BEXIS 2.4.0

Data Collection Module

User Guide

Authors

Nafiseh Navabpour, Roman Gerlach, David Blaa

Contact

Website: <http://fusion.cs.uni-jena.de/bexis>

Email: bexis-support@uni-jena.de

Phone: +49-(0)3641-948968

Acknowledgement

The development of the BExIS 2.4.0 software would not be possible without the German Research Foundation (DFG) funding the BExIS++ project. BExIS++ is a collaboration of the Friedrich-Schiller-University Jena, Germany (Dept. of Computer Science, Dept. for Geography, Dept. of Ecology) and the Max-Planck-Institute for Biogeochemistry Jena, Germany.

Content

[1. Overview 3](#_Toc385425176)

[2. Create a new dataset 3](#_Toc385425177)

[2.1 Content 4](#_Toc385425178)

[2.2 Step List 5](#_Toc385425179)

[2.3 Navigation 5](#_Toc385425180)

[2.4 Messages 5](#_Toc385425181)

[3. Upload Data 6](#_Toc385425182)

[3.1. Start 6](#_Toc385425183)

[3.2. Select File 6](#_Toc385425184)

[3.3. Get File Information 7](#_Toc385425185)

[3.4. Specify Dataset 8](#_Toc385425186)

[3.5. Define Primary Key 8](#_Toc385425187)

[3.6. Validation 8](#_Toc385425188)

[3.7. Summary 9](#_Toc385425189)

# Overview

The Data Collection Module provides tools for creating new datasets, enter metadata, and upload data to the system. There are two workflows available under the **Collect** tab:

* Create a dataset
* Upload data to a dataset

# Create a new dataset

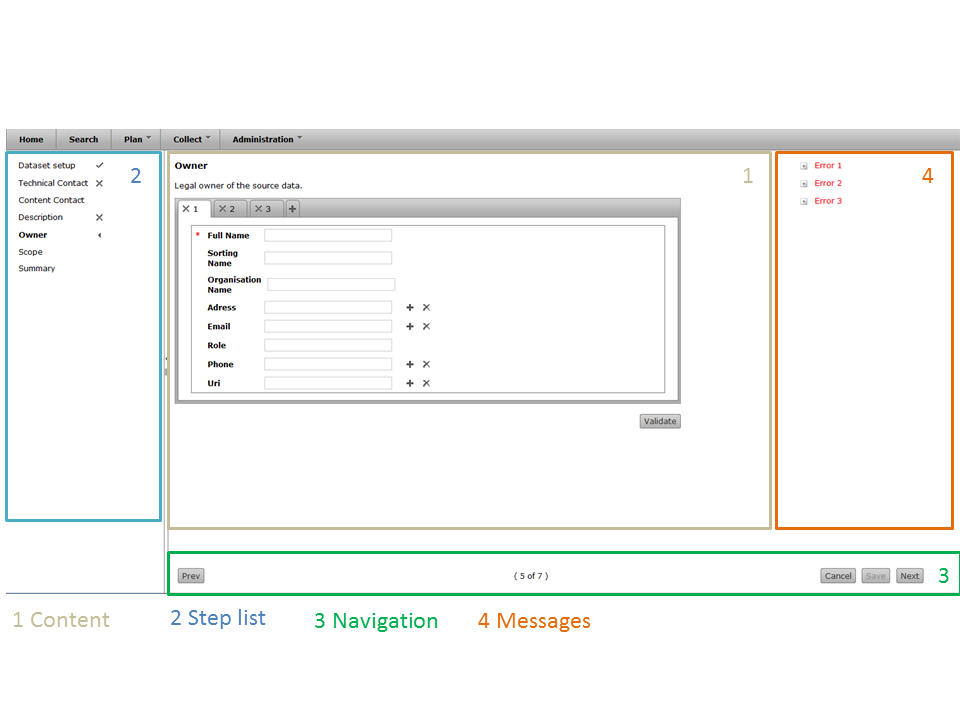
This wizard will assist you in creating a new dataset in BExIS. The Wizard is very flexible and builds up differently depending on the selected Metadata structure. Therefore, we describe only the basic functions here. The first and last steps are always the same, however.

