



Informatica[®] Data Integration Hub
10.5

Getting Started

Informatica Data Integration Hub Getting Started

10.5

March 2021

© Copyright Informatica LLC 2014, 2021

This software and documentation are provided only under a separate license agreement containing restrictions on use and disclosure. No part of this document may be reproduced or transmitted in any form, by any means (electronic, photocopying, recording or otherwise) without prior consent of Informatica LLC.

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation is subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License.

Informatica, the Informatica logo, Informatica Cloud, PowerCenter, PowerExchange, and Data Engineering Integration are trademarks or registered trademarks of Informatica LLC in the United States and many jurisdictions throughout the world. A current list of Informatica trademarks is available on the web at <https://www.informatica.com/trademarks.html>. Other company and product names may be trade names or trademarks of their respective owners.

Portions of this software and/or documentation are subject to copyright held by third parties. Required third party notices are included with the product.

See patents at <https://www.informatica.com/legal/patents.html>.

DISCLAIMER: Informatica LLC provides this documentation "as is" without warranty of any kind, either express or implied, including, but not limited to, the implied warranties of noninfringement, merchantability, or use for a particular purpose. Informatica LLC does not warrant that this software or documentation is error free. The information provided in this software or documentation may include technical inaccuracies or typographical errors. The information in this software and documentation is subject to change at any time without notice.

NOTICES

This Informatica product (the "Software") includes certain drivers (the "DataDirect Drivers") from DataDirect Technologies, an operating company of Progress Software Corporation ("DataDirect") which are subject to the following terms and conditions:

1. THE DATADIRECT DRIVERS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT.
2. IN NO EVENT WILL DATADIRECT OR ITS THIRD PARTY SUPPLIERS BE LIABLE TO THE END-USER CUSTOMER FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, CONSEQUENTIAL OR OTHER DAMAGES ARISING OUT OF THE USE OF THE ODBC DRIVERS, WHETHER OR NOT INFORMED OF THE POSSIBILITIES OF DAMAGES IN ADVANCE. THESE LIMITATIONS APPLY TO ALL CAUSES OF ACTION, INCLUDING, WITHOUT LIMITATION, BREACH OF CONTRACT, BREACH OF WARRANTY, NEGLIGENCE, STRICT LIABILITY, MISREPRESENTATION AND OTHER TORTS.

The information in this documentation is subject to change without notice. If you find any problems in this documentation, report them to us at infa_documentation@informatica.com.

Informatica products are warranted according to the terms and conditions of the agreements under which they are provided. INFORMATICA PROVIDES THE INFORMATION IN THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT.

Publication Date: 2021-09-20

Table of Contents

Preface	8
Informatica Resources.	8
Informatica Network.	8
Informatica Knowledge Base.	8
Informatica Documentation.	8
Informatica Product Availability Matrices.	9
Informatica Velocity.	9
Informatica Marketplace.	9
Informatica Global Customer Support.	9
Chapter 1: Introduction to Data Integration Hub	10
Data Integration Hub Overview.	10
Data Integration Hub Big Data.	12
Applications Overview.	13
Topics Overview.	13
Data Integration Hub Publications and Subscriptions.	13
Publications Overview.	14
Subscriptions Overview.	14
Operator User Role.	14
Operation Console.	15
Changing the Operation Console Language.	15
Data Integration Hub Overview Diagram.	15
Chapter 2: Getting Started with Data Integration Hub	17
Getting Started Overview.	17
Step 1. Log In to the Data Integration Hub Operation Console.	18
Step 2. Create a Data Integration Hub Application.	18
Chapter 3: Creating Topics	19
Creating Topics Overview.	19
Creating a Topic with a Relational Database Publication Repository.	21
Step 1. Access the New Relational Database Topic Wizard.	21
Step 2. Define Basic Topic Properties.	21
Step 3. Create Topic Data Structure.	22
Step 4. Define Topic Table Relations.	26
Step 5. Define Data Retention Period and Storage Location.	27
Step 6. Assign Topic Permissions.	27
Step 7. Review Topic Settings and Save the Topic.	27
Creating a Topic with a Hadoop Publication Repository.	28
Step 1. Access the New Big Data Topic Wizard.	28

Step 2. Define Basic Topic Properties.	28
Step 3. Create Topic Data Structure.	29
Step 4. Define Topic Table Relations.	33
Step 5. Define Data Retention Period.	33
Step 6. Assign Topic Permissions.	34
Step 7. Review Topic Settings and Save the Topic.	34
Creating a Topic with a File Store Publication Repository.	34
Step 1. Access the New File Store Topic Wizard.	34
Step 2. Define Basic Topic Properties.	35
Step 3. Create Topic Data Structure.	35
Step 4. Define Topic Table Relations.	36
Step 5. Define Data Retention Period and Storage Location.	36
Step 6. Assign Topic Permissions.	37
Step 7. Review Topic Settings and Save the Topic.	37
Creating a Topic with a Real-time Publication Repository.	37
Chapter 4: Creating Publications.	39
Creating Publications Overview.	39
Creating an Automatic Relational Database Publication.	40
Step 1. Access the Create Automatic Database Publication Wizard.	40
Step 2. Define Basic Publication Properties.	40
Step 3. Select Publication Source.	41
Step 4. View and Edit Publication Field Mapping.	41
Step 5. Define a Filter.	43
Step 6. Define Publication Schedule.	44
Step 7. Assign Publication Permissions.	44
Step 8. Review Publication Settings and Save the Publication.	45
Creating an Automatic Flat File Publication.	45
Step 1. Access the Create Automatic File Publication Wizard.	45
Step 2. Define Basic Publication Properties.	46
Step 3. Select and Configure Publication Source.	46
Step 4. Define Joins.	48
Step 5. View and Edit Publication Field Mapping.	49
Step 6. Define a Filter.	50
Step 7. Define Publication Schedule.	50
Step 8. Assign Publication Permissions.	51
Step 9. Review Publication Settings and Save the Publication.	52
Creating an Automatic Pass-through File Publication.	52
Step 1. Access the Create Automatic Pass-through File Publication Wizard.	52
Step 2. Define Basic Publication Properties.	53
Step 3. Configure Publication Source.	53
Step 4. Define Publication Schedule.	53
Step 5. Assign Publication Permissions.	54

Step 6. Review Publication Settings and Save the Publication.	55
Creating an Automatic Data-driven Publication.	55
Step 1. Access the Create Automatic Data-driven Publication page.	55
Step 2. Define Publication Properties.	56
Step 3. Create a Request to Run the Publication.	56
Creating a Custom Batch Publication.	57
Step 1. Access the Create Custom Batch Publication Wizard.	57
Step 2. Define Basic Publication Properties.	58
Step 3. Select Publication Workflow.	58
Step 4. Define Publication Schedule.	58
Step 5. Assign Publication Permissions.	59
Step 6. Review Publication Settings and Save the Publication.	59
Creating a Custom Multi-latency Publication.	60
Step 1. Access the Create Custom Multi-latency Publication Wizard.	60
Step 2. Define Basic Publication Properties.	61
Step 3. Select Publication Workflow.	61
Step 4. Define Publication Intervals.	61
Step 5. Assign Publication Permissions.	61
Step 6. Review Publication Settings and Save the Publication.	61
Creating a Custom Cloud Publication.	62
Step 1. Access the Custom Cloud Publication Wizard.	62
Step 2. Define Basic Publication Properties.	62
Step 3. Select Publication Mapping.	63
Step 4. Define Publication Schedule.	63
Step 5. Assign Publication Permissions.	64
Step 6. Review Publication Settings and Save the Publication.	64
Creating a Custom Pass-through Kafka Publication.	64
Step 1: Access the Create Custom Pass-Through Apache Kafka Publication Page.	65
Step 2. Define the Pass-through Kafka Publication Properties.	65
Creating a Modular Cloud Publication.	65
1. Access the Create Modular Cloud Publication Wizard.	66
Step 2. Define Basic Publication Properties.	66
Step 3. Select Publication Mapping.	66
Step 4. Configure Publication Source.	67
Step 5. Configure Publication Target.	67
Step 6. Set Publication Parameters.	67
Step 7. Configure Publication Field Mapping.	67
Step 8. Define Publication Schedule.	68
Step 9. Assign Publication Permissions.	68
Step 10. Review Publication Settings and Save the Publication.	69
Chapter 5: Creating Subscriptions.	70
Creating Subscriptions Overview.	70

Creating an Automatic Relational Database Subscription.	71
Step 1. Access the Create Automatic Database Subscription Wizard.	71
Step 2. Define Basic Subscription Properties and Select a Topic.	72
Step 3. Select Post-process Workflow.	72
Step 4. Define Joins.	72
Step 5. Select Subscription Target.	73
Step 6. View and Edit Subscription Field Mapping.	73
Step 7. Define a Filter.	74
Step 8. Define Subscription Schedule.	75
Step 9. Define Delivery Options.	76
Step 10. Assign Subscription Permissions.	76
Step 10. Review Subscription Settings and Save the Subscription.	77
Creating an Automatic Flat File Subscription.	77
Step 1. Access the Create Automatic Flat File Subscription Wizard.	78
Step 2. Define Basic Subscription Properties and Select a Topic.	78
Step 3. Select Post-process Workflow.	78
Step 4. Define Joins.	78
Step 5. Select and Configure Subscription Target.	79
Step 6. View and Edit Subscription Field Mapping.	80
Step 7. Define a Filter.	81
Step 8. Define Subscription Schedule.	82
Step 9. Define Delivery Options.	83
Step 10. Assign Subscription Permissions.	83
Step 10. Review Subscription Settings and Save the Subscription.	83
Creating an Automatic Pass-through File Subscription.	84
Step 1. Access the Create Automatic Pass-through File Subscription Wizard.	84
Step 2. Define Basic Subscription Properties and Select a Topic.	84
Step 3. Select Post-process Workflow.	85
Step 4. Configure Subscription Target.	85
Step 5. Define a Filter.	86
Step 6. Define Subscription Schedule.	87
Step 7. Define Delivery Scope.	87
Step 8. Assign Subscription Permissions.	87
Step 9. Review Subscription Settings and Save the Subscription.	88
Creating an Automatic Data-driven Subscription.	88
Step 1. Access the Create Automatic Data-driven Subscription Page.	89
Step 2. Define Subscription Properties.	89
Step 3. Create a Request to Run the Subscription.	89
Creating a Custom Batch Subscription.	90
Step 1. Access the Create Custom Batch Subscription Wizard.	90
Step 2. Define Basic Subscription Properties and Select Topics.	91
Step 3. Select Subscription Workflows.	91

Step 4. Define Subscription Schedule.	91
Step 5. Assign Subscription Permissions.	92
Step 6. Review Subscription Settings and Save the Subscription.	92
Creating a Custom Cloud Subscription.	93
Step 1. Access the Custom Cloud Subscription Wizard.	93
Step 2. Define Basic Subscription Properties and Select Topics.	93
Step 3. Select Subscription Mappings.	94
Step 4. Define Subscription Schedule.	94
Step 5. Assign Subscription Permissions.	95
Step 6. Review Subscription Settings and Save the Subscription.	95
Creating a Custom Pass-through Kafka Subscription.	95
Step1. Access the Create Custom Pass-through Kafka Subscription Page.	96
Step2. Define Basic Subscription Properties.	96
Creating a Modular Cloud Subscription.	96
Step 1. Access the Modular Cloud Subscription Wizard.	97
Step 2. Define Basic Subscription Properties and Select a Topic.	97
Step 3. Select Subscription Mapping.	97
Step 4. Configure Subscription Source.	98
Step 5. Configure Subscription Target.	98
Step 6. Set Subscription Parameters.	98
Step 7. Configure Subscription Field Mapping.	98
Step 8. Define Subscription Schedule.	99
Step 9. Assign Subscription Permissions.	99
Step 10. Review Subscription Settings and Save the Subscription.	100
Appendix A: Glossary.	101
Index.	104

Preface

Complete the lessons in the *Data Integration Hub Getting Started Guide* to learn the basics of the Data Integration Hub Operation Console. The tutorial teaches you how to log in to the Data Integration Hub Operation Console and how to create applications, topics, publications, and subscriptions.

Informatica Resources

Informatica provides you with a range of product resources through the Informatica Network and other online portals. Use the resources to get the most from your Informatica products and solutions and to learn from other Informatica users and subject matter experts.

Informatica Network

The Informatica Network is the gateway to many resources, including the Informatica Knowledge Base and Informatica Global Customer Support. To enter the Informatica Network, visit <https://network.informatica.com>.

As an Informatica Network member, you have the following options:

- Search the Knowledge Base for product resources.
- View product availability information.
- Create and review your support cases.
- Find your local Informatica User Group Network and collaborate with your peers.

Informatica Knowledge Base

Use the Informatica Knowledge Base to find product resources such as how-to articles, best practices, video tutorials, and answers to frequently asked questions.

To search the Knowledge Base, visit <https://search.informatica.com>. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team at KB_Feedback@informatica.com.

Informatica Documentation

Use the Informatica Documentation Portal to explore an extensive library of documentation for current and recent product releases. To explore the Documentation Portal, visit <https://docs.informatica.com>.

If you have questions, comments, or ideas about the product documentation, contact the Informatica Documentation team at infa_documentation@informatica.com.

Informatica Product Availability Matrices

Product Availability Matrices (PAMs) indicate the versions of the operating systems, databases, and types of data sources and targets that a product release supports. You can browse the Informatica PAMs at <https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Informatica Velocity

Informatica Velocity is a collection of tips and best practices developed by Informatica Professional Services and based on real-world experiences from hundreds of data management projects. Informatica Velocity represents the collective knowledge of Informatica consultants who work with organizations around the world to plan, develop, deploy, and maintain successful data management solutions.

You can find Informatica Velocity resources at <http://velocity.informatica.com>. If you have questions, comments, or ideas about Informatica Velocity, contact Informatica Professional Services at ips@informatica.com.

Informatica Marketplace

The Informatica Marketplace is a forum where you can find solutions that extend and enhance your Informatica implementations. Leverage any of the hundreds of solutions from Informatica developers and partners on the Marketplace to improve your productivity and speed up time to implementation on your projects. You can find the Informatica Marketplace at <https://marketplace.informatica.com>.

Informatica Global Customer Support

You can contact a Global Support Center by telephone or through the Informatica Network.

To find your local Informatica Global Customer Support telephone number, visit the Informatica website at the following link:

<https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>.

To find online support resources on the Informatica Network, visit <https://network.informatica.com> and select the eSupport option.

CHAPTER 1

Introduction to Data Integration Hub

This chapter includes the following topics:

- [Data Integration Hub Overview, 10](#)
- [Data Integration Hub Big Data, 12](#)
- [Applications Overview, 13](#)
- [Topics Overview, 13](#)
- [Data Integration Hub Publications and Subscriptions, 13](#)
- [Operator User Role, 14](#)
- [Operation Console, 15](#)
- [Data Integration Hub Overview Diagram, 15](#)

Data Integration Hub Overview

Data Integration Hub is an application integration solution that your organization can use to share and synchronize data between different applications in the organization.

To publish data to Data Integration Hub, first define the data set that you want to manage, for example, sales, customers, or orders. You define a data set by defining a topic. A topic defines the structure of the data that Data Integration Hub stores in the publication repository and the type of publication repository where data is stored. You can manage multiple topics that represent different data sets in Data Integration Hub. Applications publish data to topics and subscribe to data sets that are represented by topics.

Multiple applications can publish to the same topic, for example, different ordering applications can publish their orders to the same Orders topic. Multiple subscribers can consume the data from a topic. Different subscribing applications can consume the data in different formats and in different latencies based on a defined schedule.

Data Integration Hub stores the data that applications publish to topics in the Data Integration Hub publication repository. Data Integration Hub keeps the data in the publication repository until all subscribers consume the data and the retention period expires, and then deletes the data from the publication repository.

Applications can use PowerExchange® adapters and Informatica Intelligent Cloud Services™ connectors to share data from different sources, such as database tables, files, or any sources that Informatica supports. Each application can be a publisher and a subscriber to different topics.

Publications publish to a specific topic. A publication defines the data source type and the location from where Data Integration Hub retrieves the data that the application publishes. Subscriptions subscribe to one or more topics. A subscription defines the data target type and the location in the subscribing application to where Data Integration Hub sends the published data.

When you create a publication or a subscription, you can choose to use either an automatic Data Integration Hub mapping or a custom Data Integration Hub mapping. Data Integration Hub creates automatic mappings based on the data structure that you define in the topic. Custom Data Integration Hub mappings are based on PowerCenter® workflows, Data Engineering Integration mappings, or Data Integration tasks that the developer creates and maintains for the publication or the subscription.

Data Integration Hub operator uses Enterprise Data Catalog to discover and leverage existing Data Integration Hub objects, and understand their lineage and impact on other entities in the enterprise.

Examples

You run a data center for a major retail chain. The main office has multiple applications. Some of the applications are located on-premises and some are located on the cloud. Each retail branch has a point-of-sale (POS) application and an inventory application. Your applications and branches require the following data:

Customer service applications

Require up-to-date customer order data.

Sales applications

Require up-to-date product sales data.

Marketing application

Requires a weekly deals report.

Accounting application

Requires a monthly deals report.

Branch applications

Require up-to-date inventory and pricing data.

Business Intelligence (BI) application

Requires a weekly report of sales and marketing data and of user interaction data from the corporate website, for the preceding 12 months.

With Data Integration Hub, you can address the following use-cases:

Share product catalog and prices.

You can share product price updates from the sales department with each branch, as follows:

1. Create a Products topic.
2. For the Product Information Management (PIM) application, define a publication that publishes product details and prices to the Products topic and set the schedule to publish the data daily.
3. For each branch application, define a subscription to the Products topic and set the subscription to consume the published data when it is available in Data Integration Hub.

Share daily sales details.

You can share the daily sales details that you receive from the stores with your central sales application and your customer service applications, as follows:

1. Create a Sales topic.

2. For each branch application, define a publication to the Sales topic, and set the schedule to publish daily.
3. For the sales application, define a subscription to the Sales topic, and set the schedule to consume the data when it is published.
4. For the customer service application, define a subscription to the Sales topic, and set the schedule to consume the data once a week.

Share deal details from Salesforce.

You can share deal details from a Salesforce cloud application with the marketing and accounting applications, as follows:

1. Create a Deals topic.
2. For the Salesforce application, define a cloud publication to the Deals topic, and set the schedule to publish weekly.
3. For the marketing application, define a subscription to the Deals topic, and set the schedule to consume the data once a week.
4. For the accounting application, define a subscription to the Deals topic, and set the schedule to consume the data once a month.

Share business intelligence data.

You can share sales data from Salesforce, marketing data from the marketing application, and user interaction data from the corporate website with the BI application as follows:

1. Create a Business Intelligence topic with a Hadoop publication repository and set the publication data retention period to 365 days.
2. For the Salesforce application, define a cloud publication to the Business Intelligence topic, and set the schedule to publish once a week.
3. For the marketing application, define a publication to the Business Intelligence topic, and set the schedule to publish once a week.
4. For the corporate website application, define a cloud publication to the Business Intelligence topic, and set the schedule to publish once a week.
5. For the BI application, define an aggregated subscription to the Business Intelligence topic, and set the schedule to consume the data once a week.

Data Integration Hub Big Data

Publish and subscribe to high volumes of data, data streams, and data that you want to store for a long period of time with Data Integration Hub. For example, store business intelligence data that you need to review over time on the Data Integration Hub Hadoop publication repository, or publish from and subscribe to Hadoop Distributed File System (HDFS) and Hive data warehouses.

If you want to keep the published data in the Hadoop publication repository after the data is consumed by all subscribers, you can configure Data Integration Hub not to delete published data from the repository.

You can use both automatic mappings and custom mappings to publish and consume big data with Data Integration Hub. For custom mapping publications you can use Informatica Data Engineering Integration mappings and workflows and Informatica Data Engineering Streaming mappings. For custom mapping subscriptions you use Informatica Data Engineering Integration mappings and workflows.

Applications Overview

An application represents an entity in your organization that needs to share data with other applications in your organization, such as sales applications or customer service applications. In Data Integration Hub, an application is a container for publications and subscriptions.

Applications can reside either inside or outside the Data Integration Hub network, and can be located on-premises or on the cloud. To share file-based data between applications that reside outside the Data Integration Hub network, you can use the SSH File Transfer Protocol (SFTP).

An application can publish data to a defined topic and can subscribe to data from a topic. For example, a sales application can publish sales reports and subscribe to inventory updates from an operations application. When you add a publication to an application, you define the schedule according to which topic data will be published from the application. You also define the schedule according to which topic data will be retrieved from the application and published to the Data Integration Hub publication repository. When you add a subscription to an application, you define the topic to which the application subscribes and the schedule and scope of data that the application consumes from the topic. The topic defines the structure of the data that the associated publications and subscriptions publish and consume.

You can use permissions to restrict access to applications to specific users. Only users with permissions to the application can define publications and subscriptions for this application.

Topics Overview

A topic is an entity that represents a data domain that is published and consumed in Data Integration Hub. A topic defines the data structure and additional data definitions, such as the data retention period. Multiple applications can publish to the same topic. An application can subscribe to multiple topics.

