BEXIS 2.11.2
Data Dissemination Module

User Guide

Authors

Nafiseh Navabpour, Roman Gerlach, David Blaa

Contact

Website: http://bexis2.uni-jena.de
Email: bexis-support@uni-jena.de
Phone: +49-(0)3641-948968

Acknowledgement

*The development of the BEXIS 2 software would not be possible without the German Research Foundation (DFG) funding the BEXIS++ project. BEXIS++ is a collaboration of the Friedrich-Schi*ller-University Jena *(Dept. of Computer Science, Dept.* for Geography*), the Max-Planck-Inst*itute for Biogeochemistry Jena, the Technical University Munich (Dept. Ecology and Ecosystem Management), the Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG), and the Georg-August-University Göttingen.

Content

1 Overview ..... **Error! Bookmark not defined.**

2 Mapping tool 3

2.1 Overview 4

2.1.1 Source and Target 5

2.1.2 Mapping 6

3 Publishing a Dataset Version 9

3.1 Administration ..... **Error! Bookmark not defined.**

3.1.1 Configure ..... **Error! Bookmark not defined.**

3.1.2 Publish 9

3.1.3 Zip ..... **Error! Bookmark not defined.**

3.2 User ..... **Error! Bookmark not defined.**

1 Overview

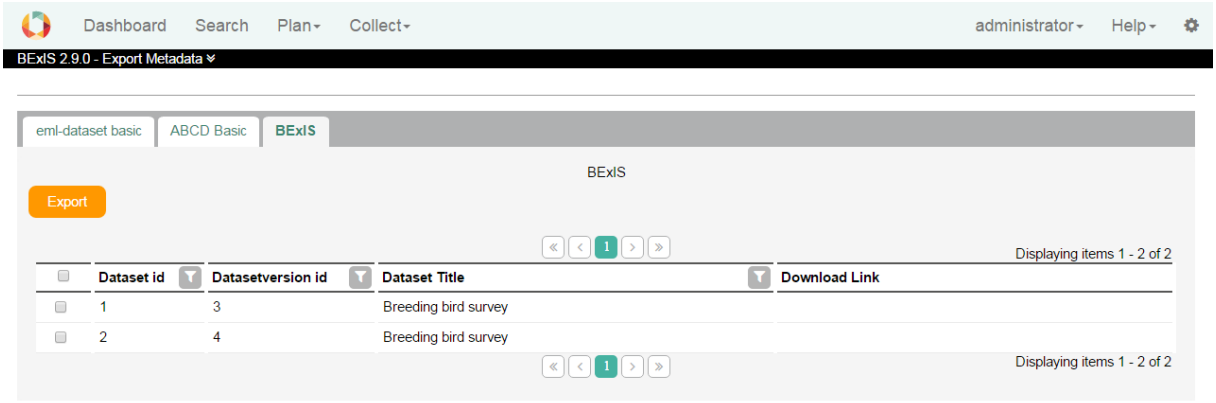
The Data Dissemination Module, available via **Setup** (Cog button) and **Export Metadata**, provides a tool to export metadata to a standard compliant XML file. For every metadata structure in the system there is one tab in the tab strip.

The data grid in one tab shows all datasets belonging to the selected metadata structure.

2 Metadata Export

Select a checkbox to mark the datasets you would like to export.

Please click the Export button and wait until the metadata XML file has been created successfully and a download link is available in-line.