The first step is to generate an empty dataset based on your selection of the three mandatory elements: research plan, data structure, and metadata structure.

On the last page a summary of the entered content is provided for your review; also pointing out potential errors (e.g. missing values in mandatory fields).

The intervening steps are determined by the selected metadata structure.

The user interface (UI) is divided into four parts which are described hereafter.



## 2.1 Content

The content area in the center is where you enter metadata describing your dataset. The forms provided here may look different and contain different attributes depending on the metadata schema (structure) you have chosen in the first step.

In each step, there is a button titled Validate to examine whether required attributes have been filled and whether the information complies with the business logic. The validation may also be triggered by using the Next button in the navigation panel or by navigating through the step list in the left panel.



 Check whether the entered values ​​are valid. Messages are displayed on the right.

Add an attribute.

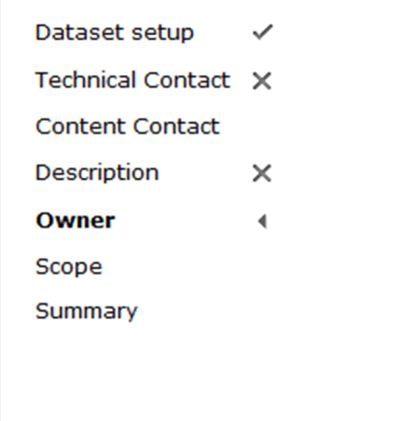


Remove an attribute.



Change order of the attribute.



 Change the active tab. Active tab highlighted in white. Remove tab

Add another attribute collection (package) of the same type



Required attributes



## 2.2 Step List

It shows you an overview of the steps you need to do. The icons next to the step show you how the status of these steps.

|  |  |
| --- | --- |
|  | Successfully validated.  Indicates an error in the respective form  Shows the active step.  Go directly to a step when you click on it. |

## 2.3 Navigation

Below you will find more navigation options.

Return to previous step



Move one step forward



Deletes everything and starts with the first step.



Show the total number of steps and the position of the active step.



## 2.4 Messages

On the right side information will be displayed if the steps were incorrectly processed or successful.

|  |  |
| --- | --- |
|  | Messages are displayed in two different colors.  Red means there is an error  Green indicates no error present.  For getting more information about an error, click the plus button to open the details. |

# Upload Data

This wizard will assist you in uploading data into the BExIS repository. A dataset can be structured or unstructured

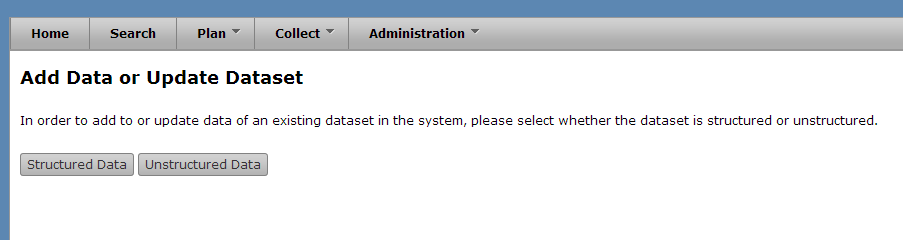
**Structured Dataset**

With the currently supported workflow it is assumed that your data file is based on an Excel template defined and downloaded earlier on from the BExIS system. In that way it is assured that the data structure of the data file and the data structure of the repository are identical.

