	Do this action
Browse the data servers and databases whose metadata is imported into the metadata repository	On the Browse tab, click Hosts .
Find assets by using the name or short description	On the Discover tab, click Find , then select an asset type and optionally specify additional information.
Run an existing query to find assets	On the Discover tab, select an existing query from the Run Query list and click Run .
Create a query to find assets	On the Discover tab, click Query .

Related concepts

“Tips for navigating the metadata workbench” on page 7

You navigate to view assets in the metadata repository and to create reports about assets and their relationships.

“Information assets in the metadata workbench” on page 11

The objects stored in the metadata repository of InfoSphere Information Server are referred to as information assets in the metadata workbench. Each asset is an instance of an asset type.

“Asset types that are displayed in the metadata workbench” on page 14

Asset information pages show properties, relationships, and available actions that are appropriate to the selected type of asset.

Related tasks

“Querying the metadata repository”

You can use queries to find and report on objects in the metadata repository.

Chapter 5, “Creating data lineage, business lineage, and impact analysis reports in the metadata workbench,” on page 45

Users of the metadata workbench can create reports that analyze the flow of data from data sources, through jobs and stages, and into databases, data files, and business intelligence reports. You can report on the dependencies between assets of certain types. In addition, you can create a business lineage report that displays only the flow of data, without the details of a full data lineage report.

Querying the metadata repository

You can use queries to find and report on objects in the metadata repository.

Related concepts

“Tips for navigating the metadata workbench” on page 7

You navigate to view assets in the metadata repository and to create reports about assets and their relationships.

“Information assets in the metadata workbench” on page 11

The objects stored in the metadata repository of InfoSphere Information Server are referred to as information assets in the metadata workbench. Each asset is an instance of an asset type.

Related tasks

“Finding assets in the metadata workbench” on page 34

You locate particular assets to investigate their properties and relationships and to run reports on them. You can browse, search, and query to locate assets.

Queries

Queries help users find and display information assets, their properties, and their relationships.

The metadata workbench has prebuilt queries that you can run to find information about assets. In addition, you can create and edit your own queries and use existing published queries as the basis for new queries.

Structure

Each query is based on a single asset type, but you can build queries so that they primarily display the features of related assets. For example, you might base a query on the database table asset type. But you might structure the display properties of the query to return detailed information about the business intelligence reports that are related to a single table.

You build a query by selecting an asset type and then selecting the criteria and the display properties from the list of properties that are available for that asset type.

Available Properties

The **Available Properties** list displays the following information:

- Properties of the selected asset type, such as name and description.
- Possible relationships that the asset type can have to other asset types. For example, a term can have four types of relationships to other terms and two relationships to categories.
- Properties and relationships of all related asset types, and of their related asset types, and so on.

You can expand the list at any related asset type to select properties or objects to display in the query result or to use as criteria for returning results. For example, for a query based on database table you can specify that the results display the term that classifies the job that contains the stage that writes to the database table. Or you can specify that the query return only those tables that are written to by stages in jobs that are classified by a term that begins with the letter E.

You use the **Available Properties** list to populate the **Criteria** and **Select** tabs.

Criteria

On the **Criteria** tab, you specify the conditions under which results are

returned. For example, for a query that is based on database tables, you might specify that the query returns only those tables that meet the following conditions:

- Read by a stage in a job
- Has no steward
- Classified by a term whose short description contains “customer”

You can add multiple conditions and subconditions, selecting properties and related objects from the **Properties** list to make the query results as precise as needed. By default, all conditions and subconditions must be met, but you can change the setting so that any condition can be met. You can set this value separately for each condition or each subcondition.

Depending on the type of the property, you can refine your query. For example, if the property is type text, you can narrow your query by using “Begins with”, “Is null”, “Is not”, “Contains”, and so on. If the property is type date, you can choose a date range by using “Is between” with begin and end dates. If the property is a relationship, you can choose “Is null” or “Is not null”.

Select

On the **Select** tab, you specify the properties of the selected asset type, and the properties of related asset types that you want to display in the query results. If you create a complex display, the query results are presented in tabs.

Apply criteria to selected properties

You can limit the query results to display only those assets that match the criteria in the **Criteria** tab. If you select this check box, the criteria is applied to the asset on which you are making the query and on all selected relationships.