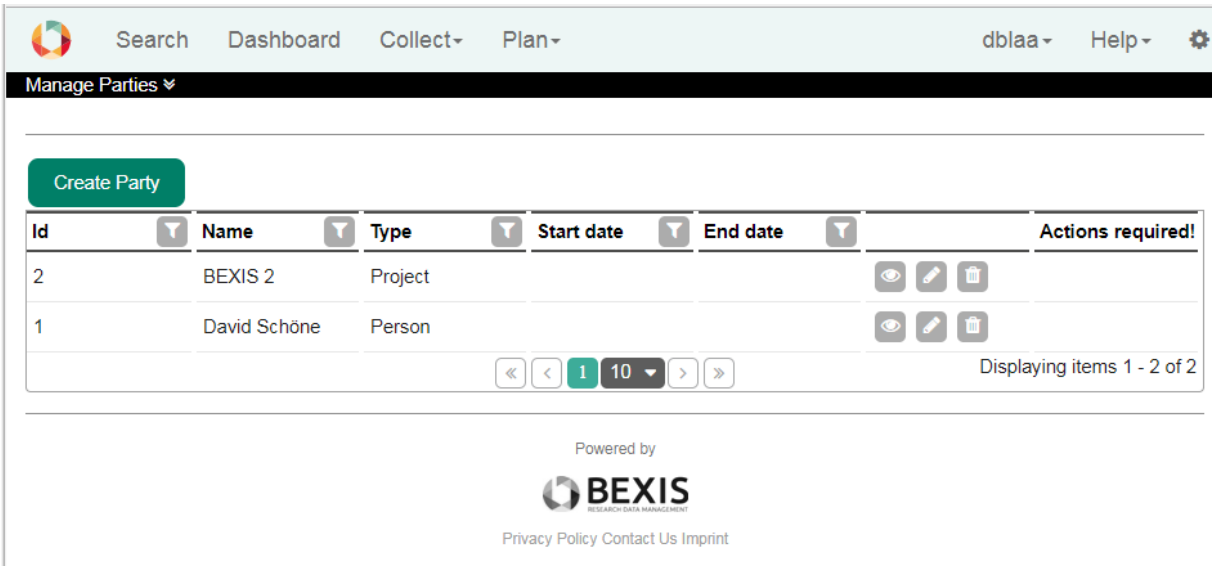
3 Mapping tool

In the mapping tool in BEXIS it is possible to set each metadata structure to predefined keys and party types (for more information about party types see the manual about parties).

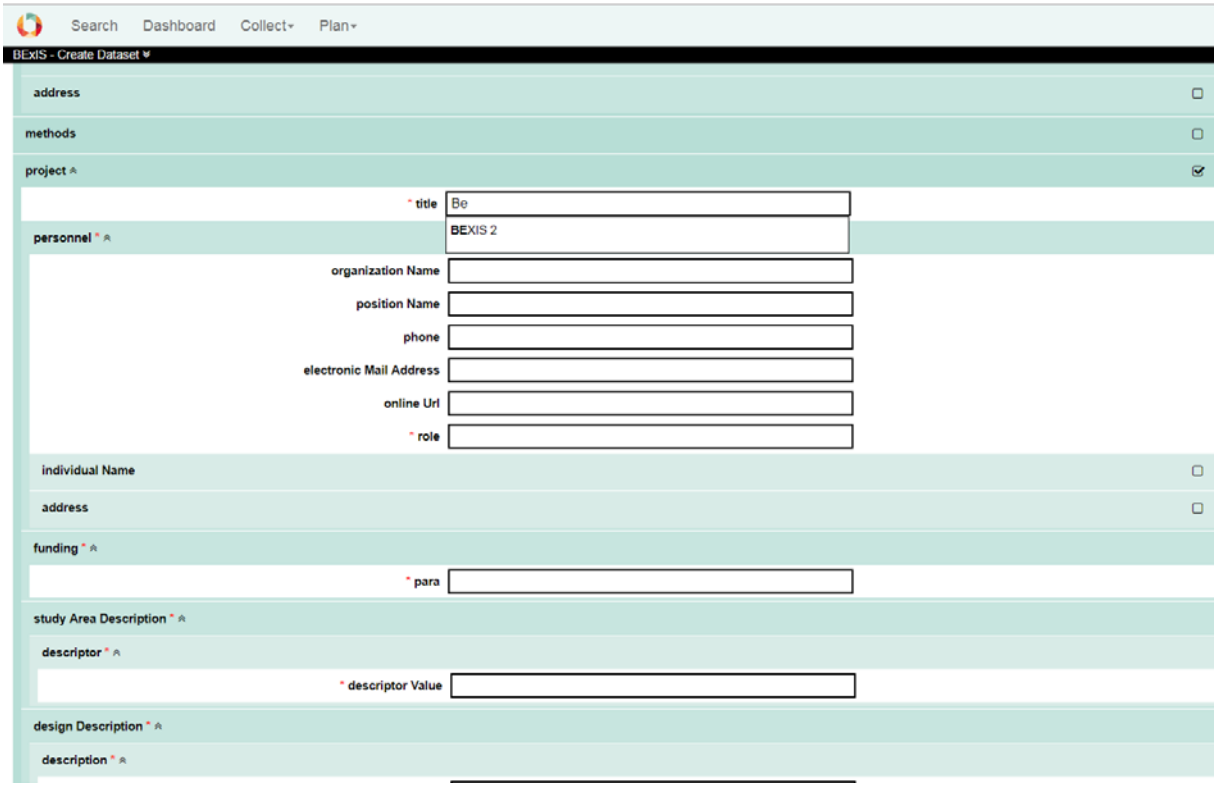
- Keys are attributes such as title or description.
- Party types are defined objects such as persons, institutes, organization or workshops.