For example, you can create a Sales topic that represents sales data. Applications from all the stores in the organization publish sales data to the Sales topic. The accounting application subscribes to the Sales topic and consumes published sales data from all stores, or, if a filter is applied, from specific stores.

Data Integration Hub Publications and Subscriptions

Publications and subscriptions are entities that define how applications publish data to Data Integration Hub and how applications consume data from Data Integration Hub. Publications publish data to a defined topic and subscriptions subscribe to topics.

Publications and subscriptions control the data flow and the schedule of data publication or data consumption. An application can be a publisher and a subscriber. Multiple applications can publish to the same topic. Multiple applications can consume data from the same topic.

You can use automatic, custom, and modular publications and subscriptions to publish data and to consume data. You can publish from and subscribe to different sources of data. Because the publishing process and the consuming process are completely decoupled, the publishing source and the consuming target do not have to be of the same data type. For example, you can publish data from a file and consume it into a database.

Automatic publications and subscriptions can publish from and subscribe to a relational database, a file, or a cloud application, or over a REST API.

Custom publications and subscriptions can publish from and subscribe to on-premises applications.

Modular publications and subscriptions can publish from and subscribe to cloud applications.

Publications Overview

Publications are entities that define how applications publish data to Data Integration Hub. Publication definitions can include the type, format, and schedule of data publication. Publications publish data to topics. Multiple publications can publish to the same topic. The topic defines the structure to which the data is published.

You create and edit publications in the Data Integration Hub Operation Console. During the configuration you choose the topic to which to publish the data. The steps that you take when you configure a publication depend on the publication type.

Subscriptions Overview

Subscriptions are entities that define how applications consume data from Data Integration Hub. Subscriptions subscribe to topics. A subscription can subscribe to multiple topics. Multiple subscriptions can consume data from the same topic.

When you create a subscription, choose the topic or topics to which the application subscribes. If applicable, define the schedule and the delivery scope of the data to consume and the delivery behavior for the published data, for example, to aggregate all data sets to a single data set, or to consume the latest published data set.

In the Data Integration Hub Operation Console, you use the subscription wizard or page to create and edit subscriptions. During the configuration you choose the topic to which to subscribe. The steps that you take when you configure a subscription depend on the subscription type and on the type of mapping that the subscription uses.

Operator User Role

The Data Integration Hub operator is responsible for managing operational entities in the Operation Console. The entities include applications, topics, publications, and subscriptions. The operator also monitors publication and subscription processing status with events and event monitoring rules.

The operator creates and modifies applications that need to publish or consume data. For each application, the operator creates and defines publications or subscriptions. The operator defined the data domains that Data Integration Hub manages by creating and modifying topics.

The operator configures and maintains source connections for publications and target connections for subscriptions. The operator can define connections when the operator creates a publication or a subscription, or manage connections directly in the Operation Console.

While publications and subscriptions are running, the operator monitors the status of the publication and subscription process with events and monitors, and in the Data Integration Hub Dashboard. If errors occur during the publication or the subscription process, the operator can perform actions on the events to troubleshoot issues.

The operator can create rules that monitor publication and subscription events, and perform actions on events that are in a defined status. For example, you can create a rule that disables publications whose events are in an Error status, a rule that invokes a PowerCenter workflow when a publication event is in an Error status, or a rule that sends an email to the Data Integration Hub administrator when a subscription event is in a Critical status.

Operation Console

Use the Operation Console user interface to manage applications, topics, publications, and subscriptions, and to monitor publications, subscriptions, and events. Administrators also use the Operation Console to manage users and system settings. Developers use the Operation Console to manage workflows and connections.

You can view the Operation Console in English or in Japanese. You can switch between the display languages.

The Operation Console contains two areas:

Navigator

Use the navigator to navigate between tasks that you can perform in the Operation Console. The navigator shows in the left pane of the Operation Console.

Current page

Main work area in which you perform the tasks that you select in the Navigator. The current page shows in the right pane of the Operation Console.

Changing the Operation Console Language

You can view the Operation Console in English or in Japanese. You can switch between the display languages.

1. In the browser from where you access Data Integration Hub, set the language to the required language.
2. The **Help** link opens the online help in English. To view the Japanese online help access the following URL:

```
http(s)://<host>:<port>/dih-help-ja
```

Where:

- <host> is the host name or the IP address of the Data Integration Hub server.
- <port> is the port number of the Data Integration Hub server.

For example:

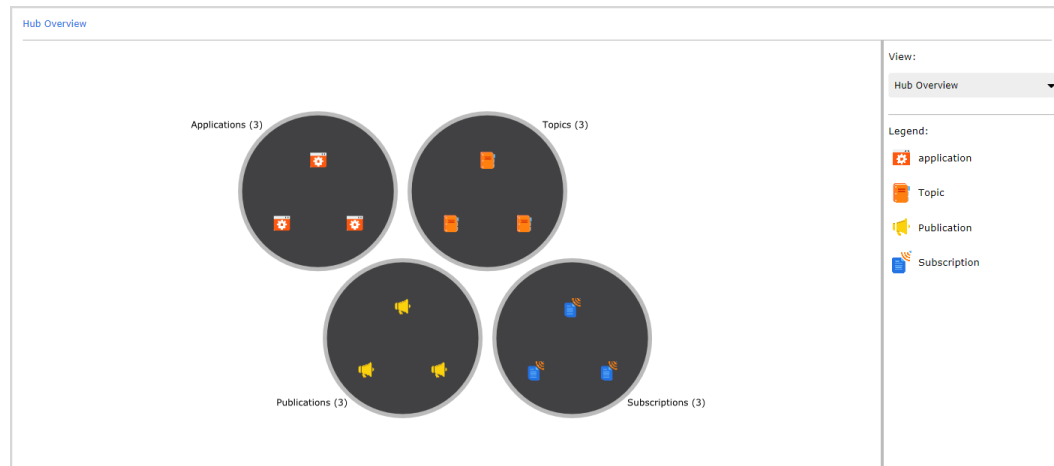
```
https://dih-releases:19443/dih-help-ja/
```

Data Integration Hub Overview Diagram

The Data Integration Hub Operation Console Home page displays the Hub Overview diagram when Data Integration Hub contains entities such as applications, topics, publications, or subscriptions.

The Hub Overview diagram provides a visual overview of the existing entities, grouped into categories.

The following image shows a sample Hub Overview diagram:



When you hover over an entity in the diagram, all related entities are highlighted. For example, when you hover over a topic, the applications and the publications that publish to the topic and the subscriptions that subscribe to the topic are highlighted. When you click an entity, a drill down view of the entity and its relations to other entities appears. For example, when you click a publication, the drill down view shows the publishing application, the topic to which the publication publishes data, and the subscriptions that subscribe to the topic. Click any entity in the drill down view. The edit window of the selected entity displays.

You can filter the Hub Overview diagram to the following views:

- Hub Overview. All entities including topics, related applications, publications, and subscriptions that are configured in Data Integration Hub.
- Process errors. Entities with current error events.
- Non-valid entities. Entities that are not valid.
- Topics with no publications. Topics with no associated publications.
- Topics with no subscriptions. Topics with no associated subscriptions.
- Most used topics. Three most used topics, based on the number of publications and subscriptions that use the topic.

When you filter the diagram, entities that are not relevant to the selected filter appear in view only mode.

To search topics in the Hub Overview diagram, click the **Find Topic** button and enter the name of the topic. Topics that match the search criteria and its related entities are displayed in Hub Overview diagram.

CHAPTER 2

Getting Started with Data Integration Hub

This chapter includes the following topics:

- [Getting Started Overview, 17](#)
- [Step 1. Log In to the Data Integration Hub Operation Console, 18](#)
- [Step 2. Create a Data Integration Hub Application, 18](#)

Getting Started Overview

In this chapter, you start the Data Integration Hub Operation Console and create a Data Integration Hub application.

Chapter Concepts

The Data Integration Hub Operation Console is a web client user interface where you manage applications, topics, publications, and subscriptions, and monitor events.

An application represents an entity in your organization that needs to share data with other entities. A Data Integration Hub application is a container for publications and for subscriptions. To publish data and to subscribe to published data, you must first create an application that represents the entity that provides the data that you want to publish and to consume.

Chapter Objectives

In this chapter, you perform the following tasks:

- Log in to the Data Integration Hub Operation Console.
- Create a Data Integration Hub application.

Step 1. Log In to the Data Integration Hub Operation Console

Start the Data Integration Hub Operation Console to begin this task.

1. From the Windows desktop, click **Start > All Programs > Informatica Data Integration Hub > Operation Console**.

The default browser opens to the Data Integration Hub Operation Console.

2. Enter the user name and password provided to you.
3. Click **Log In**.

The Data Integration Hub Overview page appears. If no entities exist in Data Integration Hub, the Overview page displays a Welcome page. If Data Integration Hub contains entities such as applications, topics, publications, or subscriptions, the Overview page displays the Data Integration Hub Overview diagram. The Overview diagram provides a visual overview of the existing entities, grouped into categories.

Note: The **B2B Gateway** link appears on the **My Services** page if your organization has the required licences. If the link doesn't appear on the **My Services** page, contact your administrator.

Step 2. Create a Data Integration Hub Application

Create an application in the Data Integration Hub Operation Console. An application can be a publisher, a subscriber, or both a publisher and a subscriber. In this task, create at least one application.

1. In the Navigator, click **Hub Management > Applications**.

The **Applications** page appears.

2. Click **New Application**.

The **Create Application** page appears.

3. In the **General** tab, enter the application name.
4. Optionally, enter a description of the application.
5. Click **Save**.

The **Create Application** page closes. The **Applications** page shows the application in the application list.

CHAPTER 3

Creating Topics

This chapter includes the following topics:

- [Creating Topics Overview, 19](#)
- [Creating a Topic with a Relational Database Publication Repository, 21](#)
- [Creating a Topic with a Hadoop Publication Repository, 28](#)
- [Creating a Topic with a File Store Publication Repository, 34](#)
- [Creating a Topic with a Real-time Publication Repository, 37](#)

Creating Topics Overview

In this section, you create topics to which applications publish data and from which applications consume data. You must have first completed the chapter "Getting Started with Data Integration Hub."

When you create a topic, you choose the topic type and the type of repository on which to store data for the topic, define the data structure and the data retention period, select a data storage location, and assign topic permissions.

Chapter Concepts

Data Integration Hub can manage and store topic data on the following types of publication repository:

- **Relational database.** Choose this type of repository to store published data in a relational database structure that represents the structure in which you want to keep the data. For example, data that is published from a relational database or from files. A relational database publication repository usually stores the data for a short intermediate period after the data is consumed by all subscribers. Data Integration Hub supports the following databases on which to store relational database topic data: Oracle, Microsoft SQL Server.
- **Big Data.** Choose this type of repository if you publish high volumes of data that you want to store for a long period of time or if you do not want Data Integration Hub to delete published data after the data is consumed. The availability of the Hadoop repository depends on whether or not the Hadoop component is installed on your system.

To publish and subscribe to a Hadoop-based repository with custom publications and subscriptions, you must use workflows that are based on a Data Engineering Integration mapping and workflow. When you create a custom publication, if one of the topics that you select for the publication is a Hadoop-based topic, only workflows that are based on a Data Engineering Integration mapping or workflow are listed for selection as the publication mapping.

When you create a compound subscription, that subscribes to multiple topics, all topics that you select must be Hadoop-based, and only workflows that are based on a Data Engineering Integration mapping or

workflow are listed for selection as the subscription mapping. You also enable the mandatory option for topics in compound subscription to prioritize a few topics in the compound subscription over other topics.

Data Integration Hub triggers a processing of subscription after the publication event for all topics are completed. If the wait time of the publication event is complete and Data Integration Hub has not published all mandatory topics, an error event is generated during run-time.

Before you use a Hadoop-based publication repository for publications and subscriptions, consider the following restrictions:

- You cannot assign a pre-process to a custom publication that publishes to a Hadoop-based repository.
- You cannot configure a custom publication that publishes files to a Hadoop-based repository to run immediately when the files are ready to be published.
- You cannot use a Hadoop repository to publish and subscribe to pass-through files and Hadoop Distributed File System (HDFS) files.
- **File Store.** Choose this type of repository to publish files that you want to keep as-is without loading the data into a relational database. For example, if you publish PDF or .zip files into a file repository, Data Integration Hub delivers the files without processing them.
- **Real-time.** Choose this type of repository to monitor real-time Apache Kafka data streaming. Apache Kafka is a distributed streaming platform that can publish and subscribe to stream of records, store and process streams of record. In order to track the Apache Kafka flows, you must configure the Apache Kafka server URL in the System property of the Data Integration Hub. You must then create a topic with the publication repository type of Real-time and create an application to define the publisher and subscriber. Also, create a workflow that maps to the Apache Kafka server. The publication and subscription of the Data Integration hub associated with the source and target of the Kafka server.

The Data Integration Hub records streaming of data in the Apache Kafka server at regular intervals. The Data Integration Hub operator configures the interval at which Data Integration Hub must record the data streaming value in the topic. The Events List stores the log of events. The Processing Information tab in the Events List, stores the Offset and LogEndOffset values that define the difference between data values at intervals in each partition.

When you create the structure of a topic, you define the data structure on the Data Integration Hub publication repository to where the publications that are associated with the topic publish data, and from where subscribers to the topic consume the data. The topic structure must contain at least one table and can consist of multiple tables.

The data retention period defines how long Data Integration Hub retains the data in the publication repository after the data is consumed.

Topic permissions control who can access the topic. The Data Integration Hub administrator creates categories and assigns categories to user groups to determine the users that can view or change topics. You assign categories to a topic to permit users to view or change the topic. Because publications and subscriptions are associated with topics, they inherit the permissions from the associated topic. When you configure permissions for a topic, only user groups with permissions to the topic can access the associated subscriptions and publications.

Chapter Objectives

In this chapter, you perform the following tasks:

- Create a topic where the published data is stored on a relational database.
- Create a topic where the published data is stored on a Hadoop repository.
- Create a topic where the published data is stored on a file repository.
- Create a topic where the published data is stored in a Apache Kafka server.

Creating a Topic with a Relational Database Publication Repository

To create a topic with a relational database publication repository, perform the following tasks:

1. Access the **New Topic** wizard.
2. Define basic topic properties.
3. Create the topic data structure.
4. Optionally, define topic table relations.
5. Define the data retention period and the storage location.
6. Optionally, assign topic permissions.
7. Review the topic settings and save the topic.

Task Prerequisites

Before you start this task, obtain the details of the topic structure that you want to create.

Step 1. Access the New Relational Database Topic Wizard

Access the **New Relational Database Topic** wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Topics**.
The **Topics** page appears.
2. Click **New Topic**, select **Relational Database**, and then click **Create**.
The **New Relational Database Topic** wizard appears.

Step 2. Define Basic Topic Properties

Define topic properties in the **General** page of the **New Topic** wizard.

1. Enter the topic name.
2. Optionally, enter a description of the topic.
3. Optionally, assign a tag to the topic so that you can search for the topic on the **Topics** page and the **Catalog** page based on the tag.
4. Choose the topic type.
 - Delta. The topic instance contains only the latest data changes. If you choose this topic type, verify that the data source includes delta indicators.
 - Full. The topic instance contains all of the data changes that occurred after the last publication. Choose this topic type if you associate publications with custom mapping with the topic.
5. Under **Publication Repository** select **Relational database**.
Note: If you select the option **Prevent new publications and new subscriptions to this topic** you cannot create publications and subscriptions that publish to and subscribe from the topic.
6. Click **Next**.
The **Structure** page appears.

Step 3. Create Topic Data Structure

Create and preview the data structure of the topic in the **Structure** page of the **New Topic** wizard.

You can use one or more of the following methods to add tables to the topic:

Add tables from a database

Use this method if the structure of the data domain that the topic represents exists in a database in your organization. For example, if the structure exists in one of the applications that publish to the topic.

Add tables from Enterprise Data Catalog

Use this method when the structure of the data domain that the topic represents exists in Enterprise Data Catalog. Select tables from an Enterprise Data Catalog asset that the Data Integration Hub topic must use.

Add a table from a flat file

Use this method when the structure of the data domain that the topic represents exists in a flat file. You can only use delimited files to add a topic table with this method.

Add tables from a PowerCenter workflow

Use this method if your organization has a PowerCenter workflow that extracts data from the source applications, and the targets of the workflow represent the structure in which you want to keep the data. The developer creates the workflow in PowerCenter Workflow Manager. Data Integration Hub generates the topic structure based on the targets in the workflow.

Add tables from a Metadata file

Use this method when the structure of a table in the data domain that the topic represents exists in a JSON, XML, XLS, or XLSX file.

Create a table

Use this method when the structure of the data domain that the topic represents does not exist in a database in your organization, in a PowerCenter workflow, or in a file, and you want to define the structure manually in Data Integration Hub.

Note: You must add at least one table to the topic structure.

1. Click **Add Tables** and then select the method by which to add the tables to the structure.

Note: Data Integration Hub adds the columns DIH_PUBLICATION_INSTANCE_DATE, DIH_PUBLICATION_INSTANCE_ID, and DIH__UPDATE_STRATEGY_FLAG to each topic table. When you finish creating the table, the columns show in the **Structure Preview** area of the **Structure** page of the topic wizard.

2. Click **Next**.

The **Table Relations** page appears.

Adding Topic Tables from a Database

Add tables to the topic structure from a database that exists in your organization.

Perform the following steps to add topic tables from a database:

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **From Database**.

The **Add Tables from Database** page appears.

2. Click the browse button next to **Database Connection**.

The **Select Metadata Connection** dialog box appears.

3. Choose a connection from the list of connections, or click **New Connection** to create a connection.
Note: To use a Teradata connection you must create the connection before you create the topic. You cannot create a new Teradata connection here.
Data Integration Hub uses the metadata access connection details to access the database schema information.
4. In the **Select Metadata Connection** dialog box click **OK**.
The **Search Results** area shows tables in the database that use the default schema.
5. Optionally, search for tables in the database. Note the following guidelines:
 - To search for tables by table name, enter a string in the **Find Topic Table** text box and then click **Search**.
 - To search for tables that use a schema other than the default schema, clear the option **Show default schema only**, enter a string in the **Schema** text box, and then click **Search**.
 - To clear the search results and show only tables that use the default schema, select the option **Show default schema only** and then click **Show All**.
 - The search is not case sensitive.
 - You can search for a substring.
6. In the **Select Topic Tables** area, select tables in the database to include in the topic structure and then click **Add Tables**.
The **Add Tables from Database** page closes. The **Structure** page shows a list of topic tables and a preview of the topic structure.

Adding Tables from Enterprise Data Catalog

You add tables from Enterprise Data Catalog when the structure of the data domain that the topic represents exists in Enterprise Data Catalog.

Perform the following steps in the **Add Tables from Enterprise Data Catalog** window to add tables from Enterprise Data Catalog:

1. Search and select a resource in the **Enterprise Data Catalog Resource** field and click **OK**.
Data Integration Hub retrieves all the objects of the asset from Enterprise Data Catalog and displays them as options in the Select Topic Tables pane.
2. **Select Topic Tables**. Select tables to include in the topic structure.
3. Click **OK**.

Note: To use tables from Enterprise Data Catalog, verify that the Data Integration Hub administrator has configured the Enterprise Data Catalog connectivity with Data Integration Hub. For more information about configuring the Enterprise Data Catalog connection, see the *Data Integration Hub Administrator Guide*.

Adding a Topic Table from a Flat File

Add a table to the topic structure from a sample flat file. You can only use delimited files to add a topic table.

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **From Flat File**.
The **Add Table from File** wizard appears, showing the **Definition** page.
2. Click the browse button next to **Sample file**.
The **Sample File** dialog box appears.
3. Click **Browse**, select the file, and then click **Upload**.

The **Sample File** dialog box closes.

4. In the **File Format** area, define the following parameters:

Code page

Character encoding used in the file.

Import column names

Optional. Select **Yes** to use the column names in the file as the default column headers in the table. Enter the number of the row that serves as the file's header row in **From row**.

Default text length

Optional. Length of the text fields in the table.

Delimiter

Delimiter used in the file to separate between columns. Select a predefined delimiter or select **Custom** to define a custom delimiter. For information about the supported column delimiters, see the sections about importing delimited flat files and about updating delimited file properties in the *PowerCenter Designer Guide*.

Text qualifier

Optional. Symbols used in the file to enclose a string.

The **Preview** area shows the columns in the table.

5. Click **Next**.

The **Structure** page of the **Add Table from File** wizard appears.

6. Review the table structure and adjust it if required. The following list describes the elements of the **Structure** page of the **Add Table from File** wizard:

Table Name

Name of the table. By default, Data Integration Hub derives the table name from the name of the file that you select in the **Definition** page. You can edit the default name.

The table name can contain only ASCII alphanumeric characters and underscores and cannot start with a numeric character.

The name must be unique in the Data Integration Hub publication repository.

Column table

Use the add, up, down, and remove buttons to add, order, and remove table columns. The table must contain at least one column.

Each column must contain the following information:

- **Column Name.** Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
If you select the option **Import column names** in the **Definition** page, Data Integration Hub populates the column names with the strings of the defined row. If you do not select the option **Import column names**, Data Integration Hub assigns default names to the columns. For example, **Field1, Field2, Fieldn**.
- **Data Type.** Select from the list of available data types. By default, Data Integration Hub reads the data as string.
- **Precision.** Enabled only for data types that support precision.
- **Scale.** Enabled only for data types that support data scaling.

