



Informatica® PowerExchange for Twitter
10.4.0

User Guide

Informatica PowerExchange for Twitter User Guide
10.4.0
December 2019

© Copyright Informatica LLC 2012, 2019

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, and PowerExchange 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.

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: 2019-12-12

Table of Contents

Preface	5
Informatica Resources.	5
Informatica Network.	5
Informatica Knowledge Base.	5
Informatica Documentation.	5
Informatica Product Availability Matrices.	6
Informatica Velocity.	6
Informatica Marketplace.	6
Informatica Global Customer Support.	6
 Chapter 1: Introduction to PowerExchange for Twitter.....	 7
PowerExchange for Twitter Overview.	7
Historical Twitter Data Extraction.	8
Real-Time Twitter Data Extraction.	8
 Chapter 2: PowerExchange for Twitter Configuration.....	 9
PowerExchange for Twitter Configuration Overview.	9
Twitter Application Creation.	9
Open Authentication Configuration.	10
Open Authentication Configuration Properties.	10
Configuring Open Authentication in the OAuth Utility.	11
Configuring Open Authentication in the Developer Tool.	11
Configuring HTTP Proxy Options at Design-Time.	11
Configuring HTTP Proxy Options at Design-Time.	11
Configuring HTTP Proxy Options while Using the OAuth Utility.	12
Configuring HTTP Proxy Options at Run-Time.	12
 Chapter 3: Twitter and Twitter Streaming Connections.....	 14
Twitter and Twitter Streaming Connections Overview.	14
Twitter Connection Properties.	14
Twitter Streaming Connection Properties.	15
Creating a Twitter Connection.	16
Creating a Twitter Streaming Connection.	17
 Chapter 4: Twitter and Twitter Streaming Data Objects.....	 18
Twitter and Twitter Streaming Data Objects Overview.	18
Twitter Data Object.	18
Twitter Data Object Views.	19
Twitter Data Object Overview Properties.	19
Twitter Data Object Operation Properties.	19

Twitter Resources.	21
Query Parameter for Twitter Resources.	22
Token File.	24
Pagination.	24
Twitter Streaming Data Object.	25
Twitter Streaming Data Object Views.	25
Twitter Streaming Data Object Overview Properties.	25
Twitter Streaming Data Object Operation Properties.	26
Twitter Streaming Resources.	27
Query Parameter for Twitter Streaming Resources.	28
Parameterization.	29
Creating a Data Object.	29
Creating a Data Object Operation.	30
Chapter 5: Twitter Mappings.	31
Twitter Mappings Overview.	31
Twitter Streaming Mappings.	31
Twitter API Rate Limits.	32
Twitter Mapping Example.	32
Appendix A: Datatype Reference.	34
Datatype Reference Overview.	34
Twitter and Transformation Datatypes.	34
Index.	35

Preface

Use the *Informatica® PowerExchange® for Twitter User Guide* to learn how to read from Twitter by using the Developer tool. Learn to create a Twitter connection and develop and run mappings in an Informatica domain.

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 PowerExchange for Twitter

This chapter includes the following topics:

- [PowerExchange for Twitter Overview, 7](#)
- [Historical Twitter Data Extraction, 8](#)
- [Real-Time Twitter Data Extraction, 8](#)

PowerExchange for Twitter Overview

Use PowerExchange for Twitter to read historical or real-time Twitter data through the Data Integration Service.

Use PowerExchange for Twitter to search and extract Twitter data such as tweets that are six to nine days old. You can create a Twitter connection to read historical Twitter data into a Twitter data object. You use open authentication (OAuth) to send secure authorized requests to Twitter.

You can extract near real-time Twitter data such as public statuses or a sample of statuses. You can create a Twitter streaming connection to continuously read real-time Twitter data into a Twitter streaming data object. A Twitter streaming connection requires OAuth authentication to send secure authorized requests to Twitter.

PowerExchange for Twitter uses Twitter API methods and resources to search and extract Twitter data. You can use Twitter search operators to define the search criteria. For example, you create a Twitter connection and define a query that searches for tweets with a positive attitude containing a keyword such as "Informatica." For real-time data, use a Twitter streaming connection to continuously read tweets that include a keyword. You can load the extracted data to a target and then use the data for data mining and analysis.

For information about the Twitter APIs, see the Twitter API documentation at the following website:

<https://dev.twitter.com/>

For information about the Twitter Developer usage policies, see the Twitter Developer Rules at the following website:

<https://dev.twitter.com/terms/api-terms>

Historical Twitter Data Extraction

Complete the following tasks to use PowerExchange for Twitter to extract historical Twitter data:

1. Create a Twitter application in the Twitter developer web site.
2. Configure OAuth.
3. Create a Twitter connection.
4. Create a Twitter data object.
5. Create a Twitter data object operation.
6. Create a mapping and use the Twitter data object operation as a source to extract historical Twitter data.

Real-Time Twitter Data Extraction

Complete the following tasks to use PowerExchange for Twitter to extract real-time Twitter data:

1. Configure OAuth.
2. Create a Twitter streaming connection.
3. Create a Twitter streaming data object.
4. Create a Twitter streaming data object operation.
5. Create a mapping and use the Twitter streaming data object operation as a source to extract real-time Twitter data.

CHAPTER 2

PowerExchange for Twitter Configuration

This chapter includes the following topics:

- [PowerExchange for Twitter Configuration Overview, 9](#)
- [Twitter Application Creation, 9](#)
- [Open Authentication Configuration, 10](#)
- [Configuring HTTP Proxy Options at Design-Time, 11](#)
- [Configuring HTTP Proxy Options at Run-Time, 12](#)

PowerExchange for Twitter Configuration Overview

PowerExchange for Twitter is installed with the Informatica services. You enable PowerExchange for Twitter with a license key.