Advantages after mapping a metadata structure:

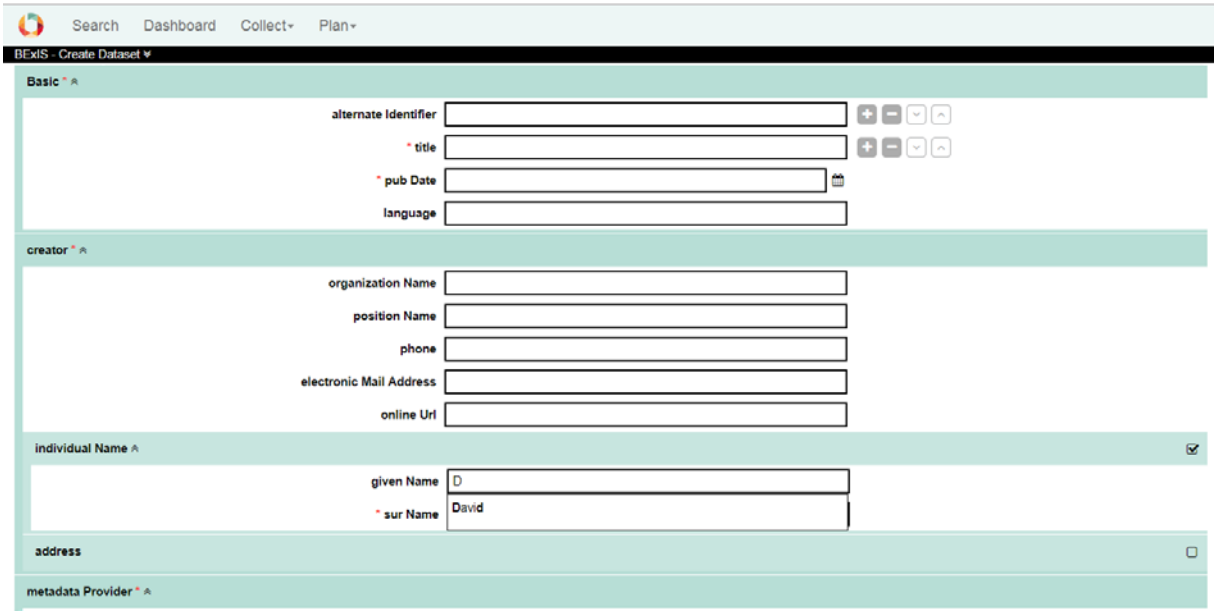
1. When publishing a dataset, BEXIS must retrieve information from the metadata and convert it to the requested formats. The more keys and party types are defined, the better the information can be prepared for publication.
2. In the BEXIS there are party types like people, project, etc.
   In the metadata form, according to the mapping, appropriate results are suggested. If a user encapsulates a person in the metadata form, all matching persons are made available for selection. This simplifies the input of metadata.



This image shows an overview of the existing Parties. In this case only one person and one exists.