Preview

This area shows the columns in the table. If you select a sample file that represents the structure of the flat file, this area shows the data in the sample file.

7. Click **Finish**.

The **Add Table from File** wizard closes. The **Structure** page of the **New Topic** wizard shows the topic table and a preview of the topic structure.

Adding Topic Tables from a PowerCenter Workflow

Add tables to the topic structure from a PowerCenter workflow.

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **From PowerCenter**.
The **Add Tables from PowerCenter Workflow** page appears.
2. Choose one of the following options to add tables to the topic structure:
 - Choose **Select a workflow from the PowerCenter repository** to select a workflow from the PowerCenter repository. Either enter a string in the **Folder name** text box and then click **Search**, or click **Show All**, and then select a workflow from the list.
 - Choose **Select a workflow definition file (.xml)** to select a workflow definition file. Browse to select the file and then click **Upload**.
3. Click **Save**.

The **Add Tables from PowerCenter Workflow** page closes. The **Structure** page shows a list of topic tables and a preview of the topic structure.

Adding Topic Tables from a Metadata File

You can load a metadata file to Data Integration Hub and create a topic table that is based on the structure of the file. Data Integration Hub supports JSON, XML, XLS, and XLXS file formats.

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **From Metadata File**.
The **Add table from Metadata File** page appears.
2. Browse to select the file and click **Upload**.
3. Enter the table name, and click **Save**.

The **Structure** page shows a list of topic tables and a preview of the topic structure.

Creating a Topic Table

Add a table to the topic structure by creating the table.

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **Create**.
The **Create Table** page appears.
2. Enter the logical table name in the **Table Name** field.
3. If required, change the physical table name.
4. Optionally, enter a description of the table.
5. To apply delta detection to the table click **Apply Delta Detection**.

Note: Data Integration Hub applies delta detection to tables for which a primary key is defined. You assign primary keys to tables in the **Table Relations** tab of the topic wizard.

6. Use the add, up, and down buttons to add and order table columns. You must add at least one column to the table.

For each column, enter the following information:

- **Column Name.** Logical name of the column. Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
- **Physical Name.** Physical name of the column. Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
- **Data Type.** Select from the list of available data types.
- **Precision.** Enabled only for data types that support precision.
- **Scale.** Enabled only for data types that support data scaling.
- **Description.** Optional description of the column.

Note: Data Integration Hub adds the columns `DIH_PUBLICATION_INSTANCE_DATE`, `DIH_PUBLICATION_INSTANCE_ID`, and `DIH__UPDATE_STRATEGY_FLAG` to each topic table. When you finish creating the table, the columns show in the **Structure Preview** area of the **Structure** page of the topic wizard.

7. Click **OK**.

The **Create Table** page closes. The **Structure** page shows the topic table and a preview of the topic structure.

Step 4. Define Topic Table Relations

You can define relations between topic tables, to filter the data that automatic database and flat file subscriptions consume from the topic in the **Table Relations** page of the **New Topic** wizard. You can also define filter accelerators for topics that you plan to use for unbound subscriptions.

Tip: In the **Show** list, you can select to show all topic tables or to show a specific table.

1. In the **Filter Accelerator** column, define a column or multiple columns as filter accelerators. A filter accelerator indicates that the column will be used in subscription queries and requires performance-related handling by Data Integration Hub. Use this indicator with topics that you plan to use for unbound subscriptions.

When you use filter accelerators, consider the impact that performance-related handling has on system performance:

- Filter accelerators speed up the handling of subscriptions that use both filters and PowerCenter pushdown optimization. By default, the PowerCenter pushdown optimization option is selected for unbound subscriptions only.
 - Filter accelerators slow down the writing of publication data to the Data Integration Hub publication repository.
 - Filter accelerators have no impact on subscriptions that do not use filters.
2. In the **Key** column, define a column from a topic table for which to define relations as a primary key.
 3. Define primary keys for as many tables as required, one primary key for each table.
 4. For each table for which you defined a primary key, define one or more foreign keys. A foreign key must relate to a primary key from a different table.
 5. Click **Next**.

The **Data Retention** page appears.

Step 5. Define Data Retention Period and Storage Location

Define the data retention period and the data storage location in the **Data Retention** page of the **New Topic** wizard.

1. Enter the number of days that Data Integration Hub retains the data in the Data Integration Hub publication database after the data is consumed in the **Publication data retention period** field.
2. Click **Advanced**.
The **Data Storage Location** area appears.
3. Choose the location in the database where Data Integration Hub stores data that is published to the topic.
 - Default. Data Integration Hub stores the data based on the default storage configuration. For example, the default File Group.
 - Custom. Browse to select an available data storage group in the publication repository database. For example, browse the available file groups.
4. Click **Next**.
The **Permissions** page appears.

Step 6. Assign Topic Permissions

Control access to the topic in the Operation Console in the **Permissions** page of the **New Topic** wizard. If you do not assign permissions, the topic is accessible by all Data Integration Hub users.

1. Select the category to which you want to assign permission to the topic under **Available Categories** and click the right arrow.
The category appears under **Selected Categories**.
2. Repeat step [1](#) to assign additional categories.
3. Click **Next**.
The **Summary** page appears.

Step 7. Review Topic Settings and Save the Topic

Review the topic settings and save the topic in the **Summary** page of the **New Topic** wizard.

1. Review the topic settings.
2. Click **Finish**.

The **New Topic** wizard closes. The **Topics** page shows the topic you created. You can create publications that publish data to the topic and subscriptions that consume data from the topic. The topic is listed in the catalog. In the catalog you can view all the data sets that are available in Data Integration Hub and verify that the data that the target applications require exists in Data Integration Hub.

Creating a Topic with a Hadoop Publication Repository

To create a topic with a Hadoop publication repository, perform the following tasks:

1. Access the **New Topic** wizard.
2. Define basic topic properties.
3. Create the topic data structure.
4. Optionally, define topic table relations.
5. Define the data retention period and the storage location.
6. Optionally, assign topic permissions.
7. Review the topic settings and save the topic.

Task Prerequisites

Before you start this task, obtain the details of the topic structure that you want to create.

Step 1. Access the New Big Data Topic Wizard

Access the **New Big Data Topic** wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Topics**.
The **Topics** page appears.
2. Click **New Topic**, select **Big Data**, and then click **Create**.
The **New Big Data Topic** wizard appears.

Step 2. Define Basic Topic Properties

Define topic properties in the **General** page of the **New Topic** wizard.

1. Enter the topic name.
2. Optionally, enter a description of the topic.
3. Optionally, assign a tag to the topic so that you can search for the topic on the **Topics** page and the **Catalog** page based on the tag.
4. Choose the topic type.
 - Delta. The topic instance contains only the latest data changes. If you choose this topic type, verify that the data source includes delta indicators.
 - Full. The topic instance contains all of the data changes that occurred after the last publication. Choose this topic type if you associate publications with custom mapping with the topic.
5. Under **Publication Repository** select **Hadoop**.
Note: If you select the option **Prevent new publications and new subscriptions to this topic** you cannot create publications and subscriptions that publish to and subscribe from the topic.
6. Click **Next**.
The **Structure** page appears.

Step 3. Create Topic Data Structure

Create and preview the data structure of the topic in the **Structure** page of the **New Topic** wizard.

You can use one or more of the following methods to add tables to the topic:

Add tables from a database

Use this method if the structure of the data domain that the topic represents exists in a database in your organization. For example, if the structure exists in one of the applications that publish to the topic.

Add tables from Enterprise Data Catalog

Use this method when the structure of the data domain that the topic represents exists in Enterprise Data Catalog. Select tables from an Enterprise Data Catalog asset that the Data Integration Hub topic must use.

Add a table from a flat file

Use this method when the structure of the data domain that the topic represents exists in a flat file. You can only use delimited files to add a topic table with this method.

Add tables from a PowerCenter workflow

Use this method if your organization has a PowerCenter workflow that extracts data from the source applications, and the targets of the workflow represent the structure in which you want to keep the data. The developer creates the workflow in PowerCenter Workflow Manager. Data Integration Hub generates the topic structure based on the targets in the workflow.

Add tables from a Metadata file

Use this method when the structure of a table in the data domain that the topic represents exists in a JSON, XML, XLS, or XLSX file.

Create a table

Use this method when the structure of the data domain that the topic represents does not exist in a database in your organization, in a PowerCenter workflow, or in a file, and you want to define the structure manually in Data Integration Hub.

Note: You must add at least one table to the topic structure.

1. Click **Add Tables** and then select the method by which to add the tables to the structure.

Note: Data Integration Hub adds the columns `DIH_PUBLICATION_INSTANCE_DATE`, `DIH_PUBLICATION_INSTANCE_ID`, and `DIH__UPDATE_STRATEGY_FLAG` to each topic table. When you finish creating the table, the columns show in the **Structure Preview** area of the **Structure** page of the topic wizard.

2. Click **Next**.

The **Table Relations** page appears.

Adding Topic Tables from a Database

Add tables to the topic structure from a database that exists in your organization.

Perform the following steps to add topic tables from a database:

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **From Database**.

The **Add Tables from Database** page appears.

2. Click the browse button next to **Database Connection**.

The **Select Metadata Connection** dialog box appears.

3. Choose a connection from the list of connections, or click **New Connection** to create a connection.

Note: To use a Teradata connection you must create the connection before you create the topic. You cannot create a new Teradata connection here.

Data Integration Hub uses the metadata access connection details to access the database schema information.
4. In the **Select Metadata Connection** dialog box click **OK**.

The **Search Results** area shows tables in the database that use the default schema.
5. Optionally, search for tables in the database. Note the following guidelines:
 - To search for tables by table name, enter a string in the **Find Topic Table** text box and then click **Search**.
 - To search for tables that use a schema other than the default schema, clear the option **Show default schema only**, enter a string in the **Schema** text box, and then click **Search**.
 - To clear the search results and show only tables that use the default schema, select the option **Show default schema only** and then click **Show All**.
 - The search is not case sensitive.
 - You can search for a substring.
6. In the **Select Topic Tables** area, select tables in the database to include in the topic structure and then click **Add Tables**.

The **Add Tables from Database** page closes. The **Structure** page shows a list of topic tables and a preview of the topic structure.

Adding a Topic Table from a Flat File

Add a table to the topic structure from a sample flat file. You can only use delimited files to add a topic table.

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **From Flat File**.

The **Add Table from File** wizard appears, showing the **Definition** page.
2. Click the browse button next to **Sample file**.

The **Sample File** dialog box appears.
3. Click **Browse**, select the file, and then click **Upload**.

The **Sample File** dialog box closes.
4. In the **File Format** area, define the following parameters:

Code page

Character encoding used in the file.

Import column names

Optional. Select **Yes** to use the column names in the file as the default column headers in the table. Enter the number of the row that serves as the file's header row in **From row**.

Default text length

Optional. Length of the text fields in the table.

Delimiter

Delimiter used in the file to separate between columns. Select a predefined delimiter or select **Custom** to define a custom delimiter. For information about the supported column delimiters, see

the sections about importing delimited flat files and about updating delimited file properties in the *PowerCenter Designer Guide*.

Text qualifier

Optional. Symbols used in the file to enclose a string.

The **Preview** area shows the columns in the table.

5. Click **Next**.

The **Structure** page of the **Add Table from File** wizard appears.

6. Review the table structure and adjust it if required. The following list describes the elements of the **Structure** page of the **Add Table from File** wizard:

Table Name

Name of the table. By default, Data Integration Hub derives the table name from the name of the file that you select in the **Definition** page. You can edit the default name.

The table name can contain only ASCII alphanumeric characters and underscores and cannot start with a numeric character.

The name must be unique in the Data Integration Hub publication repository.

Column table

Use the add, up, down, and remove buttons to add, order, and remove table columns. The table must contain at least one column.

Each column must contain the following information:

- **Column Name.** Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
If you select the option **Import column names** in the **Definition** page, Data Integration Hub populates the column names with the strings of the defined row. If you do not select the option **Import column names**, Data Integration Hub assigns default names to the columns. For example, **Field1, Field2, Fieldn**.
- **Data Type.** Select from the list of available data types. By default, Data Integration Hub reads the data as string.
- **Precision.** Enabled only for data types that support precision.
- **Scale.** Enabled only for data types that support data scaling.

Preview

This area shows the columns in the table. If you select a sample file that represents the structure of the flat file, this area shows the data in the sample file.

7. Click **Finish**.

The **Add Table from File** wizard closes. The **Structure** page of the **New Topic** wizard shows the topic table and a preview of the topic structure.

Adding Topic Tables from a PowerCenter Workflow

Add tables to the topic structure from a PowerCenter workflow.

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **From PowerCenter**. The **Add Tables from PowerCenter Workflow** page appears.

2. Choose one of the following options to add tables to the topic structure:
 - Choose **Select a workflow from the PowerCenter repository** to select a workflow from the PowerCenter repository. Either enter a string in the **Folder name** text box and then click **Search**, or click **Show All**, and then select a workflow from the list.
 - Choose **Select a workflow definition file (.xml)** to select a workflow definition file. Browse to select the file and then click **Upload**.
3. Click **Save**.

The **Add Tables from PowerCenter Workflow** page closes. The **Structure** page shows a list of topic tables and a preview of the topic structure.

Adding Topic Tables from a Metadata File

You can load a metadata file to Data Integration Hub and create a topic table that is based on the structure of the file. Data Integration Hub supports JSON, XML, XLS, and XLSX file formats.

1. In the **Structure** page of the New Topic wizard, click **Add Tables** and then select **From Metadata File**.

The **Add table from Metadata File** page appears.
2. Browse to select the file and click **Upload**.
3. Enter the table name, and click **Save**.

The **Structure** page shows a list of topic tables and a preview of the topic structure.

Creating a Topic Table

Add a table to the topic structure by creating the table.

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **Create**.

The **Create Table** page appears.
2. Enter the logical table name in the **Table Name** field.
3. If required, change the physical table name.
4. Optionally, enter a description of the table.
5. Use the add, up, and down buttons to add and order table columns. You must add at least one column to the table.

For each column, enter the following information:

- **Column Name.** Logical name of the column. Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
- **Physical Name.** Physical name of the column. Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
- **Data Type.** Select from the list of available data types.
- **Precision.** Enabled only for data types that support precision.
- **Scale.** Enabled only for data types that support data scaling.
- **Description.** Optional description of the column.

Note: Data Integration Hub adds the columns `DIH_PUBLICATION_INSTANCE_DATE`, `DIH_PUBLICATION_INSTANCE_ID`, and `DIH__UPDATE_STRATEGY_FLAG` to each topic table. When you finish creating the table, the columns show in the **Structure Preview** area of the **Structure** page of the topic wizard.

6. Click **OK**.

The **Create Table** page closes. The **Structure** page shows the topic table and a preview of the topic structure.

Step 4. Define Topic Table Relations

You can define relations between topic tables, to filter the data that automatic database and flat file subscriptions consume from the topic in the **Table Relations** page of the **New Topic** wizard. You can also define filter accelerators for topics that you plan to use for unbound subscriptions.

Tip: In the **Show** list, you can select to show all topic tables or to show a specific table.

1. In the **Filter Accelerator** column, define a column or multiple columns as filter accelerators. A filter accelerator indicates that the column will be used in subscription queries and requires performance-related handling by Data Integration Hub. Use this indicator with topics that you plan to use for unbound subscriptions.

When you use filter accelerators, consider the impact that performance-related handling has on system performance:

- Filter accelerators speed up the handling of subscriptions that use both filters and PowerCenter pushdown optimization. By default, the PowerCenter pushdown optimization option is selected for unbound subscriptions only.
 - Filter accelerators slow down the writing of publication data to the Data Integration Hub publication repository.
 - Filter accelerators have no impact on subscriptions that do not use filters.
2. In the **Key** column, define a column from a topic table for which to define relations as a primary key.
 3. Define primary keys for as many tables as required, one primary key for each table.
 4. For each table for which you defined a primary key, define one or more foreign keys. A foreign key must relate to a primary key from a different table.
 5. Click **Next**.

The **Data Retention** page appears.

Step 5. Define Data Retention Period

Define the data retention period in the **Data Retention** page of the **New Topic** wizard.

1. Choose one of the following options:

Retain published data for ... days after it is consumed

Enter the number of days that Data Integration Hub retains the data in the Data Integration Hub publication repository after the data is consumed. When the retention period ends, Data Integration Hub deletes the data from the repository.

Do not delete published data for this topic

Data Integration Hub does not delete the data from the publication repository.

2. Click **Next**.

The **Permissions** page appears.

Step 6. Assign Topic Permissions

Control access to the topic in the Operation Console in the **Permissions** page of the **New Topic** wizard. If you do not assign permissions, the topic is accessible by all Data Integration Hub users.

1. Select the category to which you want to assign permission to the topic under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

2. Repeat step [1](#) to assign additional categories.
3. Click **Next**.

The **Summary** page appears.

Step 7. Review Topic Settings and Save the Topic

Review the topic settings and save the topic in the **Summary** page of the **New Topic** wizard.

1. Review the topic settings.
2. Click **Finish**.

The **New Topic** wizard closes. The **Topics** page shows the topic you created. You can create publications that publish data to the topic and subscriptions that consume data from the topic. The topic is listed in the catalog. In the catalog you can view all the data sets that are available in Data Integration Hub and verify that the data that the target applications require exists in Data Integration Hub.

Creating a Topic with a File Store Publication Repository

To create a topic with a file store publication repository, perform the following tasks:

1. Access the **New Topic** wizard.
2. Define basic topic properties.
3. Create the topic data structure.
4. Optionally, define topic table relations.
5. Define the data retention period and the storage location.
6. Optionally, assign topic permissions.
7. Review the topic settings and save the topic.

Task Prerequisites

Before you start this task, obtain the details of the topic structure that you want to create.

Step 1. Access the New File Store Topic Wizard

Access the **New File Store Topic** wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Topics**.
The **Topics** page appears.

2. Click **New Topic**, select **File Store**, and then click **Create**.
The **New File Store Topic** wizard appears.

Step 2. Define Basic Topic Properties

Define topic properties in the **General** page of the **New Topic** wizard.

1. Enter the topic name.
2. Optionally, enter a description of the topic.
3. Optionally, assign a tag to the topic so that you can search for the topic on the **Topics** page and the **Catalog** page based on the tag.
4. Choose the topic type.
 - **Delta**. The topic instance contains only the latest data changes. If you choose this topic type, verify that the data source includes delta indicators.
 - **Full**. The topic instance contains all of the data changes that occurred after the last publication. Choose this topic type if you associate publications with custom mapping with the topic.
5. Under **Publication Repository** select **File repository**.
Note: If you select the option **Prevent new publications and new subscriptions to this topic** you cannot create publications and subscriptions that publish to and subscribe from the topic.
6. Click **Next**.
The **Structure** page appears.

Step 3. Create Topic Data Structure

Create and preview the data structure of the topic in the **Structure** page of the **New Topic** wizard.

1. Click **Add Tables**.
The **Create Table** page appears. The column table fields are populated with the default Data Integration Hub values, and you can only edit the column precision fields.
Note: Data Integration Hub adds the columns DIH_PUBLICATION_INSTANCE_DATE, DIH_PUBLICATION_INSTANCE_ID, and DIH__UPDATE_STRATEGY_FLAG to each topic table. When you finish creating the table, the columns show in the **Structure Preview** area of the **Structure** page of the topic wizard.
2. Enter the logical table name in the **Table Name** field.
3. If required, change the physical table name.
4. Optionally, enter a description of the table.
5. If required, edit the column precision fields.
6. Click **OK**.
The **Create Table** page closes. The **Structure** page shows the topic table and a preview of the topic structure.
7. Click **Next**.
The **Table Relations** page appears.

Step 4. Define Topic Table Relations

You can define relations between topic tables, to filter the data that automatic database and flat file subscriptions consume from the topic in the **Table Relations** page of the **New Topic** wizard. You can also define filter accelerators for topics that you plan to use for unbound subscriptions.

Tip: In the **Show** list, you can select to show all topic tables or to show a specific table.

1. In the **Filter Accelerator** column, define a column or multiple columns as filter accelerators. A filter accelerator indicates that the column will be used in subscription queries and requires performance-related handling by Data Integration Hub. Use this indicator with topics that you plan to use for unbound subscriptions.

When you use filter accelerators, consider the impact that performance-related handling has on system performance:

- Filter accelerators speed up the handling of subscriptions that use both filters and PowerCenter pushdown optimization. By default, the PowerCenter pushdown optimization option is selected for unbound subscriptions only.
 - Filter accelerators slow down the writing of publication data to the Data Integration Hub publication repository.
 - Filter accelerators have no impact on subscriptions that do not use filters.
2. In the **Key** column, define a column from a topic table for which to define relations as a primary key.
 3. Define primary keys for as many tables as required, one primary key for each table.
 4. For each table for which you defined a primary key, define one or more foreign keys. A foreign key must relate to a primary key from a different table.
 5. Click **Next**.

The **Data Retention** page appears.