Before you use PowerExchange for Twitter to access Twitter data, create a Twitter application and configure open authentication (OAuth) to send secure authorized requests to Twitter.

Optionally, you can configure HTTP proxy server authentication to access the Internet.

Twitter Application Creation

Before you configure open authentication, you must create a Twitter application.

Create a Twitter application in the following Twitter developer site:

<https://dev.twitter.com/apps>

In the application settings, you must specify the OAuth callback URL in the following format:

`http://<hostname.domain.com>:<port>/ows/jrs/callback`

Where

- `hostname` is the fully qualified system name of the master gateway node.
- `port` is the port number of the Administrator tool in an HTTP domain or an HTTPS port in an HTTPS domain. For example, 6008 if domain port is 6005 in an HTTP domain or 8443 in an HTTPS domain.

You use the consumer key and consumer secret listed in the application settings to configure OAuth.

Open Authentication Configuration

A Twitter connection and a Twitter Streaming connection requires open authentication (OAuth).

You configure OAuth in the OAuth Utility to send secure authorized requests to Twitter API. You can configure OAuth in the following tools:

- OAuth Utility. Open authentication configuration utility. You use the OAuth Utility to get the access token and access secret. You use the OAuth configuration details such as access token and access secret when you create a Twitter or Twitter streaming connection.
- Developer tool. You can create a Twitter or Twitter streaming connection in the Developer tool. You can launch the OAuth Utility from the **Connection** wizard to get the access token and access secret.

Open Authentication Configuration Properties

Use the OAuth Utility to configure open authentication.

The OAuth Utility properties that you configure to connect to Twitter can be used for a Twitter connection and a Twitter Streaming connection.

The following table describes the OAuth Utility properties that you configure to connect to Twitter:

Property	Description
Application	Social media web site. Select Twitter to configure OAuth for a Twitter application.
Consumer Key	The consumer key that you get when you create the Twitter application. Twitter uses the key to identify the application.
Consumer Secret	The consumer secret that you get when you create the Twitter application. Twitter uses the secret to establish ownership of the consumer key.
Callback Url	OAuth callback URL that the user is redirected after successful authentication. This property is read only. You must use the URL when you create the Twitter application.
Access Token	Access token that the OAuth Utility returns when you authorize the Twitter application. Twitter uses the token instead of the user credentials to access the protected resources.
Access Secret	Access secret that the OAuth Utility returns when you authorize the Twitter application. The secret establishes ownership of a token.

Configuring Open Authentication in the OAuth Utility

Use the OAuth Utility to get the access token and access secret.

Before you configure open authentication, you must create an application in Twitter.

1. In the address field of a browser, enter the following URL for the OAuth Utility page:

```
http://<hostname.domain.com>:<port>/ows/
```

The OAuth Utility page appears.

2. Select **Twitter** in the application.
3. Enter the consumer key and secret from the Twitter application settings.
4. Click **Authorize**.

The Twitter developer login page appears.

5. Log in to the Twitter developer site.
6. Navigate to the Twitter application page.
7. Click **Authorize App**.

The **Authentication Successful** page appears.

8. Close the **Authentication Successful** page.

The OAuth Utility lists the access token and access secret.

You use the access token and access secret to configure a Twitter connection.

Configuring Open Authentication in the Developer Tool

You can configure OAuth in the Developer tool when you create a Twitter connection.

In the **Connection** wizard, you can launch the OAuth Utility. The OAuth Utility opens the Twitter developer site in a browser. You authorize the Twitter application. The OAuth Utility populates the access token and access secret.

Configuring HTTP Proxy Options at Design-Time

If your organization uses a proxy server to access the internet, you can configure the HTTP proxy server authentication settings at design time. You can configure the HTTP proxy server authentication from the developerCore.ini file and from the web browser when you use the OAuth utility.

Configuring HTTP Proxy Options at Design-Time

If your organization uses a proxy server to access the internet, you can configure the HTTP proxy server authentication settings from the developerCore.ini file.

z

- Ensure that you enable the proxy server settings from your web browser.
- Access the developerCore.ini file from the following location:

```
<Informatica Installation Location>\clients\DeveloperClient
```

- Add the following properties to the developerCore.ini file:

Property	Description
-Dhttp.proxyHost=	Name of the HTTP proxy server.
-Dhttp.proxyPort=	Port number of the HTTP proxy server.
-Dhttp.proxyUser=	Authenticated user name for the HTTP proxy server. This is required if the proxy server requires authentication.
-Dhttp.proxyPassword=	Password for the authenticated user. This is required if the proxy server requires authentication. Note: The password is in plain text and not encrypted.
-Dhttp.nonProxyHosts=	List of host names or IP addresses for which you must not use the proxy server. Separate the list of IP addresses or host names with a pipe symbol (). For example, localhost:10.20.30.40 myHost Specify the IP address or name of the machine on which the Informatica gateway node runs so that the Developer tool connects to the domain.
-Dhttps.proxyHost=	Name of the HTTPS proxy server.
-Dhttps.proxyPort=	Port number of the HTTPS proxy server.

Configuring HTTP Proxy Options while Using the OAuth Utility

You can use a proxy server to access the internet when you use the OAuth utility. Enable the proxy server settings from your web browser.

Configuring HTTP Proxy Options at Run-Time

If your organization uses a proxy server to access the internet, you must configure the HTTP proxy server authentication settings for the Data Integration Service.