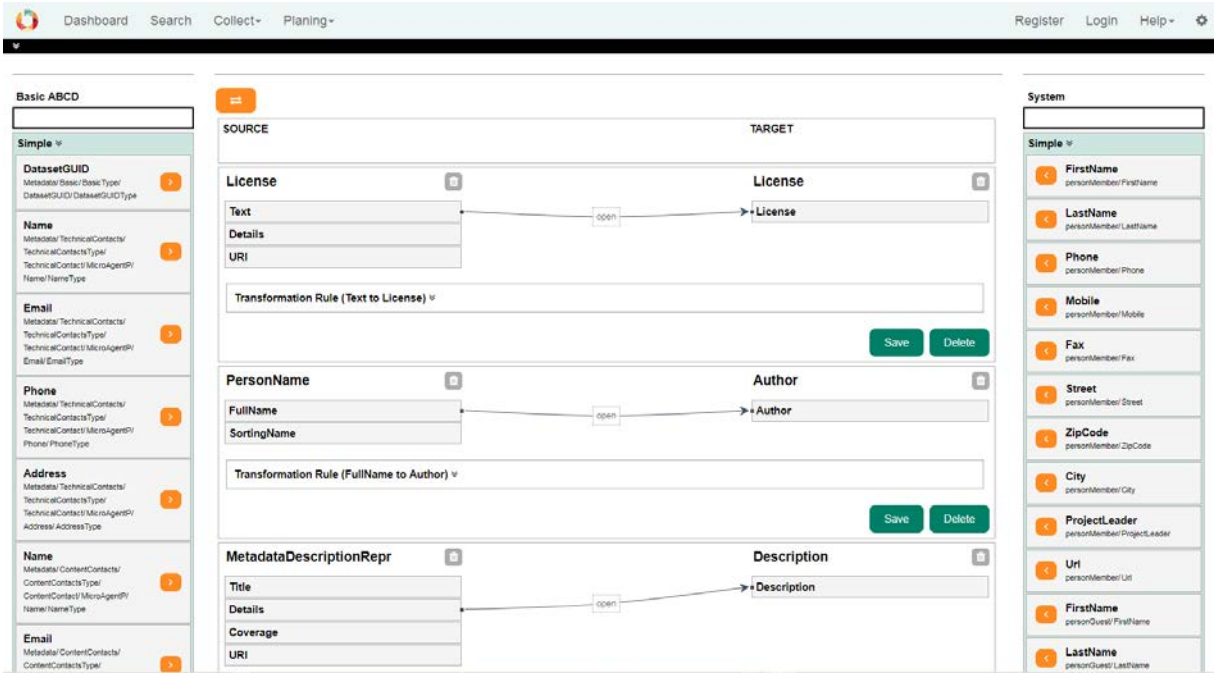


Auto complete for projects.



Auto complete for persons.

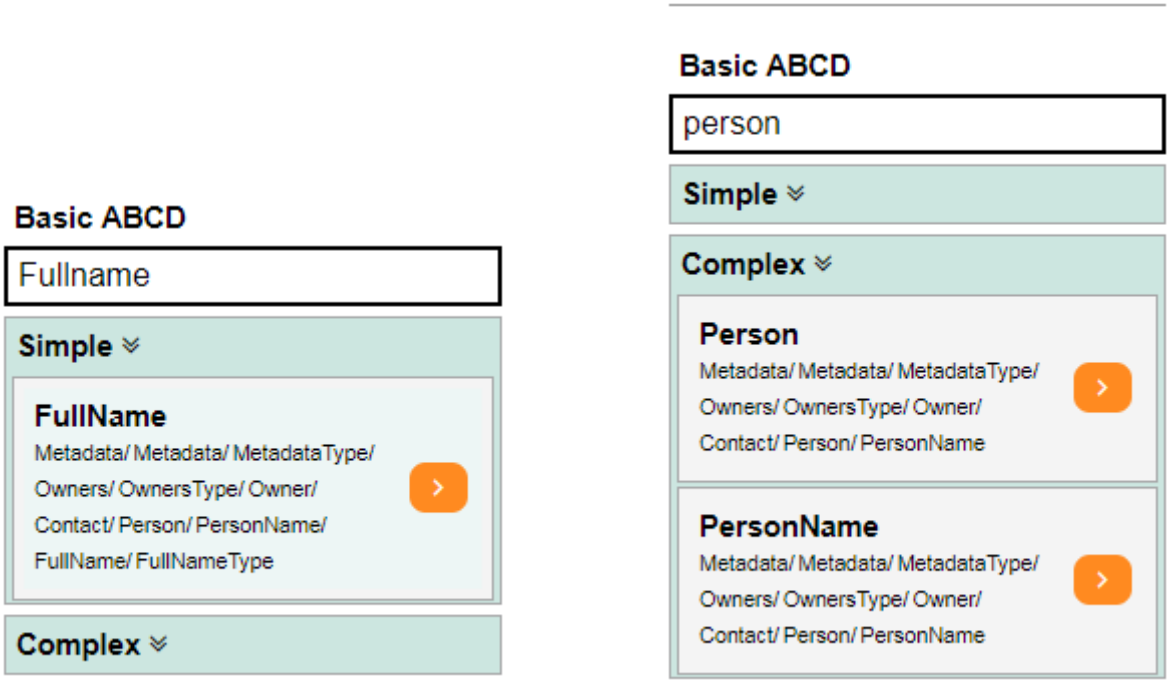
3.1 Mapping Overview



The page is divided into 3 sections. The source is displayed on the left. The target is displayed on the right side. In the middle all created mappings are displayed.

