Product Data Lake

User Guide, Data Exchange, version 1.4, February 2018

Introduction

Uploading data to the Product Data Lake upstream is done by using hot folders on FTP sites, by importing files inside the Product Data Lake or by manually typing and twisting data in the Product Data Lake user interface. Data can also be provided via put APIs. Please refer to Product Data Lake API Guide.

Downloading data downstream is done by exporting files inside the Product Data Lake or by using a pull request. A pull request is an order to provide selected data on a FTP site at given time. The pull request can be repeated daily, weekly or monthly.

This guide describes:

- What data structures are used
- How to set up hot folders
- How to upload data from your PC or network drives
- How to make pull requests
- How to interactively download (export) data
- How to view logs
- Where to place other data structures and what to do with them

Files can be formatted as XML or CSV. Note that CSV files are with tab as separator (as TXT files).

Data structures

The following structures can be exchanged:

- Attribute list
- Product list
- Attribute link
- Product link
- Attribute value
- Related product
- Digital asset
- Digital asset link
- Synonym
- Batch log (only download)
- Error log (only download)

Attribute list as CSV file / XML tag <AttributeList>, your available attributes for a product.

Per <Attribute>

- Attribute ID, <AttributeID>, your identifier for the attribute
- Attribute Type, <AttributeType>, optional, a tag to support linking with your partners *)
- Attribute Group ID, <AttributeGroupID>, optional, your identifier for a group of attributes used for a group of products
- Attribute Group Description, <AttributeGroupDescription>, optional, the description for a group of attributes
- Specific Trading Partner ID, <SpecificTradingPartnerID>, optional, your partners ID in the Product Data Lake if you only want to share this attribute with a specific trading partner **)
- Short Attribute Description, <ShortAttributeDescription>, your short description used for naming the attribute

- Long Attribute Description, <LongAttributeDescription>, your long description used for explaining the use of the attribute
- Language, <Language>, the language used in the description. Format is ll-CC where ll is the two-character language code in lower case and CC is the two-character country code in upper case.

*) Attribute Type can be:

- Product identification, for example =GTIN, =MODEL, =DESCR
- Product classification, for example #UNSPSC, #ETIMCLASS
- Product feature, for example EF123456@ETIM, ADR123@ECLASS
- Product trading data, for example !STOCK, !PRICE-LIST-USD, !PRICE-TRADE-EUR

Please refer to the User Guide for Manage Attributes for further explanation.

**) The Partner ID can be seen on the Product Data Lake in Manage Partnership, Current Partnership, either in the Upstream or Downstream pane depending on your supply chain placement.

Product list as CSV file / XML tag <ProductList>, your portfolio of products on The Product Data Lake.

Per <Product>

- Product ID, <ProductID>, your internal identifier for the product
- Product Description, <ProductDescription>, your description of the product
- Model Number, <ModelNumber>, optional, manufacturer model number for the product
- GTIN, <GTIN>, optional, Global Trade Item Number for the product *)
- Attribute Group ID, <AttributeGroupID>, optional, your identifier for the group of attributes used for this product
- Specific Trading Partner ID, <SpecificTradingPartnerID>, optional, your partners ID in the Product Data Lake if you only want to share this product with a specific trading partner **)

*) Can be Universal Product Code (UPC) or International (formerly European) Article Number (EAN). Please refer to the User Guide for Manage Product for further explanation.

**) The Partner ID can be seen on the Product Data Lake in Manage Partnership, Current Partnership, either in the Upstream or Downstream pane depending on your supply chain placement.

Attribute link as CSV file / XML tag <AttributeLink>, the mapping between your attributes and your partners attributes, if you know them beforehand.

Per <Link>

- Attribute ID, <AttributeID>, your identifier for the attribute
- Partner Attribute ID, <PartnerAttributeID>, your partners attribute identifier
- Partner ID, <PartnerID>, your partners ID in the Product Data Lake **)

Product link as CSV file / XML tag <ProductLink>, the mapping between your products and your partners products, if you know them beforehand.

Per <Link>

- Product ID, <ProductID>, your identifier for the product
- Partner Product ID, <PartnerProductID>, your partners product identifier
- Partner ID, <PartnerID>, your partners ID in the Product Data Lake **)

Attribute value as CSV file, XML tag <AttributeValue>, the actual values of an attribute for a product (in a language).