1. Open the Administrator tool.
2. Click the **Administration** tab, and then select the Data Integration Service.
3. Click the **Properties** tab.
4. Click **Edit** in the HTTP Proxy Server Properties section.
5. Configure the following properties:

Property	Description
HTTP Proxy Server Host	Name of the HTTP proxy server.
HTTP Proxy Server Port	Port number of the HTTP proxy server. Default is 8080.

Property	Description
HTTP Proxy Server User	Authenticated user name for the HTTP proxy server. This is required if the proxy server requires authentication.
HTTP Proxy Server Password	Password for the authenticated user. This is required if the proxy server requires authentication.
HTTP Proxy Server Domain	Domain for authentication.

CHAPTER 3

Twitter and Twitter Streaming Connections

This chapter includes the following topics:

- [Twitter and Twitter Streaming Connections Overview, 14](#)
- [Twitter Connection Properties, 14](#)
- [Twitter Streaming Connection Properties, 15](#)
- [Creating a Twitter Connection, 16](#)
- [Creating a Twitter Streaming Connection, 17](#)

Twitter and Twitter Streaming Connections Overview

Create a Twitter connection or a Twitter streaming connection to create data objects, preview data, and run mappings.

Use a Twitter connection to extract Twitter data that is between six and nine days old such as user profiles, friend's IDs, and tweets. A Twitter connection requires OAuth to access the Twitter data. You must create a Twitter application before you create a Twitter connection. You can configure the OAuth in the OAuth Utility or in the Developer tool.

Use a Twitter streaming connection for streaming, real-time access to public statuses. You can filter the statuses by a keyword or extract a random sample of public statuses. A Twitter streaming connection requires OAuth authentication. You can configure the OAuth in the OAuth Utility or in the Developer tool.

Twitter Connection Properties

Use a Twitter connection to extract data from the Twitter web site. The Twitter connection is a connection to social media. You can create and manage a Twitter connection in the Administrator tool or the Developer tool.

Note: The order of the connection properties might vary depending on the tool where you view them.

The following table describes Twitter connection properties:

Property	Description
Name	Name of the connection. The name is not case sensitive and must be unique within the domain. The name cannot exceed 128 characters, contain spaces, or contain the following special characters: ~ ` ! \$ % ^ & * () - + = { [] } \ : ; " ' < , > . ? /
ID	String that the Data Integration Service uses to identify the connection. The ID is not case sensitive. It must be 255 characters or less and must be unique in the domain. You cannot change this property after you create the connection. Default value is the connection name.
Description	The description of the connection. The description cannot exceed 765 characters.
Location	The domain where you want to create the connection.
Type	The connection type. Select Twitter.
Do you have OAuth details?	Indicates whether you want to configure OAuth. Select one of the following values: - Yes. Indicates that you have the access token and secret. - No. Launches the OAuth Utility.
Consumer Key	The consumer key that you get when you create the application in Twitter. Twitter uses the key to identify the application.
Consumer Secret	The consumer secret that you get when you create the Twitter application. Twitter uses the secret to establish ownership of the consumer key.
Access Token	Access token that the OAuth Utility returns. Twitter uses this token instead of the user credentials to access the protected resources.
Access Secret	Access secret that the OAuth Utility returns. The secret establishes ownership of a token.

Twitter Streaming Connection Properties

Use a Twitter Streaming connection to access near-real time data from the Twitter web site. The Twitter Streaming connection is a connection to the social media company's streaming API. You can create and manage a Twitter Streaming connection in the Administrator tool or the Developer tool.

Note: The order of the connection properties might vary depending on the tool where you view them.

The following table describes the general properties for a Twitter Streaming connection:

Property	Description
Name	Name of the connection. The name is not case sensitive and must be unique within the domain. The name cannot exceed 128 characters, contain spaces, or contain the following special characters: ~ ` ! \$ % ^ & * () - + = { [] } \ : ; " ' < , > . ? /
ID	String that the Data Integration Service uses to identify the connection. The ID is not case sensitive. It must be 255 characters or less and must be unique in the domain. You cannot change this property after you create the connection. Default value is the connection name.

Property	Description
Description	The description of the connection. The description cannot exceed 765 characters.
Location	The domain where you want to create the connection.
Type	The connection type. Select Twitter Streaming.

The following table describes the properties for hose type and OAuth authentication:

Property	Description
Hose Type	Streaming API methods. You can specify one of the following methods: <ul style="list-style-type: none"> - Filter. The Twitter <code>statuses/filter</code> method returns public statuses that match the search criteria. - Sample. The Twitter <code>statuses/sample</code> method returns a random sample of all public statuses.
Consumer Key	The consumer key that you get when you create the application in Twitter. Twitter uses the key to identify the application.
Consumer Secret	The consumer secret that you get when you create the Twitter application. Twitter uses the secret to establish ownership of the consumer key.
Do you have OAuth details?	Indicates whether you want to configure OAuth. Select one of the following values: <ul style="list-style-type: none"> - Yes. Indicates that you have the access token and secret. - No. Launches the OAuth Utility.
Access Token	Access token that the OAuth Utility returns. Twitter uses the token instead of the user credentials to access the protected resources.
Access Secret	Access secret that the OAuth Utility returns. The secret establishes ownership of a token.

Creating a Twitter Connection

Create a Twitter connection before you create a Twitter data object.