3.1.1 Source and Target

Each side as a simple and complex block as also a free text search. Simple elements are example first name, last name or full name from a person. A Complex type can be a person.



3.1.2 Mapping

Mappings are connections between the source and the target. There are different connection possibilities between the simple attributes. Generally only the connection between two simple attributes is considered.

With the aid of a transformation rule, it is possible to cover a wide range of different cases. A transformation rule consists of a RegEx and a mask. With an example you can check the values and the expected result.

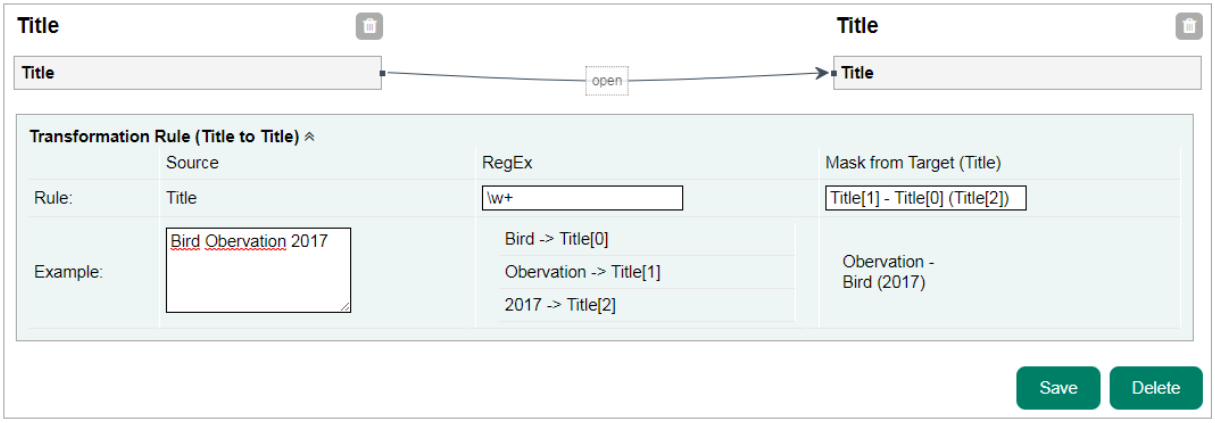
[RegEx Documenation (mdsn)]

3.2 Mapping Examples

Following is some examples of one to one, one to many and many to one mapping.