Sharing queries and results

When a query is created, it is visible only to the user who created it. The Metadata Workbench Administrator can share queries with other users of the metadata workbench.

The administrator can publish queries so that users of the same metadata workbench installation can see them. The administrator can also export queries in WBQ format so that users of other installations can use them, and the administrator can import queries in WBQ format that users have created in other installations of the metadata workbench. The WBQ format is a proprietary format, and you cannot edit the query files outside of the metadata workbench.

You can save query results in comma-separated value (CSV) and XLS format in order to use them programmatically or to distribute the results to others.

Permissions of each role

What you can do with queries depends on your metadata workbench role. The following table lists the permissions of each role.

Table 3. Query permissions by role

Task	Metadata Workbench Administrator	Metadata Workbench User
Create queries	Yes	Yes
Delete queries that you create	Yes	Yes
Delete unpublished queries that others create	Yes	No
Delete published queries	Yes	No
View and run prebuilt and published queries	Yes	Yes
Publish queries so that other users can work with them	Yes	No
Edit a prebuilt or published query and overwrite the original.	Yes	No
Edit a prebuilt or published query and save it as a new query.	Yes	Yes
Import queries in WBQ format.	Yes	Yes
Export queries in WBQ format.	Yes	No

Example

A prebuilt query, Job Run, returns the following information:

- All jobs that have runs
- All runs of each job
- The description, project, and steward of each job
- The status of each run

Any user can edit this query to display only those runs that occurred after a certain data and time. A Metadata Workbench Administrator can save the edited query and overwrite the original prebuilt query. A metadata workbench user can save the query with a new name but cannot overwrite the prebuilt query.

Related tasks

“Creating queries”

Users of the metadata workbench can create simple and complex queries to find assets in the metadata repository. Queries are based on the attributes and relationships of a selected asset type.

“Managing queries” on page 41

Metadata Workbench Administrators can edit, publish, import, and export queries. Metadata Workbench Users can import queries, edit the queries that they create, and modify published queries to save as new queries.

“Running queries” on page 42

Users of the metadata workbench can run queries to find and report on information assets.

Creating queries

Users of the metadata workbench can create simple and complex queries to find assets in the metadata repository. Queries are based on the attributes and relationships of a selected asset type.

About this task

You can create and save a query. You can share the query with other IBM InfoSphere Metadata Workbench users by publishing the query.

Procedure

1. On the **Discover** tab, click **Query**.
2. From the **Asset Type** list, select the type of asset that you want to build the query on. The query returns the specified information about assets of this type and information about any related assets that you specify.
3. Optional: Specify criteria for returning results:
 - a. On the **Criteria** tab, click  and click **Add Condition**.
 - b. In the **Available Properties** list, select an attribute or a related asset. Related assets are indicated by a plus sign (+). You can expand a related asset to select its attributes or related assets. The selected property is displayed in the condition.
 - c. On the **Criteria** tab, specify a value for the selected property.
 - d. Optional: Click the numbered arrow on a condition to add subconditions or additional conditions. Add properties and specify values for each new condition or subcondition.
 - e. Specify whether all or any of the criteria must be met for the query to return results. To change the specification, click **All** to change to **Any**, or click **Any** to change to **All**. You must do this step for all conditions and for each individual condition or subcondition that has subconditions.

If you do not specify criteria, the query returns all assets of the type that you select in step 2.

4. Optional: Specify which results to display:
 - a. In the **Available Properties** list, select an attribute or a related asset. You can expand a related asset to select its attributes or related assets.
 - b. On the **Select** tab, click the **Select** button . The selected attribute or related asset is displayed in the Displayed Properties list.
 - c. Optional: Select additional properties and add them to the list. You can reorder the displayed properties and rename the displayed properties. If you rename a displayed property, only the labels in the display results are affected by the name change.
 - d. Optional: Select the **Apply criteria to selected properties** check box if you want to limit the query results to only those assets that match the criteria in the **Criteria** tab. If you select this check box, the criteria is applied to the asset on which you are making the query and on all selected relationships. For example, you want to create a query to display all databases that have a database table with the prefix WHS. If the **Apply criteria to selected properties** check box is clear, the query results include all database tables, even those database tables without the prefix WHS.