Per <Value>

- Product ID, <ProductID>, your identifier for the product
- Attribute ID, <AttributeID>, your identifier for the attribute
- Language, <Language>, optional, language code if your profile is multi-language and values of the attribute differs per language. . Format is ll-CC where ll is the two-character language code in lower case and CC is the two-character country code in upper case.
- Value, <Value>, the value of this attribute for this product (in this language)

Digital Assets as CSV file, XML tag <DigitalAsset>, the reference to a rich media file.

Per <DigitalAssetElement>

- Digital Asset ID, <DigitalAssetID>, An ID per digital asset
- Digital Asset Description, <Digital Asset Description>, a description of the content
- Digital Asset Type ID, <DigitalAssetTypeID>, an identifier of the digital asset type
- Digital Asset Type, <DigitalAssetType>, a text describing the digital asset type
- Digital Asset Reference, <DigitalAssetReference>, a link to where the digital asset from the provider is stored
- Product ID, <ProductID>, your identifier for the product

Note: You can exchange an Attribute Value CSV file with Products as rows, Product ID as column 1, Attribute ID as further columns and Value as further cells.

Sample files are available on the Product Data Lake in Manage Data Exchange, Load Data.

Setting up hot folders

On your FTP Domain set up a folder called:

<FTP server address>\Hotfolders

Set up the following subfolders:

<FTP server address>\Hotfolders\Work

<FTP server address>\Hotfolders\Archive

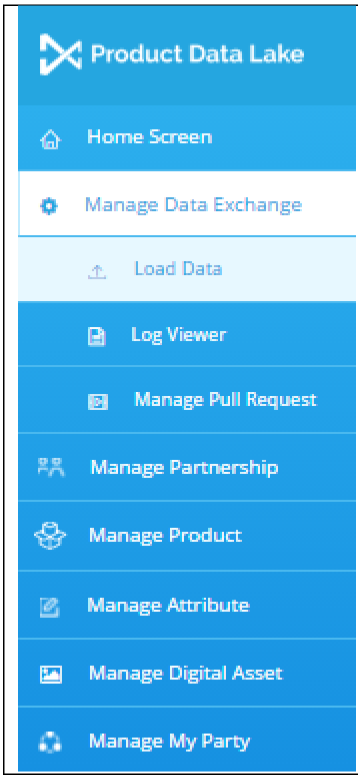
<FTP server address>\Hotfolders\Error

Files to be uploaded on a given UTC date shall be placed in a folder named:

<FTP server address>\Hotfolders\Work\YYYY-MM-DD

The Product Data Lake will place the file in the Achive of the date if upload is successful and in Errors of the date if upload failed.

Load data

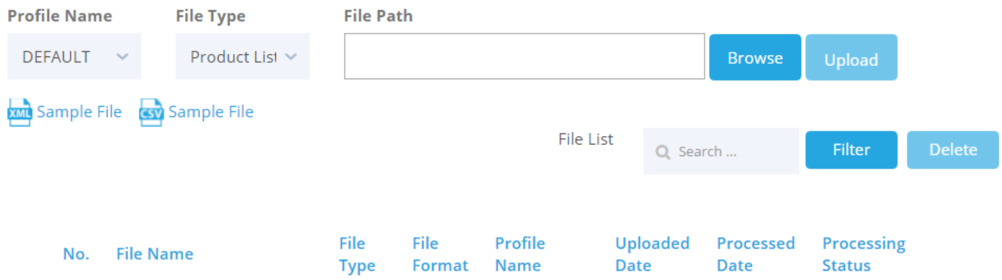


When selecting Manage Data Exchange, Load data you will get an interactive function where you can:

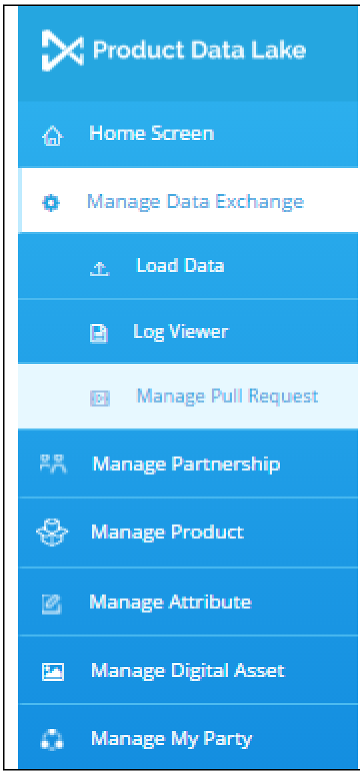
- Select the profile to be used for uploading data. The profile holds technical preferences for data exchange. Please refer to the Manage My Party User Guide for more explanation.
- Select the file type among up-loadable data structures
- Select the file path on your PC or network drives

Sample files are available on this screen.

The screen also holds a history over loaded files.

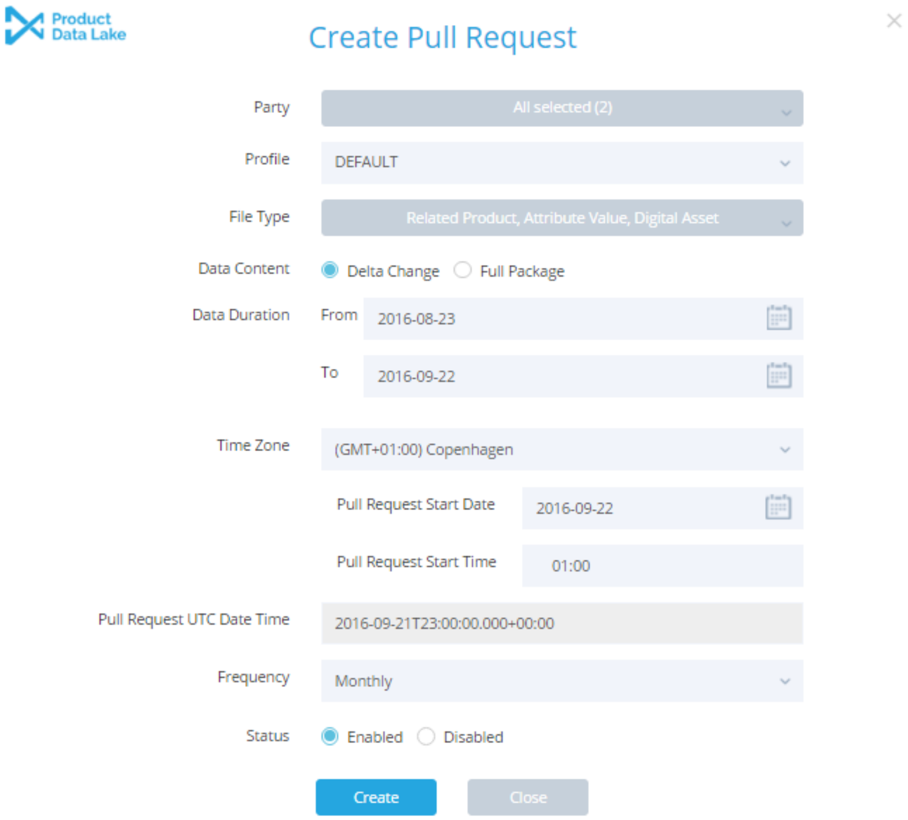


Manage Pull Request



When selecting Manage Data Exchange, Manage Pull Request you will get an interactive function where you can create pull requests and view the results of pull request.

Creating and updating a pull request is done from this screen:



The Party setting allows you to select one or several upstream trading partners from who you want to retrieve product information. You can also select your own party in order to download your own product information stored in the Product Data Lake.

The Profile setting is used to select technical preferences for the download as well as assigning the FTP domain where the results will be placed. Please refer to the Manage My Party User Guide for more explanation.