Step 5. Define Data Retention Period and Storage Location

Define the data retention period and the data storage location in the **Data Retention** page of the **New Topic** wizard.

1. Enter the number of days that Data Integration Hub retains the data in the Data Integration Hub publication database after the data is consumed in the **Publication data retention period** field.

2. Click **Advanced**.

The **Data Storage Location** area appears.

3. Choose the location in the database where Data Integration Hub stores data that is published to the topic.
 - Default. Data Integration Hub stores the data based on the default storage configuration. For example, the default File Group.
 - Custom. Browse to select an available data storage group in the publication repository database. For example, browse the available file groups.
4. Click **Next**.

The **Permissions** page appears.

Step 6. Assign Topic Permissions

Control access to the topic in the Operation Console in the **Permissions** page of the **New Topic** wizard. If you do not assign permissions, the topic is accessible by all Data Integration Hub users.

1. Select the category to which you want to assign permission to the topic under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

2. Repeat step [1](#) to assign additional categories.
3. Click **Next**.

The **Summary** page appears.

Step 7. Review Topic Settings and Save the Topic

Review the topic settings and save the topic in the **Summary** page of the **New Topic** wizard.

1. Review the topic settings.
2. Click **Finish**.

The **New Topic** wizard closes. The **Topics** page shows the topic you created. You can create publications that publish data to the topic and subscriptions that consume data from the topic. The topic is listed in the catalog. In the catalog you can view all the data sets that are available in Data Integration Hub and verify that the data that the target applications require exists in Data Integration Hub.

Creating a Topic with a Real-time Publication Repository

This topic describes how to create a topic that fetches consumer metrics from the Apache Kafka messaging platform.

Task Prerequisites

Before you start this task, obtain the details of the topic structure that you want to create.

To create a new topic with the Apache Kafka repository, perform the following tasks:

1. In the Navigator, click **Hub Management > Topics**.
The **Topics** page appears.
2. Click **New Topic**, select **Real-Time**, and then click **Create**.
The **Create Real-timeTopic** wizard appears.
3. Enter the topic name.
4. Optionally, enter a description of the topic.
5. Optionally, assign a tag to the topic so that you can search for the topic on the **Topics** page and the **Catalog** page based on the tag.
6. Enter the number of partitions.
The maximum number of partitions that the topic can have is 999.
7. Select the intervals in which the topic must fetch the consumer metrics from the Apache Kafka server.

The default value is 30 seconds.

8. Click **Permissions** to provide permissions of a category to the topic.
 - a. Select a category from the **Available Categories** and click the right arrow.

The category that you select is displayed in the list of **Selected Categories**. The topic now has permissions that are defined for the selected category.

9. Click **Save**.

The **Create Real-time Topic** wizard closes. The Topics page displays the topic you created. You can create publications that publish data to the topic and subscriptions that consume data from the topic.

The Catalog page lists the new topic.

CHAPTER 4

Creating Publications

This chapter includes the following topics:

- [Creating Publications Overview, 39](#)
- [Creating an Automatic Relational Database Publication, 40](#)
- [Creating an Automatic Flat File Publication, 45](#)
- [Creating an Automatic Pass-through File Publication, 52](#)
- [Creating an Automatic Data-driven Publication, 55](#)
- [Creating a Custom Batch Publication, 57](#)
- [Creating a Custom Multi-latency Publication, 60](#)
- [Creating a Custom Cloud Publication, 62](#)
- [Creating a Custom Pass-through Kafka Publication, 64](#)
- [Creating a Modular Cloud Publication, 65](#)

Creating Publications Overview

In this chapter, you create publications that publish data. You must have first completed the chapter "Creating Topics."

When you create a publication, you choose the topic to which the application publishes the data.

Chapter Objectives

In this chapter, you perform the following tasks:

- Create an automatic publication with a relational database source.
- Create an automatic publication with a flat file source.
- Create an automatic publication with an unstructured file source.
- Create an automatic data-driven publication that publishes data over a REST API.
- Create a custom publication with a batch workflow.
- Create a custom publication with a real-time workflow.
- Create a custom publication with an Informatica Intelligent Cloud Services task.
- Create a modular publication with a cloud source.

Creating an Automatic Relational Database Publication

To create an automatic publication with a relational database source, perform the following tasks:

1. Access the create publication wizard.
2. Define basic publication properties.
3. Select the publication source. Optionally, define joins for the publication.
4. Review the publication field mapping, and if required, edit the mapping that Data Integration Hub generates by default.
5. Optionally, define a filter for the publication.
6. Define the publication schedule.
7. Optionally, assign permissions to publication. Select the category to which you want to assign the publication. Only the categories assigned to the associated application are available to assign to the publication.
8. Review the publication settings and save the publication.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- A database that holds the source data exists.
- A connection to the source database is configured in the Data Integration Hub Operation Console.
- You have created an application.
- You have created a topic.

Step 1. Access the Create Automatic Database Publication Wizard

Access the create publication wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Automatic > Relational Database**, and then click **Create**.
The **Create Automatic Database Publication** wizard appears.

Step 2. Define Basic Publication Properties

Define publication properties in the **General** page of the publication wizard.

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.

4. Click **Next**.
The **Processing** page appears.
5. Click **Next**.
The **Source** page appears.

Step 3. Select Publication Source

Select the publication source from which Data Integration Hub reads data in the **Source** page of the publication wizard.

1. Select a connection from the **Database Connection** list.
The **Define new joins by joining existing tables** section appears.
2. Optionally, create joins for the publication. A join combines columns from two or more tables in the database. You can create multiple joins, and you can combine data from joins into new joins. Joins are virtual entities, and are not created in the source database or in the topic.

To create joins, perform the following actions:

- a. By default, the **Source** page shows the default schema only and the default schema is selected for both the left and right tables. To select non default schemas, clear the option **Show default schema only** and then select schemas for the left right tables.
Note: When you select non default schemas, do not use two source tables with the same name from different schemas to define the join. The source table names that you use to create the join must be unique across the schemas.
 - b. Select the tables to join from the **Select Left Table** and **Select Right Table** lists.
 - c. Select the join type from the **Select Join Type** list and then enter the join name in the **Join Name** field. The name that you assign to the join shows in the **Field Mapping** page, where you can map the join to a topic table.
 - d. Select the column to join from the left table and then select the column to join from the right table. To add more columns to the join click the plus sign and then select the columns to join.
 - e. Click **Create Join**.
The new join appears on the **Source** page.
 - f. Repeat step [2](#) to create as many joins as required.
3. Click **Next**.
The **Field Mapping** page appears.

Step 4. View and Edit Publication Field Mapping

View and edit the mapping of source tables and fields to topic tables and fields on the **Field Mapping** page of the publication wizard. If you created joins in "Step 3. Select Publication Source", configure field mapping of joins and fields to topic tables and fields.

1. Perform the following actions to map a join to a topic table:
 - a. Click the edit table mapping icon.
The **Edit Table Mapping** dialog box appears.
 - b. Click the row of the join that you want to map and then click **Map Source Table**.
The **Edit Table Mapping** dialog box closes. The **Field Mapping** page shows the join you mapped in the **Mapped Source Table** list.

- c. Go to Step [3](#) to map the fields in the join to fields in the topic table.
2. Perform the following actions to view and edit the mapping of a source table to a topic table:
 - a. Click the edit table mapping icon.
The **Edit Table Mapping** dialog box appears.
 - b. Search for tables in the database. Note the following guidelines:
 - To search for tables by table name, enter a string in the **Find Source Table** text box and then click **Search**.
 - To search for tables that use a schema other than the default schema, clear the option **Show default schema only**, enter a string in the **Schema** text box, and then click **Search**.
 - To clear the search results and to show only tables that use the default schema, select the option **Show default schema only** and then click **Show All**.
 - The search is not case sensitive.
 - You can search for a substring.

The **Search Results** section displays the source tables that match the search string and the schema that each table uses. The name of the topic table for which you are editing the mapping shows in the row of the source table that is mapped to the topic table, in the **Mapped To** column.
 - c. To map a different source table to the topic table, click the row of the source table that you want to map and then click **Map Source Table**.
The **Edit Table Mapping** dialog box closes. The **Field Mapping** page shows the revised mapping.
3. Perform the following actions to configure, view, and edit the mapping of fields in a source table or in a join to fields in a topic table:
 - a. Click the edit field mapping icon.
The **Edit Field Mapping** dialog box appears.
 - b. To search for fields that are used in the topic table and in the source table, enter a string for the field name in **Find Field**, and then click **Search**. The search is not case sensitive. You can search for a substring.
Choose **Display unmapped fields only** to show only unmapped fields in the search results.
Source and topic fields that match the search string appear. A green checkmark next to a field name indicates that the field is mapped.
 - c. To map a source table field to a topic table field, drag the field from the **Source Table** section to the **Topic Table** section.
 - d. To add an expression to a field, in the **Actions** column, click the expression icon to open the **Expression Editor**, and then select fields and functions to add to the field.
 - e. To clear the mapping of a field, in the **Actions** column, click the clear icon.
 - f. To view field details, rest on the details icon to the left of a field name.
 - g. To revert to the default Data Integration Hub field mapping, click **Auto Map**.
 - h. To clear all field mappings, click **Clear All**.
 - i. Click **OK** to map the fields.
The **Edit Field Mapping** dialog box closes.
4. To revert to the default Data Integration Hub table and field mapping, click **Auto Map** in the **Field Mapping** page.
5. To clear all table and field mapping, click **Clear All** in the **Field Mapping** page.
Note: The publication must contain at least one mapped source table.

- In the **Field Mapping** page, click **Next**.
The **Filter** page appears.

Step 5. Define a Filter

To define the data that an automatic relational database or flat file publication publishes, add a filter to the publication. When you add a filter to a publication, you can define filter conditions to table rows.

You can use basic expressions and advanced expressions to define filter conditions. You can use more than one method to add filters to a table row.

Basic Expression

Use this method if you want to apply a condition to a table row and the operators and values in the basic condition builder meet your requirements. For example, for an Orders table, add a condition that the date in the ShippedDate column is greater than April 1, 2016.

Advanced Expression

Add filter conditions to tables with PowerCenter expressions. For example, for an Orders table, add the following expression:

```
(ShipCountry='USA') and ((ShipCity='New York') or (ShipCity='Los Angeles'))
```

The filters use an AND logic. If multiple filters exist for a table, Data Integration Hub publishes only the rows that meet all of the conditions.

- Choose the type of expression that you want to create.
- Select the table to which to apply the filter from the **Select Table** list.
- Enter the condition parameters according to the method that you selected:
 - Basic expression: go to step [4](#).
 - Advanced expression: go to step [5](#).
- To create a basic expression, perform the following steps:
 - Select the column to filter from the **Select Column** list.
 - Select the filter operator from the **Select Operator** list. The available operators depend on the type of content in the column.
 - When **Select Value** is enabled, select or enter a value for the operator.
 - Click the plus icon to the right of the expression line.
The condition shows in the condition list.
- To create an advanced expression, perform the following steps:
 - Enter a valid PowerCenter expression in the text field. Alternatively, prepare a basic expression, select **Advanced Expression**, and then, in the expression text area, complete the expression. The expression is limited to the selected table. The expression can contain up to 1024 characters.
 - Click the plus icon to the right of the expression line.
The condition shows in the condition list.
- Repeat steps [1](#) through [5](#) to add the required conditions.
- Click **Next**.
The **Schedule** page appears.

Step 6. Define Publication Schedule

Define the method and the frequency of data publishing in the **Schedule** page of the publication wizard.

1. Select the method and the frequency of data publishing.

Manually or by an external trigger

No schedule. You can use the following methods to run the publication:

- Run manually. Click the Run arrow on the **Publications** page.
- Run by an API. Call a command-line API or a REST API that starts the publication.
- Run by a PowerCenter workflow. Start a publication transformation.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

By schedule

Runs the publication according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the publication in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the publication in intervals of up to 24 hours. You select the number of hours from the list. The publication runs at the beginning of the hour. For example, if you enter 2, the publication runs at 00:00, 02:00, and at consecutive two-hour intervals.
- Daily. Runs the publication at the same hour every day.
- Weekly. Runs the publication every week on one or more days at the same hour.
- Monthly. Runs the publication every month on a specific date or a specific day at the same hour.

Define the publication intervals in the **Repeat running** area.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

After the following publication or subscription run completes

Runs the publication after the run of the publication or subscription that you select here completes.

2. Click **Next**.

The **Permissions** page appears.

Step 7. Assign Publication Permissions

Control access to the publication on the Operation Console by assigning categories to the publication on the **Permissions** page of the publication wizard. If you do not assign categories to the publication, it inherits the categories from the associated application. If there are no categories assigned to the application, all Data Integration Hub users can access the publication.

1. Select one or more categories to assign permission to the publication under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

You can unassign categories from **Selected Categories** to **Available Categories**.

2. Click **Next**.

The **Summary** page appears.

Step 8. Review Publication Settings and Save the Publication

Review the publication settings and save the publication in the **Summary** page of the publication wizard.

1. Review the publication settings.
2. Click **Finish**.

The publication wizard closes. The **Publications** tab of the **Edit Application** page shows the publication you created. The publication is listed in the Publications catalog. If you configured a publication schedule, the publication publishes data according to the defined schedule.

Creating an Automatic Flat File Publication

To create an automatic publication with a flat file source, perform the following tasks:

1. Access the create publication wizard.
2. Define basic publication properties.
3. Select and configure the publication source.
4. Optionally, define joins for the publication.
5. Review the publication field mapping, and if required, edit the mapping that Data Integration Hub generates by default.
6. Optionally, define a filter for the publication.
7. Define the publication schedule.
8. Optionally, assign permissions to publication. Select the category to which you want to assign the publication. Only the categories assigned to the associated application are available to assign to the publication.
9. Review the publication settings and save the publication.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- A file or files that hold the source data exist.
- You know the location of the source files.
- If you are publishing from HDFS, the HDFS connection that the workflow uses is configured in the Data Integration Hub Operation Console.
- You have created an application.
- You have created a topic.

Note: If you are publishing from HDFS, Data Integration Hub does not delete the source files after it reads them. If required, you must delete the files yourself.

Step 1. Access the Create Automatic File Publication Wizard

Access the create publication wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.

The **Edit Application** page appears.

3. Select the **Publications** tab. Click **New**, then select **Automatic > Flat File**, and then click **Create**.
The **Create Automatic File Publication** wizard appears.

Step 2. Define Basic Publication Properties

Define publication properties in the **General** page of the publication wizard.

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Click **Next**.
The **Processing** page appears.
5. Click **Next**.
The **Source** page appears.

Step 3. Select and Configure Publication Source

Select the type and the location of the publication source, from where Data Integration Hub picks up the files, in the **Source** page of the publication wizard.

1. Select the source type from the **Source Type** list.
If you select the HDFS source type, the **HDFS Connection** field appears on the page.
2. If you select the HDFS source type, select a connection from the **HDFS Connection** list.
3. To use file transfer to publish files from remote servers, select **Use File Transfer**.
The **Connection** and **Delete Files from Source After Pickup** fields appear on the page.
4. If you selected to use file transfer, configure the following file transfer properties:
Connection
Select a connection from the list.
Delete Files from Source After Pickup
Data Integration Hub deletes the files after it reads them. You must have permissions on the remote server to delete the files.
5. Enter the location of the files that contain the source data in the **Directory** field. If you use file transfer, you cannot use patterns to define the directory path. You can, however, use patterns to define the file name.
6. Click **Configure**.
The **Configure Flat File Source** dialog box appears.
7. In the **Configure Flat File Source** dialog box configure the following properties and then click **OK**:
Logical Name
Description name of the source. The name identifies the source when you configure field mapping for the publication.

File Name

Name of the file that Data Integration Hub picks up and publishes to the publication repository. The file name can include asterisk wildcards and variables. Enter variables in the following format:

`($pattern)`. For example: `Input_($PublicationName)_($sequence).in`.

Configure source based on

Select one of the following options:

- Topic Table. Select the topic table that represents the structure of the flat file.
- Sample file. A sample file that represents the structure of the flat file. Browse to select and upload the file.

File Format

Expand the **File Format** area to view and edit the format of the file. Changing the file format might affect the column structure.

The **File Format** area can include the following properties:

Code page

Character encoding used in the file.

Number of initial rows to skip

Number of rows at the beginning of the file that Data Integration Hub ignores when it reads the file. Data Integration Hub published only the subsequent rows.

Import column names

Optional. Select **Yes** to use the column names in the file as the default column headers in the table. Enter the number of the row that serves as the file's header row in **From row**.

Delimiter

Delimiter used in the file to separate between columns. Select a predefined delimiter or select **Custom** to define a custom delimiter. For information about the supported column delimiters, see the sections about importing delimited flat files and about updating delimited file properties in the *PowerCenter Designer Guide*.

Text qualifier

Optional. Symbols used in the file to enclose a string.

Thousands Separator

Optional. Symbol used in the file as a thousands separator.

Decimal Separator

Symbol used in the file as a decimal separator.

Datetime Format

Date and time format used in the file. Select a predefined format or select **Other** to define a custom format. For information about the supported datetime formats, see the *PowerCenter Transformation Language Reference*.

Note: The datetime format can contain up to 50 characters.

Column Structure

The column structure depends on the file format settings that you configure in the **File Format** area. The table must contain at least one column.

If you provide a sample file, Data Integration Hub reads the file according to the file format settings and presents the columns that it detects in the sample file. Use the **Add Column** button, the up and down arrows, and the remove icon to add, order, and remove table columns.

Click the edit icon to the right of a column to edit column parameters.

Each column must contain the following parameters:

Column Name

Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.

If you select the option **Import column names**, Data Integration Hub populates the column names with the strings of the defined row. If you do not select the option **Import column names**, Data Integration Hub assigns default names to the columns. For example, **Field1**, **Field2**, **Fieldn**.

Data Type

Select from the list of available data types. By default, Data Integration Hub reads the data as string.

Precision

Enabled only for data types that support precision.

Scale

Enabled only for data types that support data scaling.

Sample File Preview

If you select a sample file that represents the structure of the flat file, this area shows the data in the sample file. The **Preview** area presents the data in the structure that Data Integration Hub applies when it distributes the data into topic fields in the topic table, according to the **File Format** parameters.

The **Configure Flat File Source** dialog box closes.

8. Repeat steps [6](#) through [7](#) for each table topic that appears on the **Source** page.
9. In the **Source** page click **Next**.

The **Join** page appears.

Step 4. Define Joins

Pull data from multiple source files into a single topic table by creating a join. You can create multiple joins, and you can combine data from joins into new joins. Joins are virtual entities, and are not created in the topic.

1. Optionally, create joins for the publication. To create a join, perform the following actions:
 - a. Select the files to join from the **Select Left File** and **Select Right File** lists.
 - b. Select the join type from the **Select Join Type** list and then enter the join name in the **Join Name** field. The name that you assign to the join shows in the **Field Mapping** page, where you can map the join to a topic table. The new join appears on the **Join** page.
 - c. Select the column to join from the left file and then select the column to join from the right file. To add more columns to the join click the plus sign and then select the columns to join.
 - d. Click **Create Join**.
2. Repeat step [1](#) to create as many joins as required.

3. Click **Next**.

The **Field Mapping** page appears.

Step 5. View and Edit Publication Field Mapping

View the mapping of source tables and fields to topic tables and fields and edit field mapping on the **Field Mapping** page of the publication wizard. If you created joins in "Step 4. Define Joins", configure field mapping of joins and fields to topic tables and fields.

1. Perform the following actions to map a source table or a join to a topic table:
 - a. Click the edit table mapping icon.
The **Edit Table Mapping** dialog box appears.
 - b. Click the row of the table or join that you want to map and then click **Map Source Table**.
The **Edit Table Mapping** dialog box closes. The **Field Mapping** page shows the table you mapped in the **Mapped Source Logical Name** list.
2. Perform the following actions to configure, view, and edit the mapping of fields in a source table or in a join to fields in a topic table:
 - a. Click the edit field mapping icon.
The **Edit Field Mapping** dialog box appears. Data Integration Hub maps the fields automatically, according to name match.
 - b. To search for fields that are used in the topic table and the source table, enter a string for the field name in **Find Field**, and then click **Search**. The search is not case sensitive. You can search for a substring.
Choose **Display unmapped fields only** to show only unmapped fields in the search results.
Source and topic fields that match the search string appear. A green checkmark next to a field name indicates that the field is mapped.
 - c. To map a source table field to a topic table field, drag the field from the **Source Table** section to the **Topic Table** section.
 - d. To add an expression to a field, in the **Actions** column, click the expression icon to open the **Expression Editor**, and then select fields and functions to add to the field.
 - e. To clear the mapping of a field, in the **Actions** column, click the clear icon.
 - f. To view field details, rest on the details icon to the left of a field name.
 - g. To revert to the default Data Integration Hub field mapping, click **Auto Map**.
3. Click **OK** to apply the mapping.
The **Edit Field Mapping** dialog box closes.
4. To clear all table and field mapping, click **Clear All** in the **Field Mapping** page.
Note: The publication must contain at least one mapped source table.
5. In the **Field Mapping** page, click **Next**.
The **Filter** page appears.

Step 6. Define a Filter

To define the data that an automatic relational database or flat file publication publishes, add a filter to the publication. When you add a filter to a publication, you can define filter conditions to table rows.