Database		
Asset Name	Schemas	Tables
<input type="checkbox"/> DW	<input type="checkbox"/> PRODUCT	<input type="checkbox"/> PLANT <input type="checkbox"/> PRODUCTION <input type="checkbox"/> WHS_PRODUCT
	<input type="checkbox"/> SALES	<input type="checkbox"/> WHS_SALES
<input type="checkbox"/> PRODUCT	<input type="checkbox"/> PRODUCT	<input type="checkbox"/> PLANT <input type="checkbox"/> WHS_PRODUCT

If the **Apply criteria to selected properties** check box is selected, the query results display only those database tables with the prefix WHS.

Database		
Asset Name	Schemas	Tables
<input type="checkbox"/> DW	<input type="checkbox"/> PRODUCT	<input type="checkbox"/> WHS_PRODUCT
	<input type="checkbox"/> SALES	<input type="checkbox"/> WHS_SALES
<input type="checkbox"/> PRODUCT	<input type="checkbox"/> PRODUCT	<input type="checkbox"/> WHS_PRODUCT

By default, **Apply criteria to selected properties** is selected in all new queries and in all queries that were created before InfoSphere Metadata Workbench, Version 8.5.

5. Optional: Save the query:
 - a. Click **Save**. The Save Query window opens.
 - b. Specify a name and description for the query.
 - c. Optional: Publish the query so that other users can see it and use it (administrators only). Published queries are displayed with a different icon.
 - d. Click **Save** in the Save Query window.
6. Optional: Click **Run**. The query runs and the results are displayed. For some complex displays, query results are presented with multiple tabs when multiple types of relationships are displayed.

What to do next

You can save query results to a CSV or XSL file.

Related concepts

“Queries” on page 36

Queries help users find and display information assets, their properties, and their relationships.

“Asset types that are displayed in the metadata workbench” on page 14

Asset information pages show properties, relationships, and available actions that are appropriate to the selected type of asset.

Related tasks

“Managing queries”

Metadata Workbench Administrators can edit, publish, import, and export queries. Metadata Workbench Users can import queries, edit the queries that they create, and modify published queries to save as new queries.

“Running queries” on page 42

Users of the metadata workbench can run queries to find and report on information assets.

Managing queries

Metadata Workbench Administrators can edit, publish, import, and export queries. Metadata Workbench Users can import queries, edit the queries that they create, and modify published queries to save as new queries.

About this task

Metadata Workbench Administrators can delete published queries. Metadata Workbench Users can delete only the queries that they create. If a query is not published, only the user who created the query can delete it.

Any user can edit a published query, but the Metadata Workbench User must save the published query with a new name as a user query, while the Metadata Workbench Administrator can change the published query.

Any user can import queries, but when a Metadata Workbench User imports a query it cannot be published.

Only the Metadata Workbench Administrator can export and publish queries.

Procedure

1. On the Discover tab, click **Query**.
2. In the query builder, click **Manage**.
3. In the Manage Queries window, do any of the following tasks:

To do this task	Do this action
Edit a query	<ol style="list-style-type: none">1. Select a query and click Edit.2. Modify the display properties and criteria as required.
Change the name or description of a query	<ol style="list-style-type: none">1. Select a query and click Edit.2. In the query builder, click Save.3. In the Save Query window, edit the description and click Save.

To do this task	Do this action
Publish a query	<ol style="list-style-type: none"> 1. Select a query and click Edit. 2. In the query builder, click Save. 3. In the Save Query window, select Publish Query and click Save.
Import a query	<ol style="list-style-type: none"> 1. Select a query and click Import. 2. In the Import Queries window, browse to select a WBQ file and click Import.
Export a query	<ol style="list-style-type: none"> 1. Select a query and click Export. 2. Save the query as a WBQ file.
Delete a query	Select a query and click Delete .

Related concepts

“Queries” on page 36

Queries help users find and display information assets, their properties, and their relationships.

“Asset types that are displayed in the metadata workbench” on page 14

Asset information pages show properties, relationships, and available actions that are appropriate to the selected type of asset.

Related tasks

“Creating queries” on page 38

Users of the metadata workbench can create simple and complex queries to find assets in the metadata repository. Queries are based on the attributes and relationships of a selected asset type.

“Running queries”

Users of the metadata workbench can run queries to find and report on information assets.

Running queries

Users of the metadata workbench can run queries to find and report on information assets.

About this task

You can also run a query when you create or edit the query.