1. Click **Window > Preferences**.
2. Select **Informatica > Connections**.
3. Expand the domain in the **Available Connections**.
4. Select the connection type in **Social Media > Twitter** and click **Add**.
5. Enter a connection name and optional description.
6. Click **Next**.
7. Enter the consumer key and consumer secret.
8. Select whether you have the access token and secret.
 - Select **Yes** if you have the access token and secret.

- Select **No** if you want to generate the access token and access secret. The OAuth Utility opens the Twitter developer login page. Authorize the application. The **Authentication Successful** window appears. Close the window. The OAuth Utility populates the access token and access secret in the **Connection** wizard.
9. Click **Test Connection** to verify the connection to the Twitter web site.
 10. Click **Finish**.

Creating a Twitter Streaming Connection

Create a Twitter streaming connection before you create a Twitter streaming data object.

1. Click **Window > Preferences**.
2. Select **Informatica > Connections**.
3. Expand the domain in the **Available Connections**.
4. Select the connection type in **Social Media > Twitter Streaming** and click **Add**.
5. Enter a connection name and optional description.
6. Select the hose type.
7. Configure the OAuth connection properties.
 - a. Enter the consumer key and consumer secret.
 - b. Select whether you have the access token and secret.
 Select **Yes** if you have the access token and secret. Select **No** if you want to generate the access token and secret. The OAuth Utility opens the Twitter developer login page. Authorize the application. The Authentication Successful window appears. Close the window. The OAuth Utility populates the access token and access secret in the Connection wizard.
8. Click **Test Connection** to verify the connection to the Twitter web site.
9. Click **Finish**.

CHAPTER 4

Twitter and Twitter Streaming Data Objects

This chapter includes the following topics:

- [Twitter and Twitter Streaming Data Objects Overview, 18](#)
- [Twitter Data Object, 18](#)
- [Twitter Streaming Data Object, 25](#)
- [Parameterization, 29](#)
- [Creating a Data Object, 29](#)
- [Creating a Data Object Operation, 30](#)

Twitter and Twitter Streaming Data Objects Overview

A Twitter data object is a physical data object that represents data based on a Twitter resource. Create a Twitter data object to read data from resources such as Followers, Profile, and Place.

You can create a Twitter data object and a Twitter streaming data object. Create a Twitter data object when you want to read Twitter data that is six to nine days old. Create a Twitter streaming data object when you want to read real-time Twitter data.

After you create a Twitter data object, you create a data object operation. You can use a Twitter data object operation as a source in mappings and mapplets.

Twitter Data Object

Create a Twitter data object to extract Twitter data that is six to nine days old.

You must create a Twitter connection before you create a Twitter data object. You can specify the resources that you want to search such as Followers and Friends. You can specify a search criteria using Twitter search operators.

After you create a Twitter data object, create a Twitter data object operation to use as a source in mappings and mapplets.

Twitter Data Object Views

After you create a Twitter data object, you can modify the data object properties in the following data object views:

- **Overview** view. Use the **Overview** view to edit the Twitter data object name, description, and resources.
- **Data Object Operation** view. Use the **Data Object Operation** view to modify the properties that the Data Integration Service uses when it reads data from Twitter.

When you create mappings that use Twitter sources, you can view the data object properties in the **Properties** view.

Twitter Data Object Overview Properties

The **Overview** properties include general properties that apply to the Twitter data object. They also include object properties that apply to the resources in the Twitter data object.

General Properties

The following table describes the general properties that you configure for Twitter data objects:

Property	Description
Name	Name of the Twitter data object.
Description	Description of the Twitter data object.
Connection	Name of the Twitter connection.

Object Properties

The following table describes the object properties that you can view for resources:

Property	Description
Name	Name of the resource.
Type	Type of the resource.
Description	Description of the resource.

Twitter Data Object Operation Properties

The data object operation properties include general, ports, sources, and advanced properties that the Data Integration Service uses to read data from Twitter.

General Properties

The general properties lists the name and description of the data object operation.

The following table describes the general properties that you can view for a Twitter data object operation:

Property	Description
Name	Name of the Twitter data object operation.
Description	Description of the Twitter data object operation.

Column Properties

The column properties lists the datatypes, precision, and scale of the source transformation.

The following table describes the column properties that you can view for a Twitter data object operation:

Property	Description
Name	Name of the Twitter data object operation.
Type	Transformation datatype of the metadata.
Precision	Maximum number of significant digits for numeric datatypes, or maximum number of characters for string datatypes. For numeric datatypes, precision includes scale.
Scale	Maximum number of digits after the decimal point for numeric values.
Description	Description of the Twitter data object operation.

Advanced Properties

The advanced properties lists the resource physical name of the source transformation.

Ports Properties

The ports properties lists the datatypes, precision, and scale of the Input transformation.

The following table describes the ports properties that you can view for a Twitter data object operation:

Property	Description
Name	Name of the metadata such as FollowersCount, FriendsCount, and Username of the resource.
Type	Datatype of the metadata.
Precision	Maximum number of characters for string datatypes.
Description	Description of the metadata.

Sources Properties

The sources properties lists the resources of the Input transformation.

The following table describes the sources properties that you configure for a Twitter data object operation:

Property	Description
Sources	List of the Twitter resources in the data object operation. You can add or delete resources.

Advanced Properties

The advanced properties includes the run-time properties of the Input transformation. You can specify the search criteria to extract Twitter data in the query parameter. When you run the mapping, the Data Integration Service uses the search criteria to extract Twitter data.