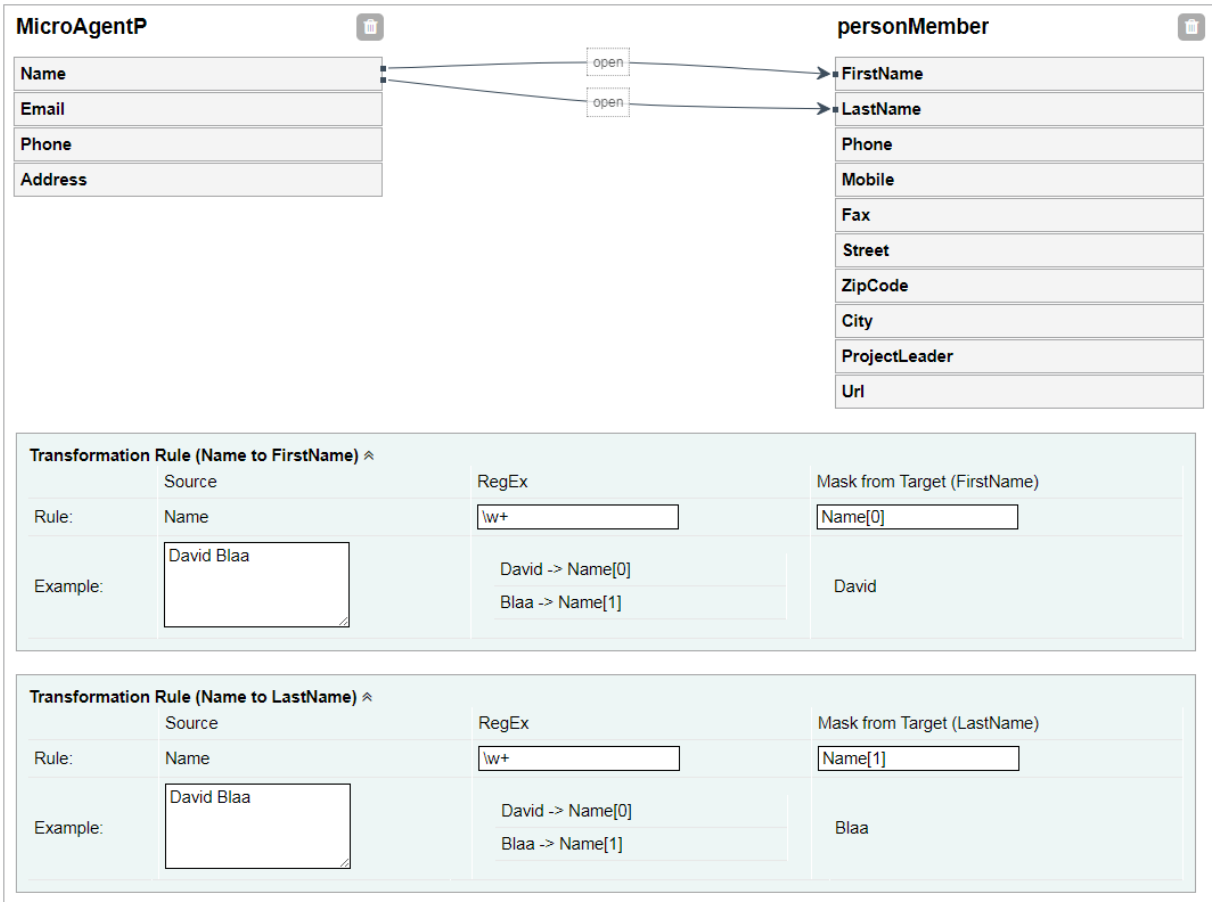
3.2.1 EXAMPLE one to one

This example creates a connection between 2 titles. All words are separated by a RegEx and then arranged differently via the mask.