Procedure

1. Select the query:
 - On the Welcome page, select the query in the **Available Queries** list.
 - On the **Discover** tab, select the query in the **Run Query** list.
2. Click **Run**.

Related concepts

“Queries” on page 36

Queries help users find and display information assets, their properties, and their relationships.

Related tasks

“Creating queries” on page 38

Users of the metadata workbench can create simple and complex queries to find assets in the metadata repository. Queries are based on the attributes and relationships of a selected asset type.

“Managing queries” on page 41

Metadata Workbench Administrators can edit, publish, import, and export queries. Metadata Workbench Users can import queries, edit the queries that they create, and modify published queries to save as new queries.

Chapter 5. Creating data lineage, business lineage, and impact analysis reports in the metadata workbench

Users of the metadata workbench can create reports that analyze the flow of data from data sources, through jobs and stages, and into databases, data files, and business intelligence reports. You can report on the dependencies between assets of certain types. In addition, you can create a business lineage report that displays only the flow of data, without the details of a full data lineage report.

Related concepts

“Tips for navigating the metadata workbench” on page 7

You navigate to view assets in the metadata repository and to create reports about assets and their relationships.

“Information assets in the metadata workbench” on page 11

The objects stored in the metadata repository of InfoSphere Information Server are referred to as information assets in the metadata workbench. Each asset is an instance of an asset type.

Related tasks

“Finding assets in the metadata workbench” on page 34

You locate particular assets to investigate their properties and relationships and to run reports on them. You can browse, search, and query to locate assets.

Data lineage, business lineage, and impact analysis report

Data lineage and impact analysis reports show the movement of data within a job or through multiple jobs and show the order of activities within a run of a job. Business lineage reports show a scaled-down view of lineage without the detailed information that is not needed by a business user. Impact analysis reports show the dependencies between assets.

You can start a report from the task list of an asset information page or from the context menu of an asset in a results list. You can track the flow from source to target or from target to source.

When you run reports, the metadata workbench displays information assets in the context of your enterprise goals. You see them not as isolated tables, columns, jobs, or stages, but as integrated parts of the process that extracts, loads, investigates, cleanses, transforms, and reports on your data.

Before you run reports, the Metadata Workbench Administrator must run the Manage Lineage utility and, where necessary, manual linking actions to set relationships between assets. To report on relationships that are created by operational metadata, you must first import operational metadata.

If a report does not return the expected results, take the following actions:

- Ensure that the Manage Lineage utility was run.
- Browse to the asset information page of information assets that you suspect are not properly linked, and expand the Design, Operational, and User-Defined information sections to validate that the correct relationships are set.
- Perform manual linking actions to set the necessary relationships.

Report types

You can run these types of reports:

Data Lineage and Impact Analysis

Both data lineage and impact analysis reports can show different types of information:

- The flow of data to or from a selected metadata asset, through stages and stage columns, through one or more jobs, into databases and business intelligence (BI) reports.

For example, a data lineage report might start with a database column that is read by a stage in a job. The report might show the following flow of data at the column level:

- A stage in the first job reads the database column
- Information flows through one or more stages in the first job until one of the stages writes to a table in a second database
- A stage in a second job reads the database column in the second database
- Information flows through one or more stages in the second job until one of the stages writes to a table in a third database
- The data in the database column in the third database is captured in a BI report

- The flow of data to or from a selected metadata asset through one or more jobs, through database tables, views, or data file structures and into BI reports and information services operations.

For example, a data lineage report might show that the sources for a BI report come from three separate jobs that write to a single database table, and that the table is bound to a BI report collection that is used by the BI report.

- The order of activities within a job run, including the tables that the jobs write to or read from and the number of rows that are written and read. You can inspect the results of each part of a job run by drilling into the job run activities to see the links, stages, and database tables, or data file structures that the job reads from and writes to.

For example, for a simple job, a data lineage report might show that the first activity read six rows from a text file, and the second activity wrote six rows to a database table. For a more complex job, the data lineage report might show the order of activities that are responsible for every read, write, or lookup.

Business Lineage

Business lineage reports show data flows only through those information assets that have been configured to be included in business lineage reports. In addition, business lineage reports do not include extension mapping documents or jobs from IBM InfoSphere DataStage and QualityStage.