You can use basic expressions and advanced expressions to define filter conditions. You can use more than one method to add filters to a table row.

Basic Expression

Use this method if you want to apply a condition to a table row and the operators and values in the basic condition builder meet your requirements. For example, for an Orders table, add a condition that the date in the ShippedDate column is greater than April 1, 2016.

Advanced Expression

Add filter conditions to tables with PowerCenter expressions. For example, for an Orders table, add the following expression:

```
(ShipCountry='USA') and ((ShipCity='New York') or (ShipCity='Los Angeles'))
```

The filters use an AND logic. If multiple filters exist for a table, Data Integration Hub publishes only the rows that meet all of the conditions.

1. Choose the type of expression that you want to create.
2. Select the table to which to apply the filter from the **Select Table** list.
3. Enter the condition parameters according to the method that you selected:
 - Basic expression: go to step [4](#).
 - Advanced expression: go to step [5](#).
4. To create a basic expression, perform the following steps:
 - a. Select the column to filter from the **Select Column** list.
 - b. Select the filter operator from the **Select Operator** list. The available operators depend on the type of content in the column.
 - c. When **Select Value** is enabled, select or enter a value for the operator. Values of string operators can contain up to 90 digits. Values of numeric operators can contain up to 15 digits.
 - d. Click the plus icon to the right of the expression line.
The condition shows in the condition list.
5. To create an advanced expression, perform the following steps:
 - a. Enter a valid PowerCenter expression in the text field. Alternatively, prepare a basic expression, select **Advanced Expression**, and then, in the expression text area, complete the expression. The expression is limited to the selected table. The expression can contain up to 1024 characters.
 - b. Click the plus icon to the right of the expression line.
The condition shows in the condition list.
6. Repeat steps [1](#) through [5](#) to add the required conditions.
7. Click **Next**.
The **Schedule** page appears.

Step 7. Define Publication Schedule

Define the method and the frequency of data publishing in the **Schedule** page of the publication wizard.

1. Select the method and the frequency of data publishing.

When the file is ready to be published

Runs the publication after the published files are ready, the next time it polls the remote sources. Define the maximal period of time that Data Integration Hub waits for the files to be available in the directory that you defined in the **Source** page in **Wait for all data to be published ... hours**. When the maximal period of time ends, Data Integration Hub discards the file events of the publication and changes the status of the publication event to Error.

Manually or by an external trigger

No schedule. You can use the following methods to run the publication:

- Run manually. Click the Run arrow on the **Publications** page.
- Run by an API. Call a command-line API or a REST API that starts the publication.
- Run by a PowerCenter workflow. Start a publication transformation.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

By schedule

Runs the publication according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the publication in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the publication in intervals of up to 24 hours. You select the number of hours from the list. The publication runs at the beginning of the hour. For example, if you enter 2, the publication runs at 00:00, 02:00, and at consecutive two-hour intervals.
- Daily. Runs the publication at the same hour every day.
- Weekly. Runs the publication every week on one or more days at the same hour.
- Monthly. Runs the publication every month on a specific date or a specific day at the same hour.

Define the publication intervals in the **Repeat running** area.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

After the following publication or subscription run completes

Runs the publication after the run of the publication or subscription that you select here completes.

2. Optionally, define a retry policy. A retry policy defines the number of times Data Integration Hub retries to run the publication in case of failure, and the retry interval. The policy does not apply to publications that you run manually.
3. Click **Next**.

The **Permissions** page appears.

Step 8. Assign Publication Permissions

Control access to the publication on the Operation Console by assigning categories to the publication on the **Permissions** page of the publication wizard. If you do not assign categories to the publication, it inherits the categories from the associated application. If there are no categories assigned to the application, all Data Integration Hub users can access the publication.

1. Select one or more categories to assign permission to the publication under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

You can unassign categories from **Selected Categories** to **Available Categories**.

2. Click **Next**.
The **Summary** page appears.

Step 9. Review Publication Settings and Save the Publication

Review the publication settings and save the publication in the **Summary** page of the publication wizard.

1. Review the publication settings.
2. Click **Finish**.

The publication wizard closes. The **Publications** tab of the **Edit Application** page shows the publication you created. The publication is listed in the Publications catalog. If you configured a publication schedule, the publication publishes data according to the defined schedule.

Creating an Automatic Pass-through File Publication

To create an automatic publication with a pass-through file source, perform the following tasks:

1. Access the create publication wizard.
2. Define basic publication properties.
3. Configure the publication source.
4. Define the publication schedule.
5. Optionally, assign permissions to publication. Select the category to which you want to assign the publication. Only the categories assigned to the associated application are available to assign to the publication. .
6. Review the publication settings and save the publication.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- A file or files that hold the source data exist.
- You know the location of the source files.
- You have created an application.
- You have created a topic.

Step 1. Access the Create Automatic Pass-through File Publication Wizard

Access the create publication wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Automatic > Pass-through File**, and then click **Create**.
The **Create Automatic Pass-through File Publication** wizard appears.

Step 2. Define Basic Publication Properties

Define publication properties in the **General** page of the publication wizard.

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.

The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.

4. Click **Next**.
The **Processing** page appears.
5. Click **Next**.
The **Source** page appears.

Step 3. Configure Publication Source

Configure the source from which Data Integration Hub publishes files in the **Source** page of the publication wizard.

1. To use file transfer to publish files from remote servers, select **Use File Transfer**.
The **Connection** and **Delete Files from Source After Pickup** fields appear on the page.
2. If you selected to use file transfer, configure the following file transfer properties:

Connection

Select a connection from the list.

Delete Files from Source After Pickup

Data Integration Hub deletes the files after it reads them. You must have permissions on the remote server to delete the files.

3. Enter the location of the file or files that Data Integration Hub publishes in the **Directory** field. If you use file transfer, you cannot use patterns to define the directory path. You can, however, use patterns to define the file name.
4. Click the edit icon next to a table topic and enter the name of the source file in the **File Name** column. The file name can include asterisk wildcards and variables. Enter variables in the following format: `($pattern)`. For example: `Input_($PublicationName)_($sequence).in`.
5. Repeat step [4](#) for each table topic that appears on the **Source** page.
6. Click **Next**.
The **Schedule** page appears.

Step 4. Define Publication Schedule

Define the method and the frequency of data publishing in the **Schedule** page of the publication wizard.

1. Select the method and the frequency of data publishing.

When the file is ready to be published

Runs the publication after the published files are ready, the next time it polls the remote sources.

Define the maximal period of time that Data Integration Hub waits for the files to be available in the directory that you defined in the **Source** page in **Wait for all data to be published ... hours**. When the maximal period of time ends, Data Integration Hub discards the file events of the publication and changes the status of the publication event to Error.

Manually or by an external trigger

No schedule. You can use the following methods to run the publication:

- Run manually. Click the Run arrow on the **Publications** page.
- Run by an API. Call a command-line API or a REST API that starts the publication.
- Run by a PowerCenter workflow. Start a publication transformation.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

By schedule

Runs the publication according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the publication in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the publication in intervals of up to 24 hours. You select the number of hours from the list. The publication runs at the beginning of the hour. For example, if you enter 2, the publication runs at 00:00, 02:00, and at consecutive two-hour intervals.
- Daily. Runs the publication at the same hour every day.
- Weekly. Runs the publication every week on one or more days at the same hour.
- Monthly. Runs the publication every month on a specific date or a specific day at the same hour.

Define the publication intervals in the **Repeat running** area.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

After the following publication or subscription run completes

Runs the publication after the run of the publication or subscription that you select here completes.

2. Optionally, define a retry policy. A retry policy defines the number of times Data Integration Hub retries to run the publication in case of failure, and the retry interval. The policy does not apply to publications that you run manually.
3. Click **Next**.

The **Permissions** page appears.

Step 5. Assign Publication Permissions

Control access to the publication on the Operation Console by assigning categories to the publication on the **Permissions** page of the publication wizard. If you do not assign categories to the publication, it inherits the categories from the associated application. If there are no categories assigned to the application, all Data Integration Hub users can access the publication.

1. Select one or more categories to assign permission to the publication under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

You can unassign categories from **Selected Categories** to **Available Categories**.

2. Click **Next**.

The **Summary** page appears.

Step 6. Review Publication Settings and Save the Publication

Review the publication settings and save the publication in the **Summary** page of the publication wizard.

1. Review the publication settings.
2. Click **Finish**.

The publication wizard closes. The **Publications** tab of the **Edit Application** page shows the publication you created. The publication is listed in the Publications catalog. If you configured a publication schedule, the publication publishes data according to the defined schedule.

Creating an Automatic Data-driven Publication

To create an automatic data-driven publication, perform the following tasks:

1. Access the create publication page.
2. Configure publication properties and select the topic into which to publish the data.
After you configure the publication properties, you can copy the following URLs from the create publication page:
 - REST URL of the API. Use this URL to publish the data.
 - REST URL of the API Swagger file with the Swagger structure for the topic into which the publication publishes data. Use the structure in the publication request.You use the URLs when you create the request that runs the publication.
3. Optionally, assign permissions to publication. Select the category to which you want to assign the publication. Only the categories assigned to the associated application are available to assign to the publication.
4. Create a request to run the publication.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- You have created an application.
- You have created a topic.

Step 1. Access the Create Automatic Data-driven Publication page

Access the create publication page in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Automatic > Data-driven**, and then click **Create**.
The **Create Automatic Data-driven Publication** page appears.

Step 2. Define Publication Properties

Define publication properties in the create publication page.

1. Enter the publication name.

The following URLs are automatically updated:

- REST URL of the API. Use this URL to publish the data.
- REST URL of the API Swagger file with the Swagger structure for the topic into which the publication publishes data. Use the structure in the publication request.

You use the URLs when you create the request that runs the publication.

2. Optionally, enter a description of the publication.

3. Select a topic from the **Topic** list.

The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.

4. Optionally, define the time interval according to which Data Integration Hub groups the published data.

- If no grouping is defined, that is, **Group publications every ... seconds** is set to zero, Data Integration Hub does not group the published data. Data Integration Hub writes the data to the publication repository as it is being published.
- If you define a grouping time interval, Data Integration Hub groups the published data and writes it to the publication repository based on the time interval. For example, if you define **Group publications every 10 seconds**, Data Integration Hub will group the published data at the end of each 10-second period.

5. Optionally, assign publication permissions.

Select one or more categories to which you want to assign permission to the publication under **Available Categories** and click the right arrow.

6. Click **Save**.

Step 3. Create a Request to Run the Publication

To create a request to run the publication, you have to copy the URL of the REST API and the URL of the API Swagger file from the publication that you created in [“Step 2. Define Publication Properties” on page 56](#).

- Create a POST request with the following details:

- Request URL: REST API URL. For example:

```
http://hostname:18080/dih-console/api/v1/publication/MyPublication/data
```

Where `MyPublication` is the publication name.

- Request body: Use the structure definition in the Swagger structure for the topic into which the publication publishes data to prepare the body of the request. Access the REST URL of the API Swagger file to view and copy the structure, for example:

```
http://hostname:18080/dih-console/api/v1/publication/MyPublication/data?swagger
```

Where `MyPublication` is the publication name.

The following is an example of request payload for a topic that contains two topic tables, each containing two columns:

```
{
  "Employee": [
    {
```

```

        "EmployeeName":"John Smith"
      },
      {
        "EmployeeId":"AA18"
      }
    ],
    "Department":[
      {
        "DepartmentName":"Computer Science"
      },
      {
        "DepartmentId":"Dep13"
      }
    ]
  }
}

```

Creating a Custom Batch Publication

To create a custom publication with a batch workflow, perform the following tasks:

1. Access the create publication wizard.
2. Define basic publication properties.
3. Select the publication workflow. If you use file transfer, select the connection to the source from which Data Integration Hub reads the files and choose whether Data Integration Hub deletes files from the source after pickup.
4. If you want to run a pre-process on the publication, select the pre-processing mapping.
5. If the workflow includes parameters, set the publication parameters.
6. Define the publication schedule.
7. Optionally, assign permissions to publication. Select the category to which you want to assign the publication. Only the categories assigned to the associated application are available to assign to the publication.
8. Review the publication settings and save the publication.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- The publication workflow is configured in the Data Integration Hub Operation Console.
- The publication workflow is a batch workflow.
- If you use file transfer, the connection that the workflow uses is configured in the Data Integration Hub Operation Console.
- You have obtained the required parameter settings.
- You have created an application.
- You have created a topic.

Step 1. Access the Create Custom Batch Publication Wizard

Access the create publication wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.

2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Custom > Batch**, and then click **Create**.
The **Create Custom Batch Publication** wizard appears.

Step 2. Define Basic Publication Properties

Define publication properties in the **General** page of the publication wizard.

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Click **Next**.
The **Processing** page appears.

Step 3. Select Publication Workflow

Select a Data Integration Hub publication workflow in the **Processing** page of the publication wizard. If required, set parameter values.

1. Select a workflow from the **Custom Mapping** list. Only publication workflows that are based on a batch workflow are available for selection.
If the workflow supports file transfer, the **Use File Transfer** option is enabled.
If the workflow contains parameters, the parameters are shown in the **Publication Parameters** area.
2. To use file transfer, select the **Use File Transfer** option and then perform the following actions:
 - a. Select the connection to the source from which Data Integration Hub reads the files in the **Connection** field.
 - b. To prevent Data Integration Hub from deleting files from the remote server after reading the files, clear the option **Delete Files from Source After Pickup**.
3. If you want to run a pre-process on the publication, select a publication pre-processing workflow from the **Pre-Processing Mapping** list.
4. If the workflow includes parameters, set the values of the parameters in the **Publication Parameters** area.
5. Click **Next**.
The **Schedule** page appears.

Step 4. Define Publication Schedule

Define the method and the frequency of data publishing in the **Schedule** page of the publication wizard.

1. Select the method and the frequency of data publishing.

Manually or by an external trigger

No schedule. You can use the following methods to run the publication:

- Run manually. Click the Run arrow on the **Publications** page.
- Run by an API. Call a command-line API or a REST API that starts the publication.
- Run by a PowerCenter workflow. Start a publication transformation.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

By schedule

Runs the publication according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the publication in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the publication in intervals of up to 24 hours. You select the number of hours from the list. The publication runs at the beginning of the hour. For example, if you enter 2, the publication runs at 00:00, 02:00, and at consecutive two-hour intervals.
- Daily. Runs the publication at the same hour every day.
- Weekly. Runs the publication every week on one or more days at the same hour.
- Monthly. Runs the publication every month on a specific date or a specific day at the same hour.

Define the publication intervals in the **Repeat running** area.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

After the following publication or subscription run completes

Runs the publication after the run of the publication or subscription that you select here completes.

2. Click **Next**.

The **Permissions** page appears.

Step 5. Assign Publication Permissions

Control access to the publication on the Operation Console by assigning categories to the publication on the **Permissions** page of the publication wizard. If you do not assign categories to the publication, it inherits the categories from the associated application. If there are no categories assigned to the application, all Data Integration Hub users can access the publication.

1. Select one or more categories to assign permission to the publication under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

You can unassign categories from **Selected Categories** to **Available Categories**.

2. Click **Next**.

The **Summary** page appears.

Step 6. Review Publication Settings and Save the Publication

Review the publication settings and save the publication in the **Summary** page of the publication wizard.

1. Review the publication settings.

2. Click **Finish**.

The publication wizard closes. The **Publications** tab of the **Edit Application** page shows the publication you created. The publication is listed in the Publications catalog. If you configured a publication schedule, the publication publishes data according to the defined schedule.

Creating a Custom Multi-latency Publication

To create a custom multi-latency publication with a real-time mapping or workflow, perform the following tasks:

1. Access the create publication wizard.
2. Define basic publication properties.
3. Select the publication workflow.
4. Define the publication intervals.
5. Optionally, assign publication permissions, if the associated application is assigned with categories.
6. Review the publication settings and save the publication.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- The publication workflow is configured in the Data Integration Hub Operation Console.
- The publication workflow is based on a Data Engineering Streaming real-time mapping or on a PowerCenter real-time workflow.
- You have obtained the required parameter settings.
- You have created an application.
- You have created a topic.

Note: You can only use a Data Engineering Steam real-time mapping if Data Integration Hub is installed on a UNIX operating system and the Data Integration Hub Data Engineering Integration / Informatica Data Quality component is installed on your system.

Step 1. Access the Create Custom Multi-latency Publication Wizard

Access the create publication wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Custom > Multi-latency**, and then click **Create**.
The **Create Custom Multi-latency Publication** wizard appears.

Step 2. Define Basic Publication Properties

Define publication properties in the **General** page of the publication wizard.

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.

The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.

4. Click **Next**.

The **Processing** page appears.

Step 3. Select Publication Workflow

Select the publication workflow in the **Processing** page of the publication wizard.

1. Select a workflow from the **Custom Mapping** list. Only real-time publication and Data Engineering Streaming workflows are available for selection.
2. If required, enter workflow parameters. Click **Next**.

The **Schedule** page appears.

Step 4. Define Publication Intervals

Define the time interval according to which Data Integration Hub groups the published data and publishes the data in the **Schedule** page of the publication wizard.

1. Select the publication interval in **Publish every**. You can set a value of 10 seconds up to 59 minutes and 50 second, with intervals of 10 seconds.
2. Click **Next**.

The **Permissions** page appears.

Step 5. Assign Publication Permissions

Control access to the publication on the Operation Console by assigning categories to the publication on the **Permissions** page of the publication wizard. If you do not assign categories to the publication, it inherits the categories from the associated application. If there are no categories assigned to the application, all Data Integration Hub users can access the publication.

1. Select one or more categories to assign permission to the publication under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

You can unassign categories from **Selected Categories** to **Available Categories**.

2. Click **Next**.

The **Summary** page appears.

Step 6. Review Publication Settings and Save the Publication

Review the publication settings and save the publication in the **Summary** page of the publication wizard.

1. Review the publication settings.

2. Click **Finish**.

The publication wizard closes. The **Publications** tab of the **Edit Application** page shows the publication you created. The publication is listed in the Publications catalog. If you configured a publication schedule, the publication publishes data according to the defined schedule.

Creating a Custom Cloud Publication

To create a custom publication with a cloud task, perform the following tasks:

1. Access the create publication wizard.
2. Define basic publication properties.
3. Select the Informatica Intelligent Cloud Services task that runs the publication mapping.
4. Define the publication schedule.
5. Optionally, assign permissions to publication. Select the category to which you want to assign the publication. Only the categories assigned to the associated application are available to assign to the publication.
6. Review the publication settings and save the publication.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- The publication mapping and the task are configured in Informatica Intelligent Cloud Services.
- You have created an application.
- You have created a topic.

Step 1. Access the Custom Cloud Publication Wizard

Access the create publication wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Custom > Cloud**, and then click **Create**.
The **Create Custom Cloud Publication** wizard appears.

Step 2. Define Basic Publication Properties

Define publication properties in the **General** page of the publication wizard.

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.

The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.

4. Click **Next**.

The **Processing** page appears.

Step 3. Select Publication Mapping

Select an Informatica Cloud task that defines the publication mapping in the **Processing** page of the publication wizard.

1. Select a task from the **Informatica Cloud Task** list.
2. Click **Next**.

The **Schedule** page appears.

Step 4. Define Publication Schedule

Define the method and the frequency of data publishing in the **Schedule** page of the publication wizard.

1. Select the method and the frequency of data publishing.

Manually or by an external trigger

No schedule. You can use the following methods to run the publication:

- Run manually. Click the Run arrow on the **Publications** page.
- Run by an API. Call a command-line API or a REST API that starts the publication.
- Run by a PowerCenter workflow. Start a publication transformation.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

By schedule

Runs the publication according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the publication in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the publication in intervals of up to 24 hours. You select the number of hours from the list. The publication runs at the beginning of the hour. For example, if you enter 2, the publication runs at 00:00, 02:00, and at consecutive two-hour intervals.
- Daily. Runs the publication at the same hour every day.
- Weekly. Runs the publication every week on one or more days at the same hour.
- Monthly. Runs the publication every month on a specific date or a specific day at the same hour.

Define the publication intervals in the **Repeat running** area.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

After the following publication or subscription run completes

Runs the publication after the run of the publication or subscription that you select here completes.

2. Click **Next**.

The **Permissions** page appears.

Step 5. Assign Publication Permissions

Control access to the publication on the Operation Console by assigning categories to the publication on the **Permissions** page of the publication wizard. If you do not assign categories to the publication, it inherits the categories from the associated application. If there are no categories assigned to the application, all Data Integration Hub users can access the publication.

1. Select one or more categories to assign permission to the publication under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

You can unassign categories from **Selected Categories** to **Available Categories**.

2. Click **Next**.

The **Summary** page appears.

Step 6. Review Publication Settings and Save the Publication

Review the publication settings and save the publication in the **Summary** page of the publication wizard.

1. Review the publication settings.
2. Click **Finish**.

The publication wizard closes. The **Publications** tab of the **Edit Application** page shows the publication you created. The publication is listed in the Publications catalog. If you configured a publication schedule, the publication publishes data according to the defined schedule.

Creating a Custom Pass-through Kafka Publication

This topic describes how to create a publication for an Apache Kafka topic.