The following table describes the advanced properties that you configure for a Twitter data object operation:

Property	Description
Query Parameter	<p>Twitter search criteria. You can specify the Twitter search criteria based on the resource. Each resource requires one of the following mandatory parameters:</p> <ul style="list-style-type: none"> - screen_name - user_id - QUERY <p>For example, you specify the screen name in the following query parameter to extract the Twitter profile of the Twitter screen name Rinnovator:</p> <pre>screen_name=Rinnovator</pre>
Token File	<p>File name and format of a file that contains a list of access tokens and access secrets. The CSV file must be on the machine where Informatica Services is installed. The Data Integration Service uses the access token and access secret at run time to establish a connection. If one set of access token and access secret fails, the Data Integration Service uses the next set. Review the session log for details.</p> <p>Specify the fully qualified path and format of the file in one of the following formats:</p> <ul style="list-style-type: none"> - <file_path>;{accesstoken,accesssecret}or <file_path> <p>Use one of the formats if the file contains only the mandatory comma-separated columns of access token and access secret.</p> <ul style="list-style-type: none"> - <file_path>;{col1,accesstoken,accesssecret,col4} <p>Use this format to specify optional columns such as col1 and col4. For example, \\hostname\OAuth\tw_tokens.csv;{accesstoken,accesssecret,ID,name} specifies the location of a file named tw_tokens.csv that contains the following mandatory comma-separated columns: accesstoken and accesssecret. The Data Integration Service ignores the optional columns, ID and name, at run time.</p>
Ends After	<p>Twitter pagination property. Controls the number of rows requested. Specify a time or a positive numeral:</p> <ul style="list-style-type: none"> - Amount of time that the Data Integration Service runs the mapping. Enter the duration in the following format: hh:mm <p>Default is blank, which indicates that the Data Integration Service runs the mapping until you stop it or until there is data.</p> <ul style="list-style-type: none"> - Number of rows that are requested. The Data Integration Service will request the minimum number of rows which is 100. The maximum number of rows requested is based on the rows returned by the Twitter application. Enter the number of rows as a positive numeral. <p>Configure for the Search resource.</p>

Twitter Resources

PowerExchange for Twitter uses the Twitter API to search and retrieve Twitter data. The Twitter API resources allow access to Twitter data. These resources provide information such as followers, tweets, and user information. Each resource includes metadata such as FollowersCount, FriendsCount, and Username.

For more information about the Twitter API, see the Twitter API documentation at <https://dev.twitter.com>.

When you create a Twitter data object or create a Twitter data object operation, you specify the resources based on the information you want to extract.

The following table describes the Twitter resources:

Resource	Description
Followers	Extracts the Twitter IDs for every user following the specified user. The Followers resource includes metadata that gives IDs of all the followers.
Friends	Extracts the Twitter IDs for every user the specified user is following. The Friends resource includes metadata that gives IDs of all whom you are following.
Search	Extracts the tweets that match a specified query. The Search resource includes metadata such as Entry, FromUser, and FromID. For more information about the Twitter Search API, see the Twitter API documentation at https://dev.twitter.com/docs/api/1.1/get/search/tweets .
Profile Lookup	Extracts the user profiles that match the specified userid or screen name. The Profile resource includes metadata such as Username, location, and ID.

Query Parameter for Twitter Resources

Use the query parameter to specify the search criteria.

When you create a Twitter data object, you specify the query parameter that the Data Integration Service uses to search for Twitter data. The query parameter must include the mandatory parameter. Each resource has one mandatory parameter.

You can use Twitter search operators with the mandatory parameter. The Twitter search operators such as OR, @, and ? that you can use are defined in the Search Operators section of the documentation for the Twitter Search API. For example, you specify the following query parameter to extract tweets that contain the keyword apple with a positive attitude: `QUERY=Apple:)`

The following table describes the mandatory parameter for each resource:

Resource	Mandatory Parameter	Description
Followers	screen_name user_id	Screen name or user ID. Specify the screen name or user ID of the Twitter account for which you want a list of followers. For example, you specify the following query parameter to extract a list of Twitter IDs of the followers of the Twitter screen name Rinnovator: <code>screen_name=Rinnovator</code>
Friends	screen_name user_id	Screen name or user ID. Specify the screen name or user ID of the Twitter account for which you want a list of friends. For example, you specify the following query parameter to extract the Twitter IDs of the friends of the Twitter screen name Rinnovator: <code>screen_name=Rinnovator</code>

Resource	Mandatory Parameter	Description
Profile Lookup	screen_name user_id	Screen names or user IDs. Specify the screen name or user IDs of the Twitter account for which you want the profiles. Specify the profile query parameter in the following formats: <ul style="list-style-type: none"> - You can specify up to 100 comma-separated screen names or user IDs. - You can specify a valid user ID or a screen name to extract the profile of a user. - You can specify a list of user IDs or screen names in a .txt or .csv format to extract the profiles of users.
Search	QUERY	Search keyword. Specify the comma-separated keywords to search for tweets. For example, you specify the following query parameter to extract the tweets that contain the keywords Apple, Lenevo, or HTC: QUERY=Apple,Lenovo,HTC

Profile Lookup Resource

You can use the profile lookup resource to extract the user profiles that match the specified user ID or screen name.

Specify the profile query parameter in the following formats:

- Comma-separated list. You can specify up to 100 comma-separated screen names or user IDs. Use the following format to specify the profile query parameter as a comma-separated list:

```
user_id=<comma separated list of user_id>
screen_name=<comma separated list of screen_name>
```

For example, you specify the screen name in the following query parameter to extract the Twitter profile of the Twitter screen name Rinnovator and nytimes:

```
screen_name=Rinnovator,nytimes
```

You can specify the user ID in the following query parameter to extract the Twitter profile of the Twitter user ID 123 and 245:

```
user_id=123,245
```

Note: The profiles might not be listed in the order they are specified in the query parameter.