You are not required to specify the flow direction of data, the analysis type, or a target asset. The business lineage report displays the graphical and textual components for only those source, target, and intermediate assets that are configured to be included in business lineage.

The Metadata Workbench Administrator configures which information assets are displayed in business lineage reports. A report is generated from

the right-click menu of an asset that is configured for business lineage. The report is read-only and you cannot get further information about the data flow or about the assets themselves.

A user of IBM InfoSphere Business Glossary Anywhere, IBM InfoSphere Business Glossary, IBM InfoSphere Metadata Workbench, and external programs such as IBM Cognos®, can create a business lineage report for an asset. The user must have at least the Business Glossary User role. The report is displayed in a new window in the web browser. For example, a business lineage report for a BI report might show the flow of data from one database table to another database table. From the second database table, the data flows into a BI report collection table and then to a BI report. The context of the database tables and the BI report collection table is displayed.

For each type of analysis, you can create a report that shows the flow of information from asset to asset that participates in the lineage or analysis flow.

Related concepts

“Extended data lineage” on page 8

You can track the flow of data across your enterprise, even when you use external processes that do not write to disk, or use tools, scripts, or other programs whose metadata is not saved to the metadata repository.

Related tasks

“Running data lineage reports” on page 48

You can run data lineage reports that combine information from job designs, operational metadata, and user-defined relationships between assets. You can see the flow of data through specific child assets, for example, the flow of data through selected database columns of a database table.

Asset types that are included in lineage and analysis reports

Certain asset types can be used to run data lineage, business lineage, or impact analysis reports.

Data lineage, business lineage, and impact analysis reports

You can run data lineage, business lineage, and impact analysis reports on the following types of assets and their children:

Table 4. Lineage and impact analysis reports on these assets and their successive child assets

Asset	Successive child assets
Application (extended data source)	<ul style="list-style-type: none"> • Object type > method > input parameter • Object type > method > output value
Business intelligence (BI)	<ul style="list-style-type: none"> • BI model > BI collection > BI collection member • BI report > BI report field
Database	<ul style="list-style-type: none"> • Schema > database table > database column • Schema > view > database column
Data file	<ul style="list-style-type: none"> • Data file structure > data file field
File (extended data source)	None

Table 4. Lineage and impact analysis reports on these assets and their successive child assets (continued)

Asset	Successive child assets
Stored procedure definition (extended data source)	<ul style="list-style-type: none"> • In parameter • Out parameter • Inout parameter • Result column

Data lineage and impact analysis reports

You can run data lineage and impact analysis reports on the following assets:

Table 5. Data lineage and impact analysis reports on these assets and their successive child assets

Asset	Successive child assets
Job	<ul style="list-style-type: none"> • Stage > stage column
Mapping in <ul style="list-style-type: none"> • Extension mapping document from IBM InfoSphere Metadata Workbench • Mapping specification from InfoSphere FastTrack • Mapping document from IBM InfoSphere Warehouse 	None

Related concepts

“Extended data lineage” on page 8

You can track the flow of data across your enterprise, even when you use external processes that do not write to disk, or use tools, scripts, or other programs whose metadata is not saved to the metadata repository.

Related tasks

“Running data lineage reports”

You can run data lineage reports that combine information from job designs, operational metadata, and user-defined relationships between assets. You can see the flow of data through specific child assets, for example, the flow of data through selected database columns of a database table.

Running data lineage reports

You can run data lineage reports that combine information from job designs, operational metadata, and user-defined relationships between assets. You can see the flow of data through specific child assets, for example, the flow of data through selected database columns of a database table.

Before you begin

The Metadata Workbench Administrator must have run the Manage Lineage and the Manage Data Source Identity utilities in the Lineage Administration section of the Administration tab. The Manage Lineage utility sets the relationships between stages, database tables, and database views so that the lineage report is accurate. The Manage Data Source Identity utility creates the relationships between the

schemas, and between the database tables and database fields of both schemas that have matching names.

About this task

Data lineage reports show the flow of information both to and from a selected asset. By default, the reports include design, operational, and user-defined relationships between assets. Duplicate assets are not displayed.

Some assets can contain child assets. The following examples of assets and their child assets can be displayed in the lineage report in a separate web browser window:

Job asset whose child assets are stages.

You can view the data as it flows through the stages that are contained in the job.