Perform the following steps to create a publication for a Kafka topic.

1. Access the create publication page.
2. Define the basic publication properties.
3. Save the publication.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- The publication mapping and the task are configured using any one of the following methods:
 - Data Engineering Streaming mapping
 - PowerCenter mapping with PWX Kafka connector
 - Any external application or tool that can read and write into the Kafka topic
- You have created an application.
- You have created a topic.

Step 1: Access the Create Custom Pass-Through Apache Kafka Publication Page

This topic describes how to create a publication for a Apache Kafka topic.

In order to create an Apache Kafka publication, you must configure the Apache Kafka server broker URL in the System Properties and then create a topic for Apache Kafka.

Perform the following steps to create an Apache Kafka publication:

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the Publications tab and click **New Publication > Custom > Pass-through Kafka > Create**.
The Create Publication page appears.

Step 2. Define the Pass-through Kafka Publication Properties

Enter a short description of the task here (optional).

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Select the topic that you have created to monitor the Apache Kafka streaming.
5. Optionally, assign publication permissions. Select one or more categories to which you want to assign permission to the publication under **Available Categories** and click the right arrow.
6. Click **Save**.

You can associate a maximum of one real-time topic to a publication. If you have created more than one publication, the publication you have created recently is active.

The publication is created and saved in the **Publication** tab.

Creating a Modular Cloud Publication

To create a modular publication with a cloud source, you perform all or some of the following tasks, based on the configuration of the Informatica Intelligent Cloud Services mapping:

1. Access the create publication wizard.
2. Define basic publication properties.
3. Select the publication mapping.
4. If required, configure the publication source.
5. If required, configure the publication target.
6. If the mapping includes parameterized transformations, set the parameter values.

7. If required, configure field mapping.
8. Define the publication schedule.
9. Optionally, assign permissions to publication. Select the category to which you want to assign the publication. Only the categories assigned to the associated application are available to assign to the publication.
10. Review the publication settings and save the publication.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- The publication mapping is configured in Informatica Intelligent Cloud Services.
- You have obtained the required parameter settings.
- You have created an application.
- You have created a topic.

1. Access the Create Modular Cloud Publication Wizard

Access the create publication wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Modular > Cloud**, and then click **Create**.
The **Create Modular Cloud Publication** wizard appears.

Step 2. Define Basic Publication Properties

Define publication properties in the **General** page of the publication wizard.

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Click **Next**.
The **Processing** page appears.

Step 3. Select Publication Mapping

Select an Informatica Cloud publication mapping in the **Processing** page of the publication wizard.

1. Select a mapping from the **Cloud Mapping** list.
2. Click **Next**.
The **Source** page appears.

Step 4. Configure Publication Source

If the publication source is not configured by the mapping, configure the source in the **Source** page of the publication wizard.

1. Configure source settings as applicable.
2. Click **Next**.

The **Target** page appears.

Step 5. Configure Publication Target

If the publication target is not configured by the mapping, configure the target in the **Target** page of the publication wizard. The target of the publication is the Data Integration Hub publication repository.

1. Configure target settings as applicable.
2. Click **Next**.

The **Input Parameters** page appears.

Step 6. Set Publication Parameters

If the mapping contains parameters, the parameters show in the **Input Parameters** page of the publication wizard. Set parameter values as applicable.

1. Click the edit icon next to the parameter for which to define a value.
2. In the **Edit Parameter** dialog box, define the parameter value in the **Expression** area. Click a field in the **Fields** area to add it to the expression.

Click **OK**.

The parameter value shows in the **Input Parameters** page.

3. Repeat steps [1](#) through [2](#) to set the required parameter values.
4. Click **Next**.

The **Field Mapping** page appears.

Step 7. Configure Publication Field Mapping

If field mapping is not configured by the mapping, map source table fields to topic table fields on the **Field Mapping** page of the publication wizard.

1. Click the edit field mapping icon.

The **Edit Field Mapping** dialog box appears.

2. To select to view all the fields, mapped fields, or unmapped fields in the topic table and the source table, select the relevant option from the **Show** lists.
3. To search for fields that are used in the topic table and the source table, enter a string for the field name in **Search Fields**, and then press Enter. The search is not case sensitive. You can search for a substring.
4. To map a source table field to a topic table field, drag the field from the **Default** section to the topic table section.
5. To add an expression to a field, in the **Mapped Field/Expression** column, click the expression icon to open the **Field Expression** dialog box, and then select fields and functions to add to the field. To validate the expression, click **Validate**.
6. Click **OK** to apply the mapping.

The **Edit Field Mapping** dialog box closes.

7. In the **Field Mapping** page, click **Next**.

The **Schedule** page appears.

Step 8. Define Publication Schedule

Define the method and the frequency of data publishing in the **Schedule** page of the publication wizard.

1. Select the method and the frequency of data publishing.

When the file is ready to be published

Runs the publication after the published files are ready, the next time it polls the remote sources. Define the maximal period of time that Data Integration Hub waits for the files to be available in the directory that you defined in the **Source** page in **Wait for all data to be published ... hours**. When the maximal period of time ends, Data Integration Hub discards the file events of the publication and changes the status of the publication event to Error.

Manually or by an external trigger

No schedule. You can use the following methods to run the publication:

- Run manually. Click the Run arrow on the **Publications** page.
- Run by an API. Call a command-line API or a REST API that starts the publication.
- Run by a PowerCenter workflow. Start a publication transformation.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

By schedule

Runs the publication according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the publication in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the publication in intervals of up to 24 hours. You select the number of hours from the list. The publication runs at the beginning of the hour. For example, if you enter 2, the publication runs at 00:00, 02:00, and at consecutive two-hour intervals.
- Daily. Runs the publication at the same hour every day.
- Weekly. Runs the publication every week on one or more days at the same hour.
- Monthly. Runs the publication every month on a specific date or a specific day at the same hour.

Define the publication intervals in the **Repeat running** area.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

After the following publication or subscription run completes

Runs the publication after the run of the publication or subscription that you select here completes.

2. Click **Next**.

The **Permissions** page appears.

Step 9. Assign Publication Permissions

Control access to the publication on the Operation Console by assigning categories to the publication on the **Permissions** page of the publication wizard. If you do not assign categories to the publication, it inherits the

categories from the associated application. If there are no categories assigned to the application, all Data Integration Hub users can access the publication.

1. Select one or more categories to assign permission to the publication under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

You can unassign categories from **Selected Categories** to **Available Categories**.

2. Click **Next**.

The **Summary** page appears.

Step 10. Review Publication Settings and Save the Publication

Review the publication settings and save the publication in the **Summary** page of the publication wizard.

1. Review the publication settings.
2. Click **Finish**.

The publication wizard closes. The **Publications** tab of the **Edit Application** page shows the publication you created. The publication is listed in the Publications catalog. If you configured a publication schedule, the publication publishes data according to the defined schedule.

CHAPTER 5

Creating Subscriptions

This chapter includes the following topics:

- [Creating Subscriptions Overview, 70](#)
- [Creating an Automatic Relational Database Subscription, 71](#)
- [Creating an Automatic Flat File Subscription, 77](#)
- [Creating an Automatic Pass-through File Subscription, 84](#)
- [Creating an Automatic Data-driven Subscription, 88](#)
- [Creating a Custom Batch Subscription, 90](#)
- [Creating a Custom Cloud Subscription, 93](#)
- [Creating a Custom Pass-through Kafka Subscription, 95](#)
- [Creating a Modular Cloud Subscription, 96](#)

Creating Subscriptions Overview

In this chapter, you create subscriptions that consume the published data. You must have first completed the chapter "Creating Topics."

When you create a subscription, you choose the topic to which the application subscribes.

Chapter Objectives

In this chapter, you perform the following tasks:

- Create an automatic subscription with a relational database target.
- Create an automatic subscription with a flat file target.
- Create an automatic subscription with an unstructured file target.
- Create an automatic data-driven subscription that consumes data over a REST API.
- Create a custom subscription with a batch workflow.
- Create a custom subscription with an Informatica Intelligent Cloud Services task.
- Create a modular subscription with a cloud target.

Creating an Automatic Relational Database Subscription

To create a subscription with an automatic mapping and a relational database target, perform the following tasks:

1. Access the create subscription wizard.
2. Define basic subscription properties and select the topic to which you want to subscribe.
3. If you want to run a post-process on the subscription, select the post-process workflow.
4. Optionally, define joins for the subscription.
5. Select the subscription target.
6. Review the publication field mapping, and if required, edit the mapping that Data Integration Hub generates by default.
7. Define a filter for the subscription.
8. Define the subscription schedule.
9. Define the delivery scope and the delivery format for the subscription.
10. Optionally, assign permissions to subscription. Select the category to which you want to assign the subscription. Only the categories assigned to the associated application are available to assign to the subscription.
11. Review the subscription settings and save the subscription.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- A target database exists.
- A connection to the target database is configured in the Data Integration Hub Operation Console.
- You have created an application.
- You have created a topic.
- You have obtained the required parameter settings for the post-process workflow, if applicable.

Step 1. Access the Create Automatic Database Subscription Wizard

Access the create subscription wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application to which you want to deliver content.
The **Edit Application** page appears.
3. Select the **Subscriptions** tab. Click **New**, then select **Automatic > Relational database**, and then click **Create**.
The **Create Automatic Database Subscription** wizard appears.

Step 2. Define Basic Subscription Properties and Select a Topic

Define subscription properties and select a topic in the **General** page of the subscription wizard.

1. Enter the subscription name.
2. Optionally, enter a description of the subscription.
3. Select a mode from the **Mode** list.
4. Select a topic from the **Topic** list.

The **Topic Structure** area shows the structure of the topic from which Data Integration Hub delivers the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.

5. Select whether to apply primary and foreign keys to topic table columns where topic table relations are defined. If joins are manually defined for the subscription, the keys are not applied.
6. Click **Next**.

The **Processing** page appears.

Step 3. Select Post-process Workflow

If you want to run a post-process on the subscription, select the post-process workflow.

1. Select a post-process workflow from the **Post-Processing Workflow** list and determine whether or not to run the workflow if the subscription fails.
2. If the workflow includes parameters, the parameters are shown in the **Post-Processing Parameters** area. If required, set the values of the parameters.
3. Click **Next**.

The **Join** page appears.

Step 4. Define Joins

Pull data from multiple topic tables into a single table by creating a join.

1. Optionally, create joins for the subscription. To create a join, perform the following actions:
 - a. Select the topic tables to join from the **Select Left Table** and **Select Right Table** lists.
 - b. Select the join type from the **Select Join Type** list and then enter the join name in the **Join Name** field. The name that you assign to the join shows in the **Field Mapping** page, where you can map the join to a target table. The new join appears on the **Join** page.
 - c. Select the column to join from the left table and then select the column to join from the right table. To add more columns to the join click the plus sign and then select the columns to join.
 - d. Click **Create Join**.
2. Repeat step [1](#) to create as many joins as required.
3. Click **Next**.

The **Target** page appears.

Step 5. Select Subscription Target

Select the type of the target to which Data Integration Hub writes data and choose the connection to the target in the **Target** page of the subscription wizard.

1. Select the data access connection and click **Next**.
The **Field Mapping** page appears.
2. Optionally, select **Create target tables if target tables do not exist**. If the target tables do not exist in the target database, Data Integration Hub creates them based on table mapping. Data Integration Hub does not create unmapped tables.

Step 6. View and Edit Subscription Field Mapping

View and edit the mapping of topic tables and fields to target tables and fields on the **Field Mapping** page of the subscription wizard. If you created joins in "Step 4. Define Joins", configure field mapping of joins to target tables and fields.

1. Perform the following actions to view and edit the mapping of a topic table or a join to a target table:
 - a. Click the edit table mapping icon.
The **Edit Table Mapping** dialog box appears.
 - b. Search for tables in the database. Note the following guidelines:
 - To search for tables by table name, enter a string in the **Find Target Table** text box and then click **Search**.
 - To search for tables that use a schema other than the default schema, clear the option **Show default schema only**, enter a string in the **Schema** text box, and then click **Search**.
 - To clear the search results and to show only tables that use the default schema, select the option **Show default schema only** and then click **Show All**.
 - The search is not case sensitive.
 - You can search for a substring.The **Search Results** section displays the target tables that match the search string and the schema that each table uses. The name of the topic table for which you are editing the mapping shows in row of the target table to which the topic table is mapped, in the **Mapped To** column.
 - c. To map the topic table to a different target table, click the row of the target table to which you want to map the topic table and then click **Map Target Table**.
The **Edit Table Mapping** dialog box closes. The **Field Mapping** page shows the revised mapping.
2. Perform the following actions on the **Field Mapping** page to view and to map fields in a topic table to fields in a target table:
 - a. Click the edit field mapping icon.
The **Edit Field Mapping** dialog box appears.
 - b. To search for fields that are used in the topic table or in the target table, enter a string for the field name in the **Find Field** text box, and then click **Search**.
Choose **Display unmapped fields only** to show only unmapped fields in the search results.
Topic and target fields that match the search string appear. A green checkmark next to a field name indicates that the field is mapped.
 - c. To map a topic table field to a target table field, drag the field from the **Topic Table** section to the **Target Table** section.

- d. To add an expression to a field, in the **Actions** column, click the expression icon to open the **Expression Editor**, and then select fields and functions to add to the field.
- e. To clear the mapping of a field, in the **Actions** column, click the clear icon.
- f. To view field details, rest on the details icon to the left of a field name.
- g. To revert to the default Data Integration Hub table and field mapping, click **Auto Map**.
- h. Click **OK** to map the fields.

The **Edit Field Mapping** dialog box closes.

3. To revert to the default Data Integration Hub table and field mapping, click **Auto Map** in the **Field Mapping** page.

4. To clear all table and field mappings, click **Clear All** in the **Field Mapping** page.

Note: The subscription must contain at least one mapped topic table.

5. Optionally, select the **Powercenter Target Load Order** to specify the order in which the integration service loads the target tables. By default, this option is disabled.

6. In the **Field Mapping** page, click **Next**.

The **Filter** page appears.

Step 7. Define a Filter

Define the data that the subscription consumes by setting filter conditions on table columns in the **Filter** page of the subscription wizard.

If data relations are defined for the topic to which the subscription subscribes, filters are applied to the tables for which the relations are defined, and the subscriber consumes only the data that is defined by the data relations.

You can use basic expressions and advanced expressions to define filter conditions. You can use more than one method to add filters to a table row.

Basic Expression

Use this method if you want to apply a condition to a table row and the operators and values in the basic condition builder meet your requirements. For example, for an Orders table, add a condition that the date in the ShippedDate column is greater than April 1, 2015.

Advanced Expression

Add filter conditions to tables with PowerCenter expressions. For example, for an Orders table, add the following expression:

```
(ShipCountry='USA') and ((ShipCity='New York') or (ShipCity='Los Angeles'))
```

The filters use an AND logic. If multiple filters exist for a table, Data Integration Hub writes only the rows that meet all of the conditions to the target.

1. Choose the type of expression that you want to create.
2. Select the table to which to apply the filter from the **Select Table** list.
3. Enter the condition parameters according to the method that you selected:
 - Basic expression: go to step [4](#).
 - Advanced expression: go to step [5](#).
4. To create a basic expression, perform the following steps:
 - a. Select the column to filter from the **Select Column** list.

- b. Select the filter operator from the **Select Operator** list. The available operators depend on the type of content in the column.
- c. When **Select Value** is enabled, select or enter a value for the operator. Values of string operators can contain up to 90 digits. Values of numeric operators can contain up to 15 digits.
- d. Click the plus icon to the right of the expression line.

The condition shows in the condition list.

5. To create an advanced expression, perform the following steps:
 - a. Enter a valid PowerCenter expression in the text field. Alternatively, prepare a basic expression, select **Advanced Expression**, and then, in the expression text area, complete the expression. The expression is limited to the selected table. The expression can contain up to 1024 characters.
 - b. Click the plus icon to the right of the expression line.

The condition shows in the condition list.

6. Repeat steps [1](#) through [5](#) to add the required conditions.
7. Click **Next**.

The **Schedule** page appears.

Step 8. Define Subscription Schedule

Define the method and the frequency of the subscription in the **Schedule** page of the subscription wizard.

1. Select the method and the frequency of the subscription.

When published data is ready

Runs the subscription immediately after the published data is ready.

Manually or by an external trigger

No schedule. You can use the following methods to run the subscription:

- Run manually. Click the Run arrow on the **Subscriptions** page.
- Run by an API. Call a command-line API or a REST API that starts the subscription.

By schedule

Runs the subscription according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the subscription in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the subscription in intervals of up to 24 hours. You select the number of hours from the list.
- Daily. Runs the subscription at the same hour every day.
- Weekly. Runs the subscription every week on one or more days at the same hour.
- Monthly. Runs the subscription every month on a specific date or a specific day at the same hour.

Define the delivery intervals in the **Repeat running** area.

After the following publication or subscription run completes

Runs the subscription after the run of the publication or subscription that you select here completes.

2. Optionally, define a retry policy. A retry policy defines the number of times Data Integration Hub retries to run the subscription in case of failure, and the retry interval. The policy does not apply to subscriptions that you run manually.
3. Click **Next**.
The **Delivery** page appears.

Step 9. Define Delivery Options

Define the delivery options of the data to consume in the **Delivery** page of the subscription wizard.

1. For subscriptions that run either manually or by an external trigger or that run by schedule, choose the data delivery scope and delivery format. For subscriptions that run immediately after the published data is ready, go to step [2](#).
 - All available publications. Processes each published data set with a separate subscription mapping.
 - All available publications - aggregated. Groups and processes all published data sets and delivers a single data set.
 - Latest publication only. Delivers only the latest published data set.
2. Choose how Data Integration Hub handles data that exists in the target application.
 - Append the new data to the existing data in the target. Data Integration Hub adds rows to the table.
 - Insert new rows and update remaining rows. Data Integration Hub inserts new rows and then updates rows that exist in the target. Applicable if the database has a primary key.
 - Insert new rows and update changed rows. Data Integration Hub inserts new rows and then updates changed rows that exist in the target. Updating only changed rows slows down system performance. Applicable if the database has a primary key.
 - Delete rows that don't exist in the topic from the target. Applicable if the type of topic to which the subscription subscribes is Full or if the subscription is an unbound subscription, and you choose to append new data or to insert new rows.
 - Overwrite existing data in the target. Data Integration Hub truncates the target tables before it inserts new rows.

Note: If the new subscription subscribes to a topic where data exists and you want to select the option **Insert new rows and update remaining rows** or the option **Insert new rows and update changed rows**, perform the following steps:

1. Select the delivery option **Overwrite existing data in the target**, complete the subscription wizard, and save the subscription.
 2. Access the application to which the subscription subscribes, open the **Subscriptions** tab, and click the Get Previous Publications icon next to the subscription.
 3. In the Get Previous Publications dialing box, select **Include consumed publications** and then click **Run**.
 4. Edit the subscription to apply the required delivery option.
3. Click **Next**.
The **Permissions** page appears.

Step 10. Assign Subscription Permissions

Control access to the subscription on the Operation Console on the **Permissions** page of the subscription wizard. If you do not assign categories to the subscription, it inherits the categories from the associated

application. If there are no categories assigned to the application, all Data Integration Hub users can access the subscription.

1. Select one or more categories to assign permission to the subscription under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

You can unassign categories from **Selected Categories** to **Available Categories**.

2. Click **Next**.

The **Summary** page appears.

Step 10. Review Subscription Settings and Save the Subscription

Review the subscription settings and save the subscription in the **Summary** page of the subscription wizard.

1. Review the subscription settings.
2. Click **Finish**.

The subscription wizard closes. The **Subscription** tab of the **Edit Application** page shows the subscription you created. If you configured a subscription schedule, the subscription consumes data according to the defined schedule.

Creating an Automatic Flat File Subscription

To create a subscription with an automatic mapping and a flat file target, perform the following tasks:

1. Access the create subscription wizard.
2. Define basic subscription properties and select the topic to which you want to subscribe.
3. If you want to run a post-process on the subscription, select the post-process workflow.
4. Optionally, define joins for the subscription.
5. Select a target type and define the structure of the target file .
6. Define a filter for the subscription.
7. Define the subscription schedule.
8. Define the delivery scope and the delivery format for the subscription.
9. Optionally, assign permissions to subscription. Select the category to which you want to assign the subscription. Only the categories assigned to the associated application are available to assign to the subscription.
10. Review the subscription settings and save the subscription.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- A target file exists.
- If you are delivering files to HDFS, the HDFS connection that the workflow uses is configured in the Data Integration Hub Operation Console.
- You have created an application.
- You have created a topic.
- You have obtained the required parameter settings for the post-process workflow, if applicable.

Step 1. Access the Create Automatic Flat File Subscription Wizard

Access the create subscription wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application to which you want to deliver content.
The **Edit Application** page appears.
3. Select the **Subscriptions** tab. Click **New**, then select **Automatic > Flat File**, and then click **Create**.
The **Create Automatic Flat File Subscription** wizard appears.

Step 2. Define Basic Subscription Properties and Select a Topic

Define subscription properties and select a topic in the **General** page of the subscription wizard.