- Valid user ID or screen name. You can specify a valid user ID or a screen name. Use the following format to specify the profile query parameter with a valid user ID or screen name:

```
user_id=<valid twitter userid>
screen_name=<valid twitter screen name>
```

For example, you specify the screen name in the following query parameter to extract the Twitter profile of the Twitter screen name Charley Rarely:

```
screen_name=CharleyRarely
```

You can specify the user ID in the following query parameter to extract the Twitter profile of the Twitter ID 544282590:

```
user_id=544282590
```

- List in a .txt or .csv file. You can specify a list of user IDs or screen names in a text or csv file to extract the profiles of users.

Use the following format to specify the profile query parameter as a list in a .txt or .csv file:

```
user_id=file:///<file_path/userid_file>
```

```
screen_name=file:///<file_path/screennames_file>
```

For example, `user_id=file:///E:/SocialMedia/twitter_userid.txt`

```
screen_name=file:///E:/SocialMedia/twitter_screenname.txt
```

Token File

You can specify the file name and format of a file that contains a list of access tokens and access secrets.

The following table provides the list of Twitter resources and whether the token file is applicable:

Resource	Token File
Followers	Yes
Friends	Yes
Profile Lookup	No
Search	No

Pagination

PowerExchange for Twitter uses the Twitter API to control the number of rows that you request when a mapping runs.

When you create a Twitter data object operation, you can specify the amount of time that the mapping runs or the number of rows that you request.

The following table specifies the pagination impact for each Twitter resource, the number of default rows that you can request, and whether the Ends After property is applicable for each Twitter resource:

Resource	Pagination Impact	Default Rows Requested	Ends After Property Applicable?
Followers	No	All rows	No
Friends	No	All rows	No
Profile Lookup	No	All rows	No
Search	Yes	100	Yes

Twitter Streaming Data Object

Create a Twitter streaming data object to extract real-time Twitter data.

You must create a Twitter streaming connection before creating a Twitter streaming data object. You can specify filter or sample hose types. You can specify a search criteria for the filter hose type.

After you create a Twitter streaming data object, create a Twitter streaming data object operation to use as a source in mappings and mapplets.

Twitter Streaming Data Object Views

After you create a Twitter streaming data object, you can modify the data object properties in the following data object views:

- **Overview** view. Use the **Overview** view to edit the Twitter streaming data object name, description, and resources.
- **Data Object Operation** view. Use the **Data Object Operation** view to configure the search criteria that the Data Integration Service uses when it reads data from Twitter.

When you create mappings that use Twitter sources, you can view the data object properties in the **Properties** view.

Twitter Streaming Data Object Overview Properties

The **Overview** properties include general properties that apply to the Twitter streaming data object. They also include object properties that apply to the resources in the Twitter streaming data object.

General Properties

The following table describes the general properties that you configure for Twitter streaming data objects:

Property	Description
Name	Name of the Twitter streaming data object.
Description	Description of the Twitter streaming data object.
Connection	Name of the Twitter streaming connection.

Object Properties

The following table describes the object properties that you can view for resources:

Property	Description
Name	Name of the resource.
Type	Type of the resource.
Description	Description of the resource.

Twitter Streaming Data Object Operation Properties

The data object operation properties include general, ports, sources, and advanced properties that the Data Integration Service uses to read data from Twitter.

General Properties

The general properties lists the name and description of the data object operation.

The following table describes the general properties that you can view for a Twitter streaming data object operation:

Property	Description
Name	Name of the Twitter streaming data object operation.
Description	Description of the Twitter data object operation.

Column Properties

The column properties lists the datatypes, precision, and scale of the source transformation.

The following table describes the column properties that you can view for a Twitter streaming data object operation:

Property	Description
Name	Name of the Twitter streaming data object operation.
Type	Transformation datatype of the metadata.
Precision	Maximum number of significant digits for numeric datatypes, or maximum number of characters for string datatypes. For numeric datatypes, precision includes scale.
Scale	Maximum number of digits after the decimal point for numeric values.
Description	Description of the Twitter streaming data object operation.

Advanced Properties

The advanced properties lists the resource physical name of the source transformation.

Ports Properties

The ports properties lists the datatypes, precision, and scale of the Input transformation.

The following table describes the ports properties that you can view for a Twitter streaming data object operation:

Property	Description
Name	Name of the metadata such as ID of the resource.
Type	Native datatype of the metadata.

Property	Description
Precision	Maximum number of characters for string datatypes.
Description	Description of the metadata.

Sources Properties

The sources properties lists the resources of the Input transformation.

The following table describes the sources properties that you configure for a Twitter streaming data object operation:

Property	Description
Sources	List of the Twitter resources in the data object operation. You can add or delete resources.

Advanced Properties

The advanced properties includes the run-time properties of the Input transformation. You can specify the search criteria and stop time to extract Twitter real-time data in the query parameter. When you run the mapping, the Data Integration Service uses the run-time properties to extract Twitter real-time data.

The following table describes the advanced properties that you configure for a Twitter streaming data object operation:

Property	Description
Query Parameter	Twitter search criteria. You can use Twitter search operators to define the search criteria for the <code>statuses\filter</code> method.
Ends After	Duration for which the Data Integration Service runs the mapping. Enter the duration in the following format: <code>hh:mm</code> For example, specify the following duration to run the mapping for ten days: <code>240:00</code> If you leave this option blank, the Data Integration Service runs the mapping continuously until you stop the mapping.

