Azure Data Lake Prerequisites

Before installing the Azure Data Lake connector, make sure that:

- You have created the necessary Service-to-Service Application account with pertinent access to your Data Lake.
- The Azure Data Lake is up and running.
- You have Admin rights to allow Read and Execute permissions on folders to Crawl.

Version of Azure Data Lake Required

The Aspire Azure Data Lake connector was created and tested using **Aspire Version 3.3**.

User Account Requirements

In order to access the Azure Data Lake, an Application Account with sufficient privileges must be supplied. The following fields must be configured in order to set up a new Data Lake connection:

- Authorization Token End Point: *p.e.* *https://login.microsoftonline.com/[yourkey]/oauth2/token*
- Application ID
- Application Secret (Application Key)
- Fully Qualified Domain Name (FQDN): *p.e [yourdomain].azuredatalakestore.net.* No HTTP prefix is required

Get an Application Account

1. See Microsoft's Use portal to create an Azure Active Directory application and service principal that can access resources for the steps on how to properly create an Application ID and its key.

> ⚠ Make sure to write down your Application Key at the time of creation. It will not be shown again after you exit the portal. **Important:** Make sure to grant the necessary Reader access to your application.

2. This connector uses a OAuth 2.0 authorization via Token End Point. Azure will supply this authorization. See Microsoft's Step 4: Get the OAuth 2.0 token endpoint (only for Java-based applications).

After these steps are completed, you will have created a valid Application.

3. Make sure to grant Read and Execute access (at least) to files and folders to crawl. See Microsoft's Step 3: Assign the Azure AD application to the Azure Data Lake Store account file or folder.

4. Follow the recommended Advance Features of the Data Lake File Explorer to recursively apply the same parent folder permissions to sub-folders using the "Apply folder permissions to sub-folders" option.

The Application does not have access to any specific folder. Aspire will log this warning during the crawl process

Windows or Linux

The Azure Data Lake connector runs on either a Windows or Linux instance of Aspire.