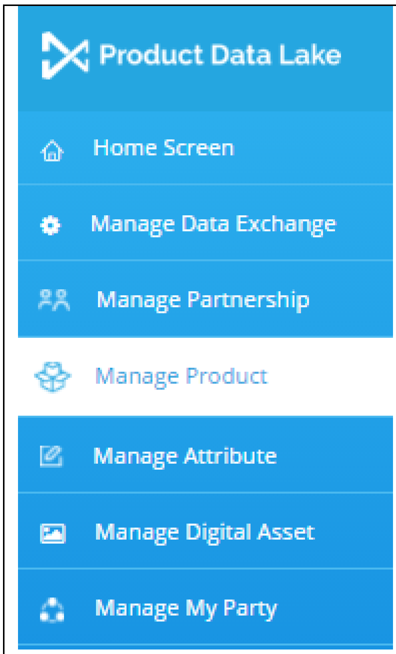
The File Type setting is used to select the data structure(s) to be downloaded.

The Data Content can be set to either Delta Change or Full Package. For Delta Change a period of Data Duration can be set to catch changes made in this time span.

The Time Zone setting allows you to schedule the pull in your preferred time zone. The pull is always converted to UTC date and time.

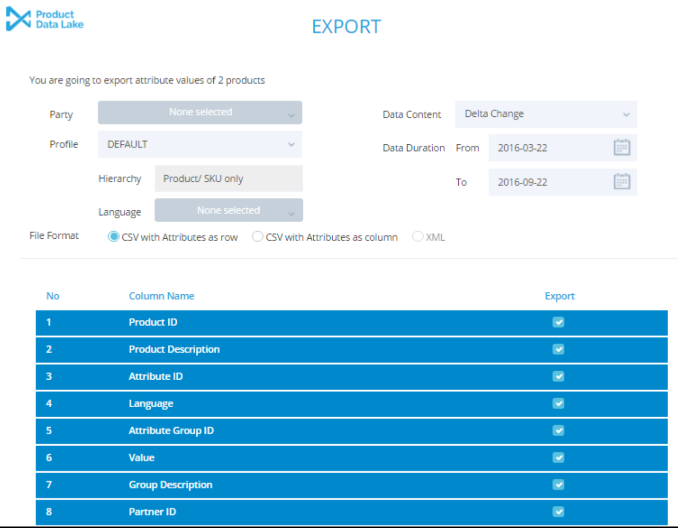
The Frequency setting can be used to repeat the pull request and thus automate the influx of product information from your trading partners and data reservoirs.

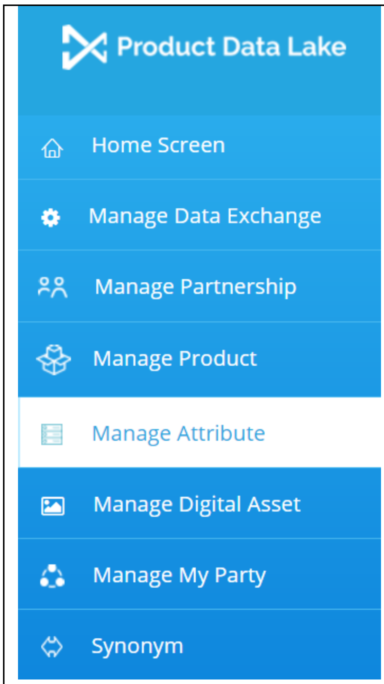
Interactive Export



An export functionality is available from the Manage Product screen.

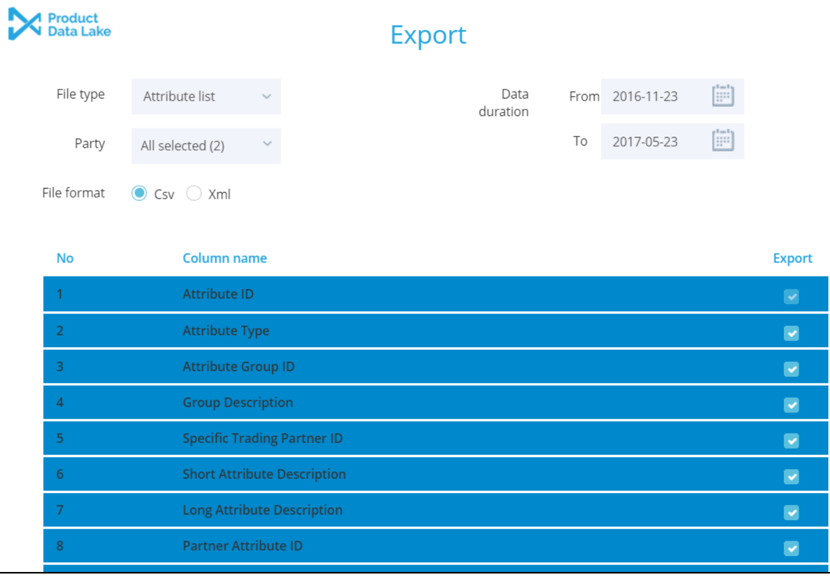
This export is anchored on the selected products and allows you to select preferences, format and included content. Please refer to the Manage Product User Guide for more explanation.