Database table asset whose child assets are database columns.

You can view the data as it flows through selected database columns of a database table. The relationship links between the selected database columns and other assets are redrawn. In addition, asset nodes that are not in the data flow are disabled. You cannot select child assets in the new browser window. For example, if you originally selected the database columns DM_SORTKEY, DM_TIMESTAMP, and DM_TXID from a database table, you cannot select a subset of these database columns from the database table in the new lineage report.

Procedure

1. Start the report by using either of the following methods:
 - In a results list, right-click the name of an asset, and choose **Data Lineage**.
 - In the task list on the asset information page for an asset, click **Data Lineage**.



2. Optional: Click  and select which types of analysis relationships to display in the report.

If you change the selection of analysis relationships, click **Refresh display** to redo the lineage graphic.

3. Optional: In the asset node, click the **Select [Column | Field | Member | Parameter]** link, if available, to trace data lineage through specific child assets. You can select no more than five child assets. You cannot select columns in a stage asset.
4. Optional: In an IBM InfoSphere DataStage job node, click the **Expand** link to see the flow of data between the stages that are contained in that job.
5. Optional: Click  to save the text of the right pane, or to save the entire graphic, the part of the graphic that is visible in the window, or the part of the graphic that you selected, to a file.

The text is saved to a file in a PDF format. The graphic is saved to a file in a JPEG image format.

If you click **Save Graph**, in the Save Graph window, select **All** to save the entire graphic, **Current View** to save the part that is visible in the window, or **Selection** to save the part that you selected.

Results

The report displays available information about the flow of data through assets.

What to do next

In the image file of the lineage graphic, click the name of an asset to display its details page in a new browser window.

In the PDF file of the text, right-click the name of an asset to get list of actions that you can do on the asset.

Related concepts

“Data lineage, business lineage, and impact analysis report” on page 45

Data lineage and impact analysis reports show the movement of data within a job or through multiple jobs and show the order of activities within a run of a job.

Business lineage reports show a scaled-down view of lineage without the detailed information that is not needed by a business user. Impact analysis reports show the dependencies between assets.

Related tasks

“Running queries” on page 42

Users of the metadata workbench can run queries to find and report on information assets.

“Finding assets in the metadata workbench” on page 34

You locate particular assets to investigate their properties and relationships and to run reports on them. You can browse, search, and query to locate assets.

Related reference

“Asset types that are included in lineage and analysis reports” on page 47

Certain asset types can be used to run data lineage, business lineage, or impact analysis reports.

Running business lineage reports

You can create business lineage reports that display the flow of information between assets that have been configured to be in business lineage reports.

Before you begin

- Only certain asset types, such as application, business intelligence (BI) report, and data file, can be included in business lineage reports. By default, all assets in these asset types are included in business lineage reports. The Metadata Workbench Administrator must configure any assets to be excluded from the business lineage report.
- You must have the Business Glossary User role to run a business lineage report.

About this task

A business lineage report is a read-only report that displays the flow of information between assets. Only those assets that have been configured to be included in business lineage reports are displayed in the lineage path.

Procedure

Run a business lineage report on a selected asset by doing either of these actions:

- In the task list on the asset information page for an asset, click **Business Lineage**.
- In a results list from a find or query action, or in the Manage Business Lineage window, right-click the name of an asset and choose **Business Lineage**.

Results

The report displays, in a new window, all of the assets that participate in the lineage path with the asset that you report on. The context, or parent, of each asset in the report is displayed. You cannot drill down into any asset in the report to obtain more information.

What to do next

You can zoom in to display selected areas of the report in more detail, save and print the report, and display the context, description, and steward of the asset. You can also send open your default email application to send feedback about the lineage report.

Contacting IBM

You can contact IBM for customer support, software services, product information, and general information. You also can provide feedback to IBM about products and documentation.

The following table lists resources for customer support, software services, training, and product and solutions information.

Table 6. IBM resources

Resource	Description and location
IBM Support Portal	You can customize support information by choosing the products and the topics that interest you at www.ibm.com/support/entry/portal/Software/Information_Management/InfoSphere_Information_Server
Software services	You can find information about software, IT, and business consulting services, on the solutions site at www.ibm.com/businesssolutions/
My IBM	You can manage links to IBM Web sites and information that meet your specific technical support needs by creating an account on the My IBM site at www.ibm.com/account/
Training and certification	You can learn about technical training and education services designed for individuals, companies, and public organizations to acquire, maintain, and optimize their IT skills at http://www.ibm.com/software/sw-training/
IBM representatives	You can contact an IBM representative to learn about solutions at www.ibm.com/connect/ibm/us/en/

Providing feedback

The following table describes how to provide feedback to IBM about products and product documentation.