**Unstructured Dataset**

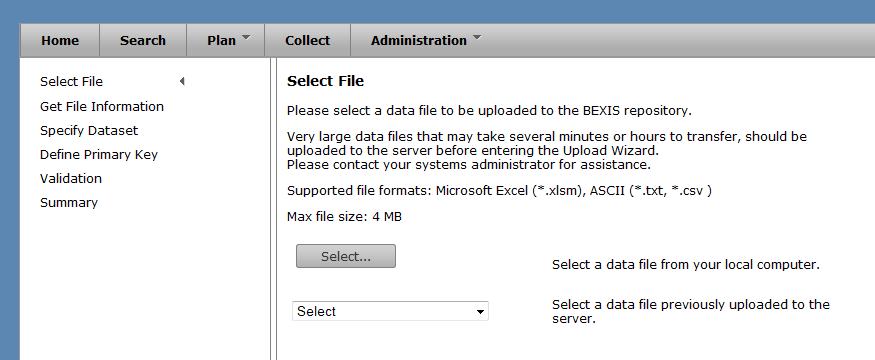
In the case of unstructured data, we do not read the contents of the data. We copy the files to the server and place them in relation to the dataset.

## Start



To upload your data, please go to the **Update data to a dataset** tabunder **Collect** tab.

## Select File



In the first step an existing file containing your data needs to be selected. You can either select a file from your local computer or a file that has been uploaded to the server prior to starting the Upload Wizard. The second option is designed for files larger than 4 MB that may take several minutes to transfer. There is currently no user interface for this separate data upload, thus you are requested to contact your systems administrator for assistance. The wizard supports file formats of Microsoft Excel (\*.xlsm) or ASCII (\*.txt, \*.csv). Microsoft Excel files are required to use a template created with the Data Planning Module (Plan) of BExIS 2 (refer to Data Planning User Guide for more details). Once a file has been successfully selected, click the **Next** button and proceed to the next step.

## Get File Information

For all Microsoft Excel files using a BExIS 2 template the file information and data structure is automatically extracted and this step is omitted. Please refer to the Data Planning User Guide for more details on how to create such a template.

For all ASCII files users need to provide information on the file structure and formatting.

First, please choose a **separator** that is being used to separate data values from each other in your ASCII file.

Depending on your language different punctuation is used for decimal values. Please choose the one present in your ASCII file.

Next please specify whether the **orientation** of your data is column-wise or row-wise (see figure below).



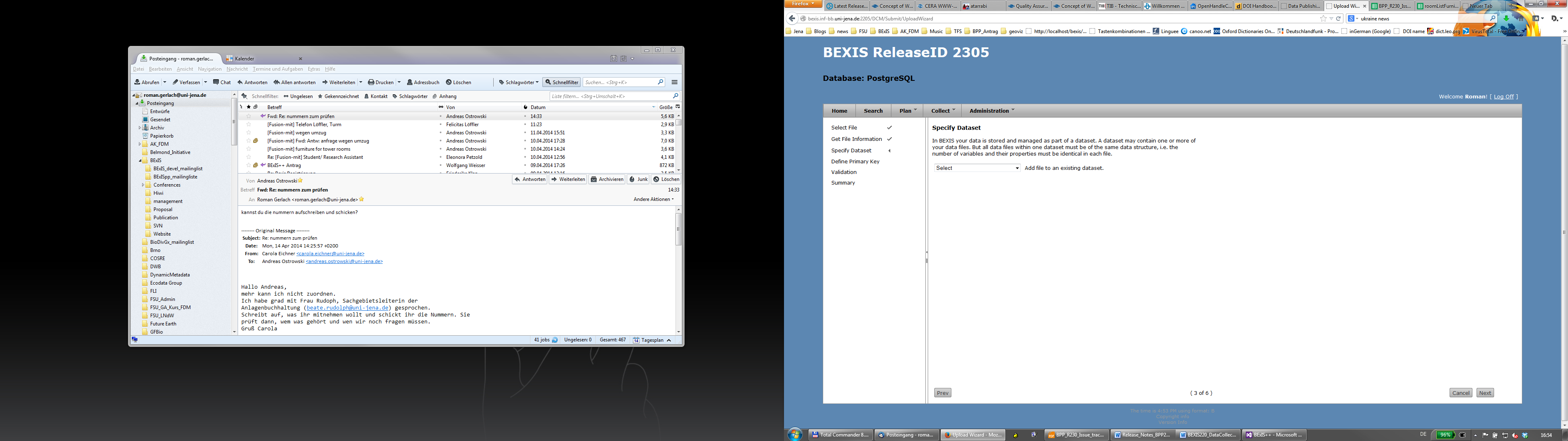
Data sets may contain empty rows or columns on top or to the left before the header and the actual data values start. Please specify this offset in number of columns or rows.

