

Service Directory Tools Documentation

Combiner App

This is a console application that has a hard coded list of service directory base URLs that are then read and put into a single database instance. However only the web service data that matches the [data package](#) is imported to ensure it matches the database structure. The database is wiped on each export.

Lessons Learned

Child data ids need to be converted either to unique ids within the database or guids and then universally updated throughout the imported data, which significantly complicates data combining. I think the question here is should all ids be guid from the start as otherwise there will always need to be a conversion to a unique id.

Google Sheets Exporter

The [Google sheets exporter](#) exports a selected service directory to a Google sheet where every entity is represented by a sheet. Only fields that are in the [data package](#) are included in the results, this ensures it can be worked out the structure correctly.

Foreign keys are then turned into constraints to ensure that any values added into a foreign key column match the column that it should be referencing.

A customizable config file has also been created that allows columns to be hidden from the export, comments to be added to column headers. As well as additional columns that can be defined to provide features such as a lookup on column values, and concatenation of values to help with the understanding of the data.

The extent column is excluded from the spreadsheet due to the size of the column.

Validator

The [Service Directory Validator](#) reads all data in the service directory to determine if it is following the standard; this is because some values may not be output because they are null.

The web service checks that it performs are:

- The service is paginatable with the form: `/services?page={nn}`
- The service has a detail page with the form: `/services/{id}`

It then applies the following rules to each column were the [data package](#) says the rule should apply, the data checks that are performed are:

- Required - it checks that there is at least one value for a column
- Unique - it checks that all of the values for a column are unique
- Format - it checks that it matches the uuid, email, url, currency or year format
- Data type - it check that it matches the number, boolean, data, time, datetime or string
- Enum - it checks that it matches the specified enum values

Additionally it provides some statistics on the service directory in the form of a count of resources by types.