Table 7. Providing feedback to IBM

Type of feedback	Action
Product feedback	You can provide general product feedback through the Consumability Survey at www.ibm.com/software/data/info/consumability-survey

Table 7. Providing feedback to IBM (continued)

Type of feedback	Action
Documentation feedback	<p>To comment on the information center, click the Feedback link on the top right side of any topic in the information center. You can also send comments about PDF file books, the information center, or any other documentation in the following ways:</p> <ul style="list-style-type: none"><li data-bbox="933 436 1414 495">• Online reader comment form: www.ibm.com/software/data/rcf/<li data-bbox="933 499 1414 531">• E-mail: comments@us.ibm.com

Accessing product documentation

Documentation is provided in a variety of locations and formats, including in help that is opened directly from the product client interfaces, in a suite-wide information center, and in PDF file books.

The information center is installed as a common service with IBM InfoSphere Information Server. The information center contains help for most of the product interfaces, as well as complete documentation for all the product modules in the suite. You can open the information center from the installed product or from a Web browser.

Accessing the information center

You can use the following methods to open the installed information center.

- Click the **Help** link in the upper right of the client interface.

Note: From IBM InfoSphere FastTrack and IBM InfoSphere Information Server Manager, the main Help item opens a local help system. Choose **Help > Open Info Center** to open the full suite information center.

- Press the F1 key. The F1 key typically opens the topic that describes the current context of the client interface.

Note: The F1 key does not work in Web clients.

- Use a Web browser to access the installed information center even when you are not logged in to the product. Enter the following address in a Web browser: `http://host_name:port_number/infocenter/topic/com.ibm.swg.im.iis.productization.iisinfsv.home.doc/ic-homepage.html`. The `host_name` is the name of the services tier computer where the information center is installed, and `port_number` is the port number for InfoSphere Information Server. The default port number is 9080. For example, on a Microsoft® Windows® Server computer named `iisdocs2`, the Web address is in the following format: `http://iisdocs2:9080/infocenter/topic/com.ibm.swg.im.iis.productization.iisinfsv.nav.doc/dochome/iisinfsv_home.html`.

A subset of the information center is also available on the IBM Web site and periodically refreshed at `http://publib.boulder.ibm.com/infocenter/iisinfsv/v8r7/index.jsp`.

Obtaining PDF and hardcopy documentation

- A subset of the PDF file books are available through the InfoSphere Information Server software installer and the distribution media. The other PDF file books are available online and can be accessed from this support document: `https://www.ibm.com/support/docview.wss?uid=swg27008803&wv=1`.
- You can also order IBM publications in hardcopy format online or through your local IBM representative. To order publications online, go to the IBM Publications Center at `http://www.ibm.com/e-business/linkweb/publications/servlet/pbi.wss`.

Providing feedback about the documentation

You can send your comments about documentation in the following ways:

- Online reader comment form: www.ibm.com/software/data/rcf/
- E-mail: comments@us.ibm.com

Product accessibility

You can get information about the accessibility status of IBM products.

The IBM InfoSphere Information Server product modules and user interfaces are not fully accessible. The installation program installs the following product modules and components:

- IBM InfoSphere Business Glossary
- IBM InfoSphere Business Glossary Anywhere
- IBM InfoSphere DataStage
- IBM InfoSphere FastTrack
- IBM InfoSphere Information Analyzer
- IBM InfoSphere Information Services Director
- IBM InfoSphere Metadata Workbench
- IBM InfoSphere QualityStage

For information about the accessibility status of IBM products, see the IBM product accessibility information at http://www.ibm.com/able/product_accessibility/index.html.

Accessible documentation

Accessible documentation for InfoSphere Information Server products is provided in an information center. The information center presents the documentation in XHTML 1.0 format, which is viewable in most Web browsers. XHTML allows you to set display preferences in your browser. It also allows you to use screen readers and other assistive technologies to access the documentation.