Twitter Streaming Resources

PowerExchange for Twitter uses the Twitter Streaming API to search and retrieve Twitter real-time data.

For more information about the Twitter Streaming API, see the Twitter API documentation at <https://dev.twitter.com>.

When you create a Twitter streaming data object or create a Twitter streaming data object operation, you specify the resources based on the information you want to extract.

The following table describes the Twitter resources you can specify:

Resource	Description
User	Extracts public statuses that reference the specified set of users. These references include tweets, retweets, replies.
Place	Extracts tweets created using the Geotagging API.
Status	Extracts tweets based on search criteria specified in a comma-separated list.

Query Parameter for Twitter Streaming Resources

Use the query parameter to specify the search criteria.

When you create a Twitter streaming data object, you specify the query parameter that the Data Integration Service uses to search for real-time data.

The sample hose type does not require a query parameter. A sample hose type returns a random sample of Twitter statuses.

The filter hose type requires a mandatory parameter. You can use Twitter search operators such as count, track, and location with the mandatory parameter. The operators are defined in the Streaming API Twitter documentation.

The following table describes the mandatory parameter for Filter hose type:

Hose Type	Mandatory Parameter	Description
Filter	TOPICS	Search keyword. Specify up to 144 comma-separated keywords for the tweets you want to extract. For example, you specify the following query parameter to extract the tweets that contain the keywords: TOPICS=Lenovo,Harley Davidson,BMW,Opel

Parameterization

You can parameterize the connection and read operation properties of Twitter and Twitter Streaming data objects to override the properties at run time.

Twitter Data Objects

The following table lists the read operation properties that you can parameterize and the type of parameterization supported:

Property	Type of Parameterization Supported
Query Parameter	Partial
Token File	Partial
Ends After	Full

Twitter Streaming Data Objects

The following table lists the read operation properties that you can parameterize and the type of parameterization supported:

Property	Type of Parameterization Supported
Query Parameter	Partial
Ends After	Full

Creating a Data Object

Create a Twitter or Twitter streaming data object to specify the Twitter resources.

1. Select a project or folder in the Object Explorer view.
2. Click **File > New > Data Object**.
3. Select **Twitter Data Object** or **Twitter Streaming Data Object** and click **Next**.
The **New Twitter Data Object** dialog box appears.
4. Click **Browse** next to the **Location** option and select the target project or folder.
5. Click **Browse** next to the **Connection** option and select a connection from which you want to import the Twitter resource.
6. To add a resource to the data object, click **Add** next to the **Resource** option.
The **Add sources to the data object** dialog box appears.
7. Navigate or search for the resources to add to the data object and click **OK**.
8. Optionally, enter a name for the data object.
9. Click **Finish**.

The data object appears under Data Object in the project or folder in the Object Explorer view. You can also add resources to a data object after you create it.

Creating a Data Object Operation

Create a data object operation from a data object. You can create multiple data object operations from a data object. Each data object operation must have only one resource.

Before you create a data object operation, you must create the data object with the resource.

1. Select the data object in the Object Explorer view.
2. Right-click and select **New > Data Object Operation**.

The **Data Object Operation** dialog box appears.

3. Enter a name for the data object operation.
4. Select the type of data object operation.
5. Click **Add**.

The **Select a resource** dialog box appears.

6. Select the resource for which you want to create the read operation and click **OK**. You can specify only one resource for a data object operation.
7. Click **Finish**.

The Developer tool creates the data object operation for the selected data object.

CHAPTER 5

Twitter Mappings

This chapter includes the following topics:

- [Twitter Mappings Overview, 31](#)
- [Twitter Streaming Mappings, 31](#)
- [Twitter API Rate Limits, 32](#)
- [Twitter Mapping Example, 32](#)

Twitter Mappings Overview

After you have created the Twitter data object operation or Twitter streaming data object operation, you can develop a mapping. You can define the following objects in the mapping:

- Twitter or Twitter streaming data object operation as the input to read data from Twitter
- Transformations
- Any target

Validate and run the mapping to extract the Twitter data and load it to a target.

Twitter Streaming Mappings

When you run a mapping with Twitter streaming data objects, the Data Integration Service extracts Twitter real-time data.

Before you run a Twitter streaming mapping, review the following guidelines:

- You can specify the duration for which the Data Integration Service runs the mapping. The mapping runs and the data extraction continues for the specified duration. Specify a duration in the Ends After property for the Twitter streaming data object operation in the following format:
`hh:mm`
- To run multiple mappings, create multiple Twitter streaming connections, each connection with unique Twitter user account credentials.

Twitter API Rate Limits

The Twitter API specifies rate limits for the number of requests per hour.

Twitter imposes different types of rate limits based on the usage such as the number of API calls for each API key, number of calls for each account type, and number of calls for each method.

Review the Twitter policies when you design the frequency of the API calls. If the number of requests exceeds the permitted limits, the Twitter API returns an error response with more information about the error. Twitter may also prohibit users from future API access if guidelines are continuously and persistently violated. Users will have to directly contact Twitter to reset their API access. To avoid being prohibited, review the Twitter guidelines at the following web site:

<https://dev.twitter.com/>

Twitter Mapping Example

You work for an organization, HypoMarket Corporation. You need to review all the tweets that mention your product HypoBasket since the time you released the product in February 2012. You need to review the users IDs and the location of the tweets and verify that the product HypoBasket is a trending topic on Twitter.

Create a mapping that identifies tweets that contain the word HypoBasket, the User ID, and location of the tweets, and writes those records to a table.