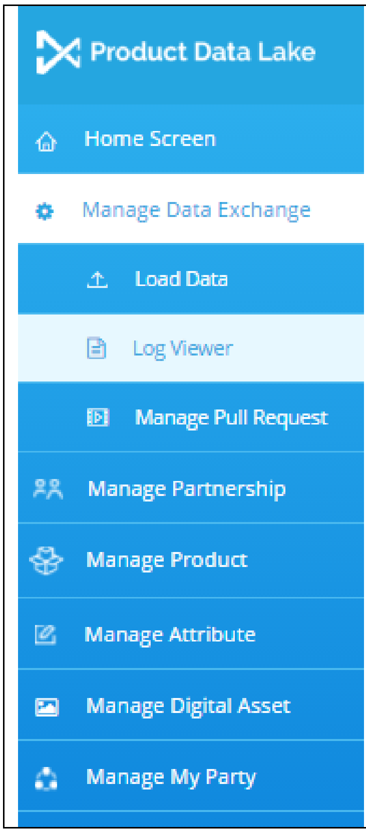


An export functionality is available from the Manage Attribute screen.

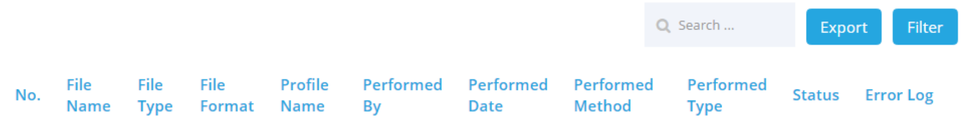
This export allows you to select encompassed partners, preferences, format and included content. Please refer to the Manage Attribute User Guide for more explanation.



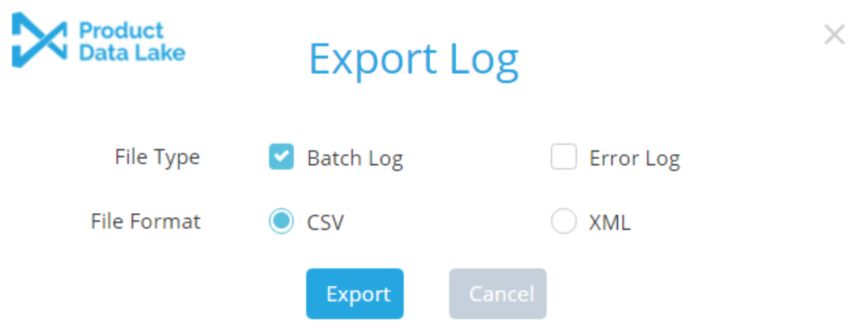
Log Viewer



The log viewer shows logs from any kind of upload to and download from the Product Data Lake.



Logs can be exported:



Other Data Structures

It is the intension of the Product Data Lake, that other data structures than those supported directly by upload functionality and linking in the Product Data Lake can be stored and processed to the supported generic formats.

One way of doing this uploading to the FTP domain referenced in the Product Data Lake and engaging with ambassadors and reservoirs that participates in the Product Data Lake, who will be able transform the data to the exchangeable format and push the data to the work folder in the hot folder for further linking and transformation.

Additional documentation available

For an introduction to the Product Data Lake, please refer to the Product Data Lake Overview.

For how to get a party account on the Product Data Lake, please refer to Product Data Lake Getting Started.

For how to operate further interactive functionality on the Product Data Lake, please refer to the other Product Data Lake User Guides.

Product Data Lake contact information:

By phone: Denmark: +45 40410528. By eMail: support@productdatalake.com