IBM and accessibility

See the IBM Human Ability and Accessibility Center for more information about the commitment that IBM has to accessibility.

Notices and trademarks

This information was developed for products and services offered in the U.S.A.

Notices

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

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

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785 U.S.A.

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

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
1623-14, Shimotsuruma, Yamato-shi
Kanagawa 242-8502 Japan

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:

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

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

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

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

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

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

IBM Corporation
J46A/G4
555 Bailey Avenue
San Jose, CA 95141-1003 U.S.A.

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

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

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

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

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

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

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to

IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Each copy or any portion of these sample programs or any derivative work, must include a copyright notice as follows:

© (your company name) (year). Portions of this code are derived from IBM Corp. Sample Programs. © Copyright IBM Corp. _enter the year or years_. All rights reserved.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at www.ibm.com/legal/copytrade.shtml.

The following terms are trademarks or registered trademarks of other companies:

Adobe is a registered trademark of Adobe Systems Incorporated in the United States, and/or other countries.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency which is now part of the Office of Government Commerce.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

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

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office

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

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

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

The United States Postal Service owns the following trademarks: CASS, CASS Certified, DPV, LACS^{Link}, ZIP, ZIP + 4, ZIP Code, Post Office, Postal Service, USPS and United States Postal Service. IBM Corporation is a non-exclusive DPV and LACS^{Link} licensee of the United States Postal Service.

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

Index

A

- asset
 - term 23
- asset types
 - child assets 47
 - sources for impact analysis reports 47
 - sources for lineage reports 47
- assets
 - BI reports 21
 - business intelligence reports 21
 - category 22
 - data file 23
 - data file structure 24
 - database 26
 - database table 27
 - deleting 12
 - displaying details 12
 - editing details 12
 - engine 33
 - host 25
 - job 30
 - schema 28
 - stage 31
 - steward 33
 - transformation project 32
 - view 29
- assets, performing tasks on
 - assets, deleting 12
 - assets, details 12
 - assets, editing details 12
 - business lineage report, creating 12
 - business lineage report, excluding asset in 12
 - business lineage report, including asset in 12
 - business name, editing 12
 - custom attributes, assigning 12
 - data lineage report, creating 12
 - graph view 12
 - impact analysis 12
 - notes, adding 12
 - notes, deleting 12
 - notes, editing 12
 - steward, assigning 12

B

- BI reports 21
- business intelligence reports 21
- business lineage report
 - creating 12
 - excluding asset from 12
 - including asset in 12
- business lineage reports 45
- business name
 - editing 12

C

- category 22
- custom attributes
 - assigning 12
- customer support
 - contacting 53

D

- data file 23
- data file structure 24
- data lineage report
 - creating 12
- data lineage reports 45
- data lineage, extended 8
- database 26
- database table 27

E

- engine 33
- extended data lineage 8

F

- finding assets
 - by creating query 34
 - by existing query 34
 - by short description 34
- data servers and databases 34
- particular type 34
- projects and jobs 34

G

- graph view 12

H

- host 25

I

- impact analysis 12
- impact analysis reports 45
- information assets
 - display context 11
 - information page 11
- InfoSphere Metadata Workbench
 - Administrator role 5
 - User role 5

J

- job 30

L

- legal notices 59
- lineage
 - running business lineage reports 50
 - running data reports 48

M

- metadata workbench 36
 - accessing metadata workbench 3
 - navigation tips 7
 - opening 3
 - overview 1
- metadata workbench reports
 - asset types as sources 47
 - business lineage 45, 50
 - data lineage 45, 48
 - impact analysis 45

N

- navigation tips for metadata workbench 7
- notes
 - adding 12
 - deleting 12
 - editing 12

P

- product accessibility
 - accessibility 57
- product documentation
 - accessing 55

Q

- queries 36
 - creating 39
 - editing 41
 - limiting results 39
 - publishing 39
 - running 42

R

- reports
 - BI reports 21
 - business intelligence reports 21
- roles
 - InfoSphere Metadata Workbench Administrator 5
 - Metadata Workbench User 5

S

- schema 28

- software services
 - contacting 53
- stage 31
- steward 33
 - assigning 12
- support
 - customer 53

T

- term 23
- trademarks
 - list of 59
- transformation project 32

V

- view 29



Printed in USA

SC19-3476-00