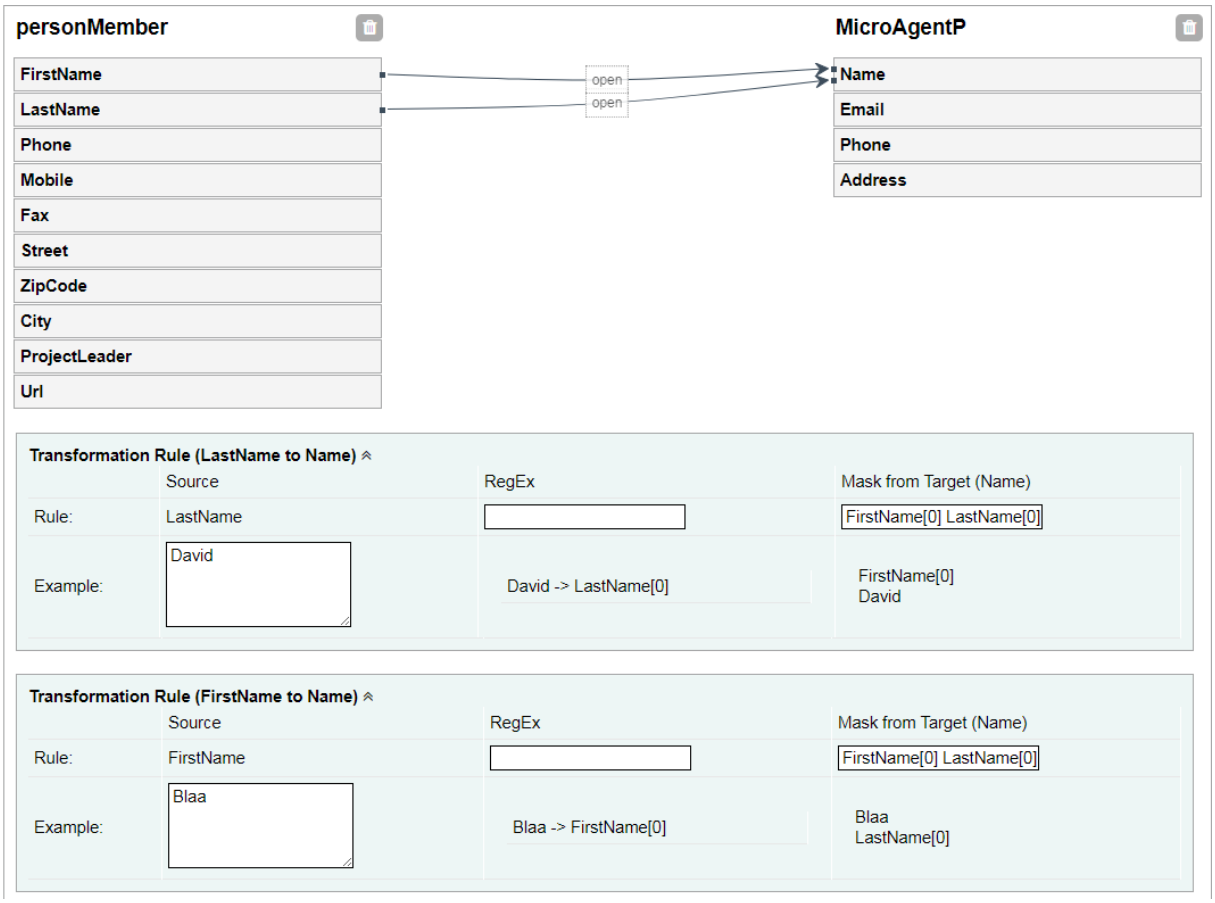
3.2.2 EXAMPLE one to many

This example creates a connection between 1 name and a FirstName and LastName. In the transformation rule, the first and last names are separated from one another by a RegEx and then positioned in the mask via the variable.



3.2.3 EXAMPLE many to one

This example creates a connection between the FirstName and LastName by name. Here is no RegEx needed but the mask ordered from both variables.



3.3 Create a mapping

1. Search a select for a simple or complex element from the source.
2. Add Element to the mapping in the middle by clicking the orange arrow next to the element.
3. Search a select for a simple or complex element from the target.
4. Search a select for a simple or complex element from the target.
5. Create the mapping by clicking the create button
6. In the mapping container there are all available simple elements for this mapping. Draw a line by clicking on one simple element from the source side and drag it to a simple element on the target side.
7. If needed, add RegEx and mask to the transformation rule. After entering values in the blocks
8. Press save

4 Publishing a Dataset Version

It is possible to prepare the data for 2 brokers and a total of 3 data repositories.

- Broker
  - GFBIO
  - Pensoft

- Data Repositories
  - GFBIO – Collections
  - Pangaea
  - Pensoft

There is a limitation for Pensoft. All data prepared for Pensoft must be generated with the GBIF metadata structure.

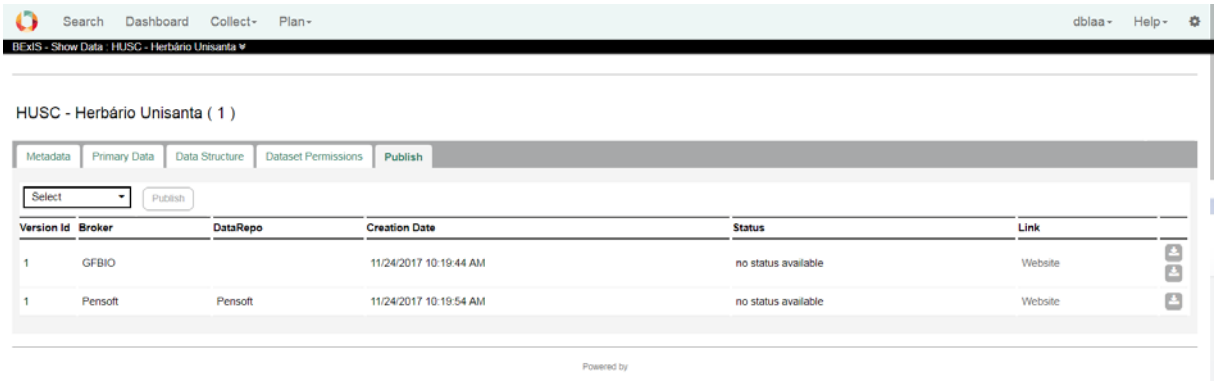
It is only possible to prepare datasets if the metadata is valid.

4.1 Publish

The user can publish a dataset version if you go to the dataset and find the publish tab.
The dropdown is showing all available data center. Select on and the system try to convert the data and the metadata as defined in the submissionConfig.xml. If something fails a message will displayed.