1. Enter the subscription name.
2. Optionally, enter a description of the subscription.
3. Select a mode from the **Mode** list.
4. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic from which Data Integration Hub delivers the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
5. Select whether to apply primary and foreign keys to topic table columns where topic table relations are defined. If joins are manually defined for the subscription, the keys are not applied.
6. Click **Next**.
The **Processing** page appears.

Step 3. Select Post-process Workflow

If you want to run a post-process on the subscription, select the post-process workflow.

1. Select a post-process workflow from the **Post-Processing Workflow** list and determine whether or not to run the workflow if the subscription fails.
2. If the workflow includes parameters, the parameters are shown in the **Post-Processing Parameters** area. If required, set the values of the parameters.
3. Click **Next**.
The **Join** page appears.

Step 4. Define Joins

Pull data from multiple topic tables into a single table by creating a join.

1. Optionally, create joins for the subscription. To create a join, perform the following actions:
 - a. Select the topic tables to join from the **Select Left Table** and **Select Right Table** lists.
 - b. Select the join type from the **Select Join Type** list and then enter the join name in the **Join Name** field. The name that you assign to the join shows in the **Field Mapping** page, where you can map the join to a target table. The new join appears on the **Join** page.

- c. Select the column to join from the left table and then select the column to join from the right table. To add more columns to the join click the plus sign and then select the columns to join.
 - d. Click **Create Join**.
2. Repeat step [1](#) to create as many joins as required.
 3. Click **Next**.
The **Target** page appears.

Step 5. Select and Configure Subscription Target

Select the type of the target to which Data Integration Hub writes data and define the location and the structure of the target file in the **Target** page of the subscription wizard.

1. Select the target type from the **Target Type** list.
If you select the HDFS target type, the **HDFS Connection** field appears on the page.
2. If you select the HDFS target type, select a connection from the **HDFS Connection** list.
3. To use file transfer to deliver files to remote servers, select **Use File Transfer**.
The **Connection** field appears on the page.
4. If you selected to use file transfer, select a connection from the **Connection** list.
5. Define the location and the structure of the target file:

Directory

Location where Data Integration Hub creates the target file or files. Data Integration Hub generates a file for each table in the topic.

Target Filename Pattern

Pattern of the names of the target files. For more information, see [“Target File Name Pattern for Flat Files” on page 80](#).

Code Page

Character encoding used in the file.

Use column names as the first line

Use the first line in the source data as the table header.

Delimiter

Delimiter used in the file to separate between columns. Select a predefined delimiter or select **Custom** and then define a custom delimiter. For information about the supported column delimiters, see the sections about delimited files in the *PowerCenter Designer Guide*.

Text Qualifier

Optional. Symbols used in the file to enclose a string.

Thousands Separator

Optional. Symbol used in the file as a thousands separator.

Decimal Separator

Symbol used in the file as a decimal separator.

Datetime Format

Date and time format used in the file. Select a predefined format or select **Other** to define a custom format. For information about the supported datetime formats, see the chapter about dates in the *PowerCenter Transformation Language Reference*.

Note: The datetime format can contain up to 50 characters.

6. Click **Next**.

The **Field Mapping** page appears.

Target File Name Pattern for Flat Files

When you define a flat file target, the pattern of the file name, which you define in the **Target Filename Pattern** field, can contain the following variables:

Variable	Description
(\$Table_Name)	Name of the table in the topic that contains the data. Use this variable in topics that include more than one table to create an output file for each table column.
(\$Time_Stamp)	Date and time when the subscription runs, in the following format: yyyy-mm-dd-hh24_mi_ss
(\$Publication_Instance_Date)	Date and time when the publication instance is created, in the following format: yyyy-mm-dd-hh24_mi_ss Use this variable to differentiate between multiple publication instances and to sort publications by date.

For example, the file name pattern `file_($Table_Name)_($Publication_Instance_Date).txt` creates the following file name:

```
file_ORDERS_2015-03-27-12_43_26.txt
```

Step 6. View and Edit Subscription Field Mapping

View and edit the mapping of topic tables and fields to target tables and fields on the **Field Mapping** page of the subscription wizard. If you created joins in "Step 4. Define Joins", configure field mapping of joins to target tables and fields.

1. Perform the following actions on the **Field Mapping** page to view and to map fields in a topic table to fields in a target table:

- a. Click the edit field mapping icon.

The **Edit Field Mapping** dialog box appears.

- b. To search for fields that are used in the topic table or in the target table, enter a string for the field name in the **Find Field** text box, and then click **Search**.

Choose **Display unmapped fields only** to show only unmapped fields in the search results.

Topic and target fields that match the search string appear. A green checkmark next to a field name indicates that the field is mapped.

- c. To map a topic table field to a target table field, drag the field from the **Topic Table** section to the **Target Table** section.
- d. To add an expression to a field, in the **Actions** column, click the expression icon to open the **Expression Editor**, and then select fields and functions to add to the field.
- e. To clear the mapping of a field, in the **Actions** column, click the clear icon.
- f. To view field details, rest on the details icon to the left of a field name.
- g. To revert to the default Data Integration Hub table and field mapping, click **Auto Map**.
- h. Click **OK** to map the fields.

The **Edit Field Mapping** dialog box closes.

2. To revert to the default Data Integration Hub table and field mapping, click **Auto Map** in the **Field Mapping** page.
3. To clear all table and field mappings, click **Clear All** in the **Field Mapping** page.
Note: The subscription must contain at least one mapped topic table.
4. In the **Field Mapping** page, click **Next**.

The **Filter** page appears.

Step 7. Define a Filter

Define the data that the subscription consumes by setting filter conditions on table columns in the **Filter** page of the subscription wizard.

If data relations are defined for the topic to which the subscription subscribes, filters are applied to the tables for which the relations are defined, and the subscriber consumes only the data that is defined by the data relations.

You can use basic expressions and advanced expressions to define filter conditions. You can use more than one method to add filters to a table row.

Basic Expression

Use this method if you want to apply a condition to a table row and the operators and values in the basic condition builder meet your requirements. For example, for an Orders table, add a condition that the date in the ShippedDate column is greater than April 1, 2015.

Advanced Expression

Add filter conditions to tables with PowerCenter expressions. For example, for an Orders table, add the following expression:

```
(ShipCountry='USA') and ((ShipCity='New York') or (ShipCity='Los Angeles'))
```

The filters use an AND logic. If multiple filters exist for a table, Data Integration Hub writes only the rows that meet all of the conditions to the target.

1. Choose the type of expression that you want to create.
2. Select the table to which to apply the filter from the **Select Table** list.
3. Enter the condition parameters according to the method that you selected:
 - Basic expression: go to step [4](#).
 - Advanced expression: go to step [5](#).
4. To create a basic expression, perform the following steps:
 - a. Select the column to filter from the **Select Column** list.

- b. Select the filter operator from the **Select Operator** list. The available operators depend on the type of content in the column.
- c. When **Select Value** is enabled, select or enter a value for the operator. Values of string operators can contain up to 90 digits. Values of numeric operators can contain up to 15 digits.
- d. Click the plus icon to the right of the expression line.

The condition shows in the condition list.

5. To create an advanced expression, perform the following steps:
 - a. Enter a valid PowerCenter expression in the text field. Alternatively, prepare a basic expression, select **Advanced Expression**, and then, in the expression text area, complete the expression. The expression is limited to the selected table. The expression can contain up to 1024 characters.
 - b. Click the plus icon to the right of the expression line.

The condition shows in the condition list.

6. Repeat steps [1](#) through [5](#) to add the required conditions.
7. Click **Next**.

The **Schedule** page appears.

Step 8. Define Subscription Schedule

Define the method and the frequency of the subscription in the **Schedule** page of the subscription wizard.

1. Select the method and the frequency of the subscription.

When published data is ready

Runs the subscription immediately after the published data is ready.

Manually or by an external trigger

No schedule. You can use the following methods to run the subscription:

- Run manually. Click the Run arrow on the **Subscriptions** page.
- Run by an API. Call a command-line API or a REST API that starts the subscription.

By schedule

Runs the subscription according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the subscription in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the subscription in intervals of up to 24 hours. You select the number of hours from the list.
- Daily. Runs the subscription at the same hour every day.
- Weekly. Runs the subscription every week on one or more days at the same hour.
- Monthly. Runs the subscription every month on a specific date or a specific day at the same hour.

Define the delivery intervals in the **Repeat running** area.

After the following publication or subscription run completes

Runs the subscription after the run of the publication or subscription that you select here completes.

2. Optionally, define a retry policy. A retry policy defines the number of times Data Integration Hub retries to run the subscription in case of failure, and the retry interval. The policy does not apply to subscriptions that you run manually.
3. Click **Next**.
The **Delivery** page appears.

Step 9. Define Delivery Options

Define the delivery options of the data to consume in the **Delivery** page of the subscription wizard.

1. For subscriptions that run either manually or by an external trigger or that run by schedule, choose the data delivery scope and delivery format. For subscriptions that run immediately after the published data is ready, go to step [2](#).
 - All available publications. Processes each published data set with a separate subscription mapping.
 - All available publications - aggregated. Groups and processes all published data sets and delivers a single data set.
 - Latest publication only. Delivers the latest published data set.
2. Choose how Data Integration Hub handles data that exists in the target application. The selection is disabled for subscriptions that subscribe to a topic with a Delta publication type and are configured to consume all available publications, and for subscriptions that use file transfer.
 - Append the new data to the existing data in the target. Data Integration Hub appends records to the file.
 - Overwrite existing data in the target. Data Integration Hub overwrites the file.
3. Click **Next**.
The **Permissions** page appears.

Step 10. Assign Subscription Permissions

Control access to the subscription on the Operation Console on the **Permissions** page of the subscription wizard. If you do not assign categories to the subscription, it inherits the categories from the associated application. If there are no categories assigned to the application, all Data Integration Hub users can access the subscription.

1. Select one or more categories to assign permission to the subscription under **Available Categories** and click the right arrow.
The category appears under **Selected Categories**.
You can unassign categories from **Selected Categories** to **Available Categories**.
2. Click **Next**.
The **Summary** page appears.

Step 10. Review Subscription Settings and Save the Subscription

Review the subscription settings and save the subscription in the **Summary** page of the subscription wizard.

1. Review the subscription settings.
2. Click **Finish**.
The subscription wizard closes. The **Subscription** tab of the **Edit Application** page shows the subscription you created. If you configured a subscription schedule, the subscription consumes data according to the defined schedule.

Creating an Automatic Pass-through File Subscription

To create a subscription with an automatic mapping and a pass-through file target, perform the following tasks:

1. Access the create subscription wizard.
2. Define basic subscription properties and select the topic to which you want to subscribe.
3. If you want to run a post-process on the subscription, select the post-process workflow.
4. Configure the subscription target.
5. Define a filter for the subscription.
6. Define the subscription schedule.
7. Define the delivery scope of the subscription.
8. Optionally, assign permissions to subscription. Select the category to which you want to assign the subscription. Only the categories assigned to the associated application are available to assign to the subscription.
9. Review the subscription settings and save the subscription.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- A target file exists.
- You have created an application.
- You have created a topic.
- You have obtained the required parameter settings for the post-process workflow, if applicable.

Step 1. Access the Create Automatic Pass-through File Subscription Wizard

Access the create subscription wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application to which you want to deliver content.
The **Edit Application** page appears.
3. Select the **Subscriptions** tab. Click **New**, then select **Automatic > Pass-through File**, and then click **Create**.
The **Create Automatic Pass-through File Subscription** wizard appears.

Step 2. Define Basic Subscription Properties and Select a Topic

Define subscription properties and select a topic in the **General** page of the subscription wizard.

1. Enter the subscription name.
2. Optionally, enter a description of the subscription.
3. Select a mode from the **Mode** list.

4. Select a topic from the **Topic** list.

The **Topic Structure** area shows the structure of the topic from which Data Integration Hub delivers the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.

5. Click **Next**.

The **Processing** page appears.

Step 3. Select Post-process Workflow

If you want to run a post-process on the subscription, select the post-process workflow.

1. Select a post-process workflow from the **Post-Processing Workflow** list and determine whether or not to run the workflow if the subscription fails.
2. If the workflow includes parameters, the parameters are shown in the **Post-Processing Parameters** area. If required, set the values of the parameters.
3. Click **Next**.

The **Target** page appears.

Step 4. Configure Subscription Target

Configure the target to which Data Integration Hub delivers files in the **Target** page of the subscription wizard.

1. To use file transfer to deliver files to remote servers, select **Use File Transfer**.

The **Connection** field appears on the page.

2. If you selected to use file transfer, select a connection from the **Connection** list.
3. Enter the location of the file or files that Data Integration Hub delivers in the **Directory** field.
4. Click the edit icon next to a topic table to assign a target file to the table. By default, Data Integration Hub assigns the name of the topic table as the file name, in the following format: `<topic_table_name>_($sequence)`. For example, if the name of the topic table is `orderId`, the default name of the file that is assigned to the table is `orderId_($sequence)`. You can edit the `<topic_table_name>` section of file name.

The pattern of the file name can contain the following variables:

Variable	Description
<code>(\$sequence)</code>	Required. Use this variable to differentiate between multiple publication instances and to sort publications by order.
<code>(\$Time_Stamp)</code>	Optional. Date and time when the subscription runs, in the following format: <code>yyyy-mm-dd-hh24_mi_ss</code>

For example:

The file name pattern `file_($Time_Stamp)_($sequence).txt` creates the following file name:

`file_2015-03-27-12_43_26_1260.txt`

You must assign a target file for at least one topic table.

5. Repeat step [4](#) for each table topic that appears on the **Target** page.

6. Click **Next**.

The **Filter** page appears.

Step 5. Define a Filter

Define the data that the subscription consumes by setting filter conditions on the file metadata.

You can use basic expressions and advanced expressions to define filter conditions. You can use more than one method to add filters to a table row.

Basic Expression

Use this method if you want to apply a condition to a table row and the operators and values in the basic condition builder meet your requirements. For example, for an Orders table, add a condition that the date in the ShippedDate column is greater than April 1, 2015.

Advanced Expression

Add filter conditions to tables with PowerCenter expressions. For example, for an Orders table, add the following expression:

```
(ShipCountry='USA') and ((ShipCity='New York') or (ShipCity='Los Angeles'))
```

The filters use an AND logic. If multiple filters exist for a table, Data Integration Hub writes only the rows that meet all of the conditions to the target.

1. Choose the type of expression that you want to create.
2. Select the table to which to apply the filter from the **Select Table** list.
3. Enter the condition parameters according to the method that you selected:
 - Basic expression: go to step [4](#).
 - Advanced expression: go to step [5](#).
4. To create a basic expression, perform the following steps:
 - a. Select the column to filter from the **Select Column** list.
 - b. Select the filter operator from the **Select Operator** list. The available operators depend on the type of content in the column.
 - c. When **Select Value** is enabled, select or enter a value for the operator. Values of string operators can contain up to 90 digits. Values of numeric operators can contain up to 15 digits.
 - d. Click the plus icon to the right of the expression line.

The condition shows in the condition list.

5. To create an advanced expression, perform the following steps:
 - a. Enter a valid PowerCenter expression in the text field. Alternatively, prepare a basic expression, select **Advanced Expression**, and then, in the expression text area, complete the expression. The expression is limited to the selected table. The expression can contain up to 1024 characters.
 - b. Click the plus icon to the right of the expression line.
6. Repeat steps [1](#) through [5](#) to add the required conditions.
7. Click **Next**.

The **Schedule** page appears.

Step 6. Define Subscription Schedule

Define the method and the frequency of the subscription in the **Schedule** page of the subscription wizard.

1. Select the method and the frequency of the subscription.

When published data is ready

Runs the subscription immediately after the published data is ready.

Manually or by an external trigger

No schedule. You can use the following methods to run the subscription:

- Run manually. Click the Run arrow on the **Subscriptions** page.
- Run by an API. Call a command-line API or a REST API that starts the subscription.

By schedule

Runs the subscription according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the subscription in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the subscription in intervals of up to 24 hours. You select the number of hours from the list.
- Daily. Runs the subscription at the same hour every day.
- Weekly. Runs the subscription every week on one or more days at the same hour.
- Monthly. Runs the subscription every month on a specific date or a specific day at the same hour.

Define the delivery intervals in the **Repeat running** area.

After the following publication or subscription run completes

Runs the subscription after the run of the publication or subscription that you select here completes.

2. Optionally, define a retry policy. A retry policy defines the number of times Data Integration Hub retries to run the subscription in case of failure, and the retry interval. The policy does not apply to subscriptions that you run manually.

3. Click **Next**.

The **Delivery** page appears.

Step 7. Define Delivery Scope

Define the delivery scope of the data to consume in the **Delivery** page of the subscription wizard.

1. Choose the data delivery scope:

- All available publications. Processes each published data set with a separate subscription mapping.
- Latest publication only. Delivers only the latest published data set.

2. Click **Next**.

The **Permissions** page appears.

Step 8. Assign Subscription Permissions

Control access to the subscription on the Operation Console on the **Permissions** page of the subscription wizard. If you do not assign categories to the subscription, it inherits the categories from the associated

application. If there are no categories assigned to the application, all Data Integration Hub users can access the subscription.

1. Select one or more categories to assign permission to the subscription under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

You can unassign categories from **Selected Categories** to **Available Categories**.

2. Click **Next**.

The **Summary** page appears.

Step 9. Review Subscription Settings and Save the Subscription

Review the subscription settings and save the subscription in the **Summary** page of the subscription wizard.

1. Review the subscription settings.
2. Click **Finish**.

The subscription wizard closes. The **Subscription** tab of the **Edit Application** page shows the subscription you created. If you configured a subscription schedule, the subscription consumes data according to the defined schedule.

Creating an Automatic Data-driven Subscription

To create an automatic data-driven subscription, perform the following tasks:

1. Access the create subscription page.
2. Configure subscription properties and select the topic from which to consume the data.
After you configure the subscription properties, you can copy the following URLs from the create subscription page:
 - REST URL of the API. Use this URL to subscribe to the data.
 - REST URL of the Swagger structure for the topic from which the subscription consumes data.You use the URLs when you create the request that runs the subscription.
3. Optionally, assign permissions to subscription. Select the category to which you want to assign the subscription. Only the categories assigned to the associated application are available to assign to the subscription.
4. Create a request to run the subscription.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- You have created an application.
- You have created a topic.

Step 1. Access the Create Automatic Data-driven Subscription Page

Access the create subscription page in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application to which you want to deliver content.
The **Edit Application** page appears.
3. Select the **Subscriptions** tab. Click **New**, then select **Automatic > Data-driven**, and then click **Create**.
The **Create Automatic Data-driven Subscription** page appears.

Step 2. Define Subscription Properties

Define subscription properties and select a topic in the create subscription page.

1. Enter the subscription name.
The following URLs are automatically updated:
 - REST URL of the API. Use this URL to subscribe to the data.
 - REST URL of the Swagger structure for the topic from which the subscription consumes data.You use the URLs when you create the request that runs the subscription.
2. Optionally, enter a description of the subscription.
3. Select a mode from the **Mode** list.
4. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic from which Data Integration Hub delivers the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
5. Enter the URL to where Data Integration Hub sends notifications when data is ready to consume.
6. Optionally, assign subscription permissions. Select one or more categories to which you want to assign permission to the subscription under **Available Categories** and click the right arrow.
7. Click **Save**.

Step 3. Create a Request to Run the Subscription

To create a request to run the subscription, you have to copy the URL of the REST API and the URL of the API Swagger file from the subscription that you created in [“Step 2. Define Subscription Properties” on page 89](#).

- ▶ Create a POST request with the following details:

- Request URL: REST API URL. For example:

```
http://hostname:18080/dih-console/api/v1/subscription/MySubscription/data
```

Where `MySubscription` is the subscription name.

- Request body:

```
{"batchSize":<records_max_number> } :
```

Where `records_max_number` is the maximal number of records from a table that can be put on the HTTP response stream before flushing it.

For example:

```
{"batchSize":5 } :
```

If `records_max_number` is left empty, Data Integration Hub applies the default batch size of 500 records with a limit of 5000 records.

Creating a Custom Batch Subscription

To create a subscription with a custom mapping that uses a batch workflow, perform the following tasks:

1. Access the create subscription wizard.
2. Define basic subscription properties and select the topic to which you want to subscribe.
3. Select the subscription workflow. If you use file transfer, select the connection to the target to which Data Integration Hub writes the files. If you want to run a post-process on the subscription, select the post-process workflow.
4. If the workflow includes parameters, set subscription parameter values.
5. Define the subscription schedule.
6. Optionally, assign permissions to subscription. Select the category to which you want to assign the subscription. Only the categories assigned to the associated application are available to assign to the subscription.
7. Review the subscription settings and save the subscription.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- The subscription workflow, and the post-process workflow, if applicable, are configured in the Data Integration Hub Operation Console.
- If you use file transfer, the connection that the workflow uses is configured in the Data Integration Hub Operation Console.
- You have obtained the required parameter settings for the subscription workflow and, if applicable, for the post-process workflow.
- You have created an application.
- You have created a topic.

Step 1. Access the Create Custom Batch Subscription Wizard

Access the create subscription wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application to which you want to deliver content.
The **Edit Application** page appears.
3. Select the **Subscriptions** tab. Click **New**, then select **Custom > Batch**, and then click **Create**.
The **Create Custom Batch Subscription** wizard appears.