You can use the following objects in a Twitter mapping:

Twitter input

The mapping source is a Twitter data object that contains the resource Search.

Create a physical data object and add the data object to the mapping. Add the Search resource to the physical data object. Change the query parameter with the following query:

```
QUERY=HypoBasket&since:2012-02-01
```

Sorter transformation

Optionally, sort the data based on the time stamp.

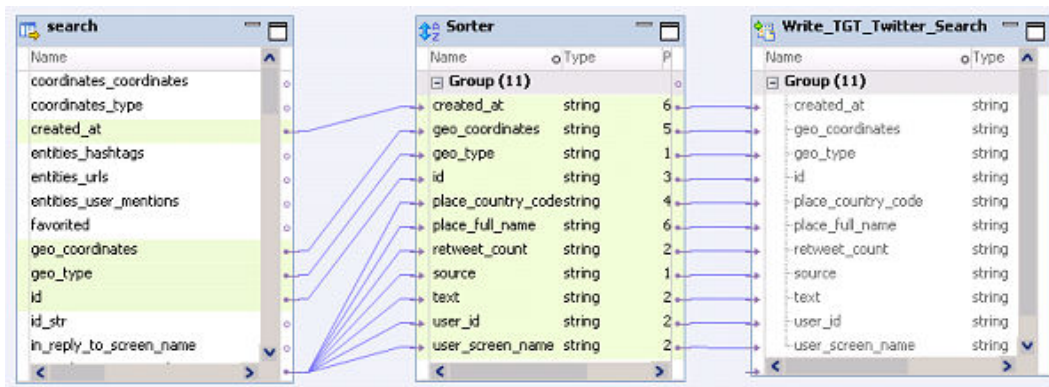
Add a Sorter transformation to the mapping. Specify the time stamp as the sort key with direction as ascending.

Mapping output

Add a flat file data object to the mapping as a target.

After you run the mapping, Data Integration Service writes the extracted tweets to the target table. You can use text analytics and sentiment analysis tools to analyze the tweets.

The following image shows the mapping example:



APPENDIX A

Datatype Reference

This appendix includes the following topics:

- [Datatype Reference Overview, 34](#)
- [Twitter and Transformation Datatypes, 34](#)

Datatype Reference Overview

Informatica Developer uses the following datatypes in Twitter mappings:

- Twitter native datatypes. Twitter datatypes appear in the physical data object column properties.
- Transformation datatypes. Set of datatypes that appear in the transformations. They are internal datatypes based on ANSI SQL-92 generic datatypes, which the Data Integration Service uses to move data across platforms. Transformation datatypes appear in all transformations in a mapping.

When the Data Integration Service reads source data, it converts the native datatypes to the comparable transformation datatypes before transforming the data. When the Data Integration Service writes to a target, it converts the transformation datatypes to the comparable native datatypes.

Twitter and Transformation Datatypes

The following table lists the Twitter native datatypes that Data Integration Service supports and the corresponding transformation datatypes:

Twitter Datatype	Transformation Datatype	Range and Description
String	String	1 to 104,857,600 characters

INDEX

C

- callback URL [9](#)
- configuring
 - HTTP proxy options [12](#)
 - OAuth [11](#)
- configuring HTTP proxy options
 - Developer tool [11](#)
- configuring OAuth
 - Developer tool [11](#)
- connections overview
 - description [14](#)
- creating
 - data object operation [30](#)
 - Twitter application [9](#)
 - Twitter connection [16](#)
 - Twitter data object [29](#)
 - Twitter streaming connection [17](#)

D

- data object operation [30](#)
- data objects
 - Twitter [18](#)
 - Twitter streaming [25](#)
- datatype reference overview
 - description [34](#)
- datatypes
 - Transformation [34](#)
 - Twitter [34](#)

E

- extraction
 - historical Twitter data [8](#)
 - real-time Twitter data [8](#)

M

- mapping example
 - Twitter [32](#)

O

- OAuth
 - configuration properties [10](#)
 - configuring [11](#)
- open authentication [10](#)

P

- pagination
 - twitter [24](#)
- PowerExchange for Twitter overview
 - description [7](#)

Q

- query parameter
 - Twitter resources [22](#)
 - Twitter streaming data object [28](#)

R

- resources
 - Twitter [21](#)
 - Twitter streaming [27](#)

T

- Twitter
 - resources [21](#)
 - token file [24](#)
- Twitter API rate limits [32](#)
- Twitter connection
 - creating [16](#)
- Twitter connections
 - properties [14](#)
- Twitter data object
 - creating [29](#)
 - general properties [19](#)
 - objects properties [19](#)
 - overview properties [19](#)
 - properties [19](#)
 - views [19](#)
- Twitter data object operation
 - advanced properties [19](#)
 - column properties [19](#)
 - creating [30](#)
 - general properties [19](#)
 - ports properties [19](#)
 - properties [19](#)
 - sources properties [19](#)
- Twitter data objects overview
 - description [18](#)
- Twitter streaming
 - ends after [31](#)
 - mappings [31](#)
 - resources [27](#)
- Twitter streaming connection
 - creating [17](#)

Twitter Streaming connections
 properties [15](#)
Twitter streaming data object
 advanced properties [26](#)
 column properties [26](#)
 creating [29](#)
 general properties [25](#), [26](#)
 objects properties [25](#)

Twitter streaming data object (*continued*)
 overview properties [25](#)
 ports properties [26](#)
 properties [25](#), [26](#)
 sources properties [26](#)
 views [25](#)
Twitter streaming data object operation
 creating [30](#)