There are two types of fails:

1) The system is not able to convert the data.
2) Metadata is not valid. This is a warning. You can go on but the metadata.xml in the zip ist not valid against the exported xsd schema



4.2 GBFIO

Via the GFBIO portal you can start a submission and publish your dataset. Depending on the subject of the data set, a suitable Data Repository is defined. There are different main Types.

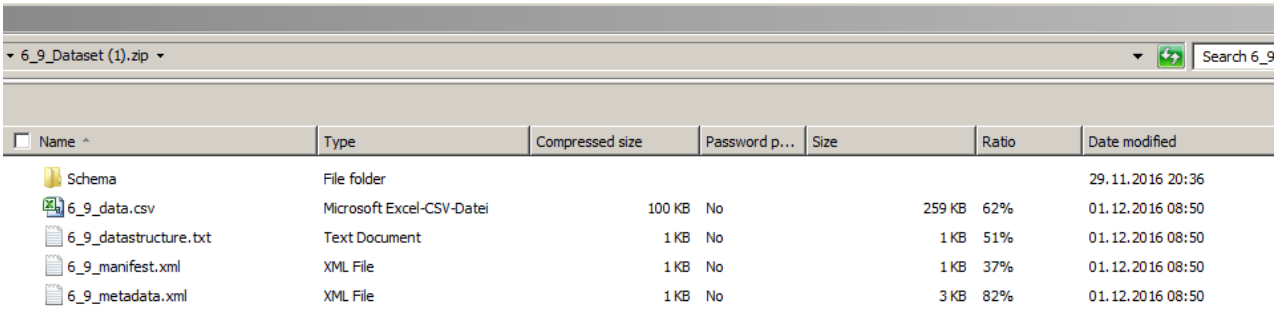
- Pangaea
- Collections

- ENA

Each data repository has different data requirements. BEXIS 2 offers an export for Pangea and Collections.

4.2.1 Collections

The data for the collections is stored in a zip file.
The following files can be found in the zip file.



1. Schema - XSD Schema for the metadata
2. Data.*** - Primary Data
3. Data structure - Structure of the primary data
4. Manifest File - General information's about the Dataset
5. Metadata - Metadata information's about the dataset

4.2.2 Pangaea

For Pangea, the metadata and primary data are stored in a text file.


```
{
  "AuthorIDs": [
    "Paulo",
    "Paulo",
    "",
    "Zélia",
    "Leonor",
    "de S. Penteado Sampaio",
    "de S. Penteado Sampaio",
    "speciesLink Network",
    "Rodrigues de Mello",
    "Costa Maia"
  ],
  "Title": "HUSC - Herbário Unisanta",
  "Abstract": "O acervo conta com 8.131 espécimes (briófitas, pteridófitas, macroalgas marinhas e angiospermas). O principal objetivo é proporcionar a comunida
  "ReferenceIDs": [],
  "ExportFilename": "HUSCHerbrioUnisanta",
  "EventLabel": "",
  "ParameterIDs": [
    {
      "Id": 3,
      "Label": "TimeUTC",
      "Description": "",
      "isOptional": true,
      "unit": {
        "Id": 12,
        "Name": "none",
        "Description": "",
        "Dimension": {
          "Name": "dimensionless",
          "Description": "",
          "Specification": ""
        },
        "MeasurementSystem": "Unknown"
      },
      "dataType": {
        "Id": 5,
        "Name": "datetime",
        "Description": "A date with / or time",
        "SystemType": "DateTime"
      }
    },
    {
      "unit": {
        "Id": 12,
        "Name": "none",
        "Description": "",
        "Dimension": {
          "Name": "dimensionless",
          "Description": "",
          "Specification": ""
        },
        "MeasurementSystem": "Unknown"
      },
      "dataType": {
        "Id": 2,
        "Name": "double",
        "Description": "A real number",
        "SystemType": "Double"
      }
    }
  ],
  "ProjectIDs": [
    "INCT - Herbário Virtual da Flora e dos Fungos"
  ],
  "TopologicTypeID": 0,
  "StatusID": 0,
  "LoginID": 0
}

TimeUTC;D8CO1_1;D8CO1_3;D8CO1_4;D8CO1_5;D8CO1_6;D8CO2_1;D8CO2_3;D8CO2_4;D8CO2_5;D8CO2_6
2017-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2018-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2019-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2020-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2021-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2022-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2023-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2024-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2025-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2026-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2027-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2028-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2029-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2030-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2031-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2032-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2033-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
2034-02-21T12:00:00;1;12;22;223;3443;434;44;43;4;4
```