Step 2. Define Basic Subscription Properties and Select Topics

Define subscription properties and select topics in the **General** page of the subscription wizard.

1. Enter the subscription name.
2. Optionally, enter a description of the subscription.
3. Select a mode from the **Mode** list.
4. Select the topic to which you want to subscribe and then click **Add Topic**.
The topic appears in the topic table.
5. Optionally, subscribe to additional topics and create a compound subscription.

In compound subscriptions that consume data from multiple topics, topic information includes a **Mandatory** option that determines whether the topic is mandatory or not. Data Integration Hub waits for all mandatory topics to be available for consumption before it runs the subscription. By default, all topics in a compound subscription are mandatory. If a topic is not mandatory for the subscription, clear the **Mandatory** option for the topic.

6. If the subscription is a compound subscription, in the **Wait for all mandatory topics to be available for consumption for ... minutes** field, specify the number of minutes to wait from the time the first topic is ready to consume until all mandatory topics are ready to consume.
7. Click **Next**.
The **Processing** page appears.

Step 3. Select Subscription Workflows

Select a Data Integration Hub subscription workflow in the **Processing** page of the subscription wizard. If you want to run a post-process on the subscription, select the post-process workflow. If required, set parameter values.

1. Select a subscription workflow from the **Custom Mapping** list. Only subscription workflows that are based on a batch workflow are available for selection.
If the workflow supports file transfer, the **Use File Transfer** option is enabled.
2. To use file transfer, select the **Use File Transfer** option and then select the connection to the target to which Data Integration Hub writes the files in the **Connection** field.
3. To run a post-process on the subscription, select a post-process workflow from the **Pos-process Mapping** list and determine whether or not to run the workflow if the subscription fails.
4. If the workflows include parameters, the parameters are shown on the page. Set the values of the parameters in the **Subscription Parameters** area and the **Post Process Parameters** area, as applicable.
5. Click **Next**.
The **Schedule** page appears.

Step 4. Define Subscription Schedule

Define the method and the frequency of the subscription in the **Schedule** page of the subscription wizard. If you create a compound subscription, where the subscription consumes data sets from multiple topics, you can only choose to consume data when it is published.

1. Select the method and the frequency of the subscription.

When published data is ready

Runs the subscription immediately after the published data is ready.

Note: This option is not applicable to subscriptions that subscribe to a Hadoop-based topic.

Manually or by an external trigger

No schedule. You can use the following methods to run the subscription:

- Run manually. Click the Run arrow on the **Subscriptions** page.
- Run by an API. Call a command-line API or a REST API that starts the subscription.

By schedule

Runs the subscription according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the subscription in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the subscription in intervals of up to 24 hours. You select the number of hours from the list.
- Daily. Runs the subscription at the same hour every day.
- Weekly. Runs the subscription every week on one or more days at the same hour.
- Monthly. Runs the subscription every month on a specific date or a specific day at the same hour.

Define the delivery intervals in the **Repeat running** area.

After the following publication or subscription run completes

Runs the subscription after the run of the publication or subscription that you select here completes.

2. Click **Next**.

The **Permissions** page appears.

Step 5. Assign Subscription Permissions

Control access to the subscription on the Operation Console on the **Permissions** page of the subscription wizard. If you do not assign categories to the subscription, it inherits the categories from the associated application. If there are no categories assigned to the application, all Data Integration Hub users can access the subscription.

1. Select one or more categories to assign permission to the subscription under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

You can unassign categories from **Selected Categories** to **Available Categories**.

2. Click **Next**.

The **Summary** page appears.

Step 6. Review Subscription Settings and Save the Subscription

Review the subscription settings and save the subscription in the **Summary** page of the subscription wizard.

1. Review the subscription settings.

2. Click **Finish**.

The subscription wizard closes. The **Subscription** tab of the **Edit Application** page shows the subscription you created. If you configured a subscription schedule, the subscription consumes data according to the defined schedule.

Creating a Custom Cloud Subscription

To create a subscription with a custom mapping and a cloud task, perform the following tasks:

1. Access the create subscription wizard.
2. Define basic subscription properties and select the topic to which you want to subscribe.
3. Select the Informatica Intelligent Cloud Services task that runs the subscription mapping.
4. Define the subscription schedule.
5. Optionally, assign permissions to subscription. Select the category to which you want to assign the subscription. Only the categories assigned to the associated application are available to assign to the subscription.
6. Review the subscription settings and save the subscription.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- The subscription mapping and the task are configured in Informatica Intelligent Cloud Services.
- You have created an application.
- You have created a topic.

Step 1. Access the Custom Cloud Subscription Wizard

Access the create subscription wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application to which you want to deliver content.
The **Edit Application** page appears.
3. Select the **Subscriptions** tab. Click **New**, then select **Custom > Cloud**, and then click **Create**.
The **Create Custom Cloud Subscription** wizard appears.

Step 2. Define Basic Subscription Properties and Select Topics

Define subscription properties and select topics in the **General** page of the subscription wizard.

1. Enter the subscription name.
2. Optionally, enter a description of the subscription.
3. Select a mode from the **Mode** list.
4. Select the topic to which you want to subscribe and then click **Add Topic**.
The topic appears in the topic table.

5. Optionally, subscribe to additional topics and create a compound subscription.
In compound subscriptions that consume data from multiple topics, topic information includes a **Mandatory** option that determines whether the topic is mandatory or not. Data Integration Hub waits for all mandatory topics to be available for consumption before it runs the subscription. By default, all topics in a compound subscription are mandatory. If a topic is not mandatory for the subscription, clear the **Mandatory** option for the topic.
6. If the subscription is a compound subscription, in the **Wait for all mandatory topics to be available for consumption for ... minutes** field, specify the number of minutes to wait from the time the first topic is ready to consume until all mandatory topics are ready to consume.
7. Click **Next**.
The **Processing** page appears.

Step 3. Select Subscription Mappings

Select an Informatica Cloud task that defines the subscription mapping in the **Processing** page of the subscription wizard. If you want to run a post-process on the subscription, select the post-process workflow.

1. Select a task from the **Informatica Cloud Task** list.
2. Click **Next**.
The **Schedule** page appears.

Step 4. Define Subscription Schedule

Define the method and the frequency of the subscription in the **Schedule** page of the subscription wizard. If you create a compound subscription, where the subscription consumes data sets from multiple topics, you can only choose to consume data when it is published.

1. Select the method and the frequency of the subscription.

When published data is ready

Runs the subscription immediately after the published data is ready.

Note: This option is not applicable to subscriptions that subscribe to a Hadoop-based topic.

Manually or by an external trigger

No schedule. You can use the following methods to run the subscription:

- Run manually. Click the Run arrow on the **Subscriptions** page.
- Run by an API. Call a command-line API or a REST API that starts the subscription.

By schedule

Runs the subscription according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the subscription in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the subscription in intervals of up to 24 hours. You select the number of hours from the list.
- Daily. Runs the subscription at the same hour every day.
- Weekly. Runs the subscription every week on one or more days at the same hour.
- Monthly. Runs the subscription every month on a specific date or a specific day at the same hour.

Define the delivery intervals in the **Repeat running** area.

After the following publication or subscription run completes

Runs the subscription after the run of the publication or subscription that you select here completes.

2. Click **Next**.

The **Permissions** page appears.

Step 5. Assign Subscription Permissions

Control access to the subscription on the Operation Console on the **Permissions** page of the subscription wizard. If you do not assign categories to the subscription, it inherits the categories from the associated application. If there are no categories assigned to the application, all Data Integration Hub users can access the subscription.

1. Select one or more categories to assign permission to the subscription under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

You can unassign categories from **Selected Categories** to **Available Categories**.

2. Click **Next**.

The **Summary** page appears.

Step 6. Review Subscription Settings and Save the Subscription

Review the subscription settings and save the subscription in the **Summary** page of the subscription wizard.

1. Review the subscription settings.
2. Click **Finish**.

The subscription wizard closes. The **Subscription** tab of the **Edit Application** page shows the subscription you created. If you configured a subscription schedule, the subscription consumes data according to the defined schedule.

Creating a Custom Pass-through Kafka Subscription

This topic describes how to create a subscription for an Apache Kafka topic.

Perform the following steps to create a subscription for a Kafka topic.

1. Access the create subscription page.
2. Define the basic subscription properties.
3. Save the subscription.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- The publication mapping and the task are configured in Informatica Intelligent Cloud Services.
- You have created an application.
- You have created a topic.

Step1. Access the Create Custom Pass-through Kafka Subscription Page

Perform the following tasks to create a custom pass-through Kafka subscription:

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Click the **Subscriptions** tab and click **New** and the select **Custom > Pass-through Kafka > Create**.
The custom pass-through Kafka subscription wizard appears.

Step2. Define Basic Subscription Properties

Perform the following tasks to define the basic subscription properties:

1. Enter the subscription name.
2. Optionally, enter a description of the subscription.
3. Select the topic to which you want to subscribe and then click **Add Topic**.
4. Optionally, assign subscription permissions if the associated application is assigned with categories. Select one or more categories to which you want to assign permission to the subscription under **Available Categories** and click right arrow.
5. Click **Save**.
The subscription is saved in the Subscriptions tab.

Creating a Modular Cloud Subscription

To create a subscription with a modular mapping and a cloud target, you perform all or some of the following tasks, based on the configuration of the subscription mapping:

1. Access the create subscription wizard.
2. Define basic subscription properties and select the topic to which you want to subscribe.
3. Select the subscription mapping.
4. If required, configure the subscription source.
5. If required, configure the subscription target.
6. If the mapping includes parameterized transformations, set the parameter values.
7. If required, configure field mapping.
8. Define the subscription schedule.
9. Optionally, assign permissions to subscription. Select the category to which you want to assign the subscription. Only the categories assigned to the associated application are available to assign to the subscription.
10. Review the subscription settings and save the subscription.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- The subscription mapping is configured in Informatica Intelligent Cloud Services.
- You have obtained the required parameter settings.
- You have created an application.
- You have created a topic.

Step 1. Access the Modular Cloud Subscription Wizard

Access the create subscription wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application to which you want to deliver content.
The **Edit Application** page appears.
3. Select the **Subscriptions** tab. Click **New**, then select **Modular > Cloud**, and then click **Create**.
The **Create Modular Cloud Subscription** wizard appears.

Step 2. Define Basic Subscription Properties and Select a Topic

Define subscription properties and select a topic in the **General** page of the subscription wizard.

1. Enter the subscription name.
2. Optionally, enter a description of the subscription.
3. Select a mode from the **Mode** list.
4. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic from which Data Integration Hub delivers the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
5. Select whether to apply primary and foreign keys to topic table columns where topic table relations are defined. If joins are manually defined for the subscription, the keys are not applied.
6. Click **Next**.
The **Processing** page appears.

Step 3. Select Subscription Mapping

Select an Informatica Cloud subscription mapping in the **Processing** page of the subscription wizard.

1. Select a mapping from the **Cloud Mapping** list.
2. Click **Next**.
The **Source** page appears.

Step 4. Configure Subscription Source

If the subscription source is not configured by the mapping, configure the source in the **Source** page of the subscription wizard. The source of the subscription is the Data Integration Hub publication repository.

1. Configure source settings as applicable.
2. Click **Next**.

The **Target** page appears.

Step 5. Configure Subscription Target

If the subscription target is not configured by the mapping, configure the target in the **Target** page of the subscription wizard.

1. Configure target settings as applicable.
2. Click **Next**.

The **Input Parameters** page appears.

Step 6. Set Subscription Parameters

If the mapping contains parameters, the parameters show in the **Input Parameters** page of the subscription wizard. Set parameter values as applicable.

1. Click the edit icon next to the parameter for which to define a value.
2. In the **Edit Parameter** dialog box, define the parameter value in the **Expression** area. Click a field in the **Fields** area to add it to the expression.

Click **OK**.

The parameter value shows in the **Input Parameters** page.

3. Repeat steps [1](#) through [2](#) to set the required parameter values.
4. Click **Next**.

The **Field Mapping** page appears.

Step 7. Configure Subscription Field Mapping

If field mapping is not configured by the mapping, map topic table fields to target table fields on the **Field Mapping** page of the subscription wizard.

1. Click the edit field mapping icon.
The **Edit Field Mapping** dialog box appears.
2. To select to view all the fields, mapped fields, or unmapped fields in the topic table and the target table, select the relevant option from the **Show** lists.
3. To search for fields that are used in the topic table and the target table, enter a string for the field name in **Search Fields**, and then press Enter. The search is not case sensitive. You can search for a substring.
4. To map a topic table field to a target table field, drag the field from the **Default** section to the target table section.
5. To add an expression to a field, in the **Mapped Field/Expression** column, click the expression icon to open the **Field Expression** dialog box, and then select fields and functions to add to the field. To validate the expression, click **Validate**.
6. Click **OK** to apply the mapping.

The **Edit Field Mapping** dialog box closes.

7. In the **Field Mapping** page, click **Next**.

The **Schedule** page appears.

Step 8. Define Subscription Schedule

Define the method and the frequency of the subscription in the **Schedule** page of the subscription wizard.

1. Select the method and the frequency of the subscription.

When published data is ready

Runs the subscription immediately after the published data is ready.

Manually or by an external trigger

No schedule. You can use the following methods to run the subscription:

- Run manually. Click the Run arrow on the **Subscriptions** page.
- Run by an API. Call a command-line API or a REST API that starts the subscription.

By schedule

Runs the subscription according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the subscription in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the subscription in intervals of up to 24 hours. You select the number of hours from the list.
- Daily. Runs the subscription at the same hour every day.
- Weekly. Runs the subscription every week on one or more days at the same hour.
- Monthly. Runs the subscription every month on a specific date or a specific day at the same hour.

Define the delivery intervals in the **Repeat running** area.

After the following publication or subscription run completes

Runs the subscription after the run of the publication or subscription that you select here completes.

2. Click **Next**.

The **Permissions** page appears.

Step 9. Assign Subscription Permissions

Control access to the subscription on the Operation Console on the **Permissions** page of the subscription wizard. If you do not assign categories to the subscription, it inherits the categories from the associated application. If there are no categories assigned to the application, all Data Integration Hub users can access the subscription.

1. Select one or more categories to assign permission to the subscription under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

You can unassign categories from **Selected Categories** to **Available Categories**.

2. Click **Next**.

The **Summary** page appears.

Step 10. Review Subscription Settings and Save the Subscription

Review the subscription settings and save the subscription in the **Summary** page of the subscription wizard.

1. Review the subscription settings.
2. Click **Finish**.

The subscription wizard closes. The **Subscription** tab of the **Edit Application** page shows the subscription you created. If you configured a subscription schedule, the subscription consumes data according to the defined schedule.

APPENDIX A

Glossary

aggregated subscription

A subscription that consumes multiple data sets from the same topic with a single batch workflow. An aggregated subscription can use an automatic mapping or a custom mapping to process data. When you use an automatic mapping, the subscription sorts the data according to the publication date and time of the publication instances.

application

An entity that represents a system in your organization that needs to share data with other systems. An application can be a publisher and a subscriber. An application can publish multiple data sets.

automatic Data Integration Hub mapping

A mapping that Data Integration Hub automatically generates to process a data set with a basic transformation logic that maintains the structure of the data. When you create a publication or a subscription with an automatic Data Integration Hub mapping, Data Integration Hub also creates the PowerCenter sources, targets, metadata folders, and connections to process the data set.

batch workflow

A workflow that runs once and stops after completion. The workflow reads from a file, a database, or another source and writes to a target. Use a batch workflow to process Data Integration Hub publications and subscriptions with a PowerCenter workflow or a Data Engineering Integration mapping.

big data

A set of data that is so large and complex that it cannot be processed through standard database management tools.

catalog

A list of all available topics in Data Integration Hub. You can subscribe to any topic in the catalog. You can also drill down to view the topic structure, and to view and edit the associated publications and subscriptions.

child event

An event within the hierarchy of another event that acts as a parent event. The child event is a subsidiary of the parent event. A child event in Data Integration Hub represents a subscription.

compound subscription

A custom subscription, either a custom batch subscription or a custom cloud subscription, that consumes data sets from multiple topics. All topics to which the subscription subscribes must be of the same topic type. For example, all the topics are managed on a relational database publication repository.

custom Data Integration Hub mapping

A mapping that processes a data set. The mapping includes the data sources and targets, metadata folders, and connections to process the data.

A custom mapping uses a PowerCenter workflow, a Data Engineering Integration mapping, or an Informatica Cloud task to process data. Workflows and tasks can perform complex transformations on the data.

A custom mapping that uses a PowerCenter workflow or a Data Engineering Integration mapping can include parameters.

Data Integration Hub repository

A relational database table set that contains the metadata required to process publications and subscriptions in Data Integration Hub. It also contains the events that Data Integration Hub generates while it processes publications and subscriptions.

Data Integration Hub server

A service that manages publication and subscription processing in Data Integration Hub. The Data Integration Hub server triggers batch workflows and sends and receives notifications from PowerCenter.

Data Integration task

A Data Integration task is a process that you configure to analyze, extract, transform, and load data. In Data Integration Hub, a Data Integration task is a task that reads from a file, a database, or another source and writes to a target. Use Data Integration tasks to process Data Integration Hub publications and subscriptions with Informatica Intelligent Cloud Services.

When you use a Data Integration task to process publications, you use the Data Integration Hub cloud connector as the publication target. When you use a Data Integration task to process subscriptions, you use the Data Integration Hub cloud connector as the subscription source.

document store

File directory where Data Integration Hub stores files that are published with a publication that is configured with the **When the file is ready to be published** scheduling option. When the application publishes from a file source with a publication with a **When the file is ready to be published** schedule, Data Integration Hub copies the file from the source directory to the document store to process the file contents. The document store directory must be accessible to the Data Integration Hub server, the Apache Tomcat server, and the PowerCenter Integration Service with the same file path.

event

An occurrence of a publication or subscription at each stage of processing. The Data Integration Hub server generates the event and updates the event status while it processes the publication or subscription.

Operation Console

Web interface to manage applications and topics, run and monitor publications and subscriptions, and administer user access in Data Integration Hub. Use a web browser to access the Operation Console.

parent event

An event at the top level of a hierarchy of events. A parent event in Data Integration Hub represents a publication. The parent event shows the status of both the parent event and of the child events that represent the subscribers to this publication.

publication

An entity that defines data flow from a data source to the Data Integration Hub publication repository and the data publishing schedule. The publication publishes the data to a topic that defines the structure of the data in the publication repository. When a publication runs, Data Integration Hub extracts the data set from the application, processes the data, and writes the data to the publication repository. You can then create one or more subscriptions to process and write the published data set to target applications.

publication repository

A relational database table set that stores published data sets that subscribers can consume. Data Integration Hub stores the data in the publication repository in the following ways:

- For each publication instance, the retention period for consumed data starts if all the subscribers have either successfully consumed or discarded the data. That is, after all the events that are associated with the publication instance are either in a Complete or in a Discarded event status. If all the subscribers consume or discard the data, Data Integration Hub stores the consumed data in the publication repository until the retention period for consumed data expires, and then deletes the consumed data from the publication repository.
- Data Integration Hub stores unconsumed data in the publication repository until the retention period for unconsumed data expires, and then deletes the unconsumed data from the publication repository.

real-time workflow

A PowerCenter workflow that is scheduled to run continuously and does not need to be started by Data Integration Hub. The workflow reads data from real-time sources and writes the data to Data Integration Hub. For example, data that is published through web-service providers or Java Message Service (JMS) queues. Use a real-time workflow to read data from real-time sources and write the data continuously to Data Integration Hub. When publishing through a real-time workflow, you group the published data into single publications at predefined time intervals.

subscription

An entity that defines the type, format, and schedule of data flow from the Data Integration Hub publication repository to a data target. When a subscription runs, Data Integration Hub extracts the data set from the publication repository, processes the data, and writes the data to the target application. You can subscribe to one or more topics. Each topic to which you subscribe can contain data from multiple publishers.

topic

An entity that represents a data domain that applications publish and consume through Data Integration Hub. A topic defines the data structure and additional data definitions such as the data retention period. Multiple applications can publish to the same topic. Multiple applications can consume data from the same topic.

unbound subscription

A subscription that is not restricted to specific publication instances. It subscribes to all the data that a publication publishes and consumes the data based on the subscription filter, regardless of when or in what batch the data was published.

INDEX

A

application
definition [13](#)

B

big data
description [12](#)

C

create
publication [39](#)
subscription [70](#)

G

glossary
of terms [101](#)

H

Hub Overview
diagram [15](#)
Hub Overview diagram
description [15](#)

O

operation console
description [15](#)
overview
description [10](#)

P

publication
creating [39](#)
definition [13](#)
overview [14](#)

S

subscription
creating [70](#)
definition [13](#)
overview [14](#)

T

topic
connection [19](#)
creating [19](#)
data retention [19](#)
definition [13](#), [19](#)
permissions [19](#)
publication repository [19](#)
structure [19](#), [22](#), [23](#), [25](#), [29–32](#)
topic structure
create table [25](#), [32](#)
from database [22](#), [29](#)
from file [23](#), [30](#)
from PowerCenter workflow [25](#), [31](#)

U

user roles
operator [14](#)