Further, your data file may contain a header defining variable names, types etc. The row/column where this header starts needs to be specified (see figure below).



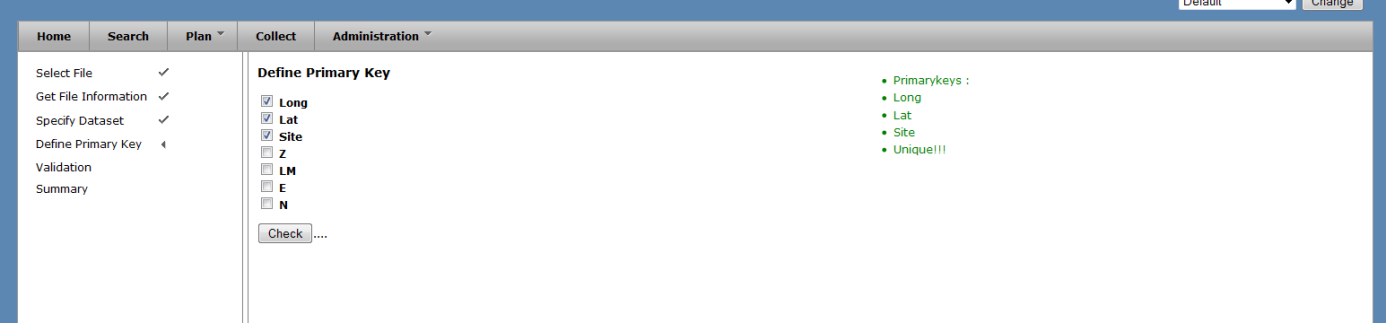
Finally, the row/column where the actual data values start needs to be specified.

## Specify Dataset



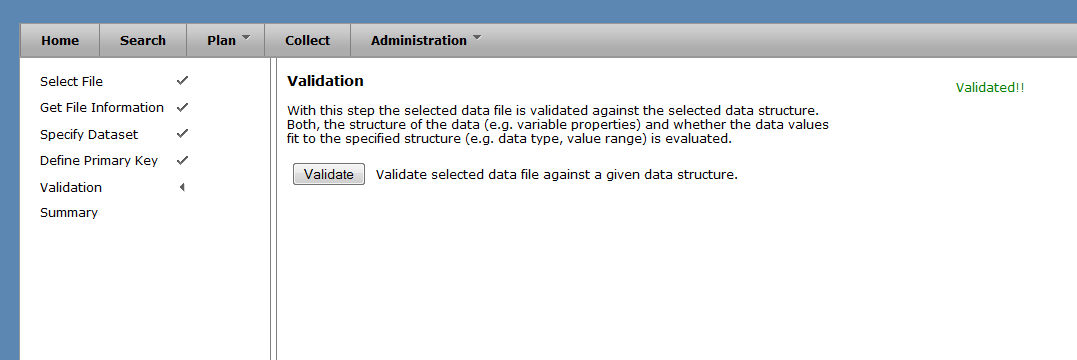
In BExIS your data is stored and managed as part of a dataset. A dataset may contain one or more of your data files. But all data files within one dataset must be of the same data structure, i.e. the number of variables and their properties must be identical in each file. To upload your data to the system, please select one of the existing dataset from the dropdown list.

## Define Primary Key



While adding data to an existing dataset you need to specify a unique identifier (e.g. primary key) for your dataset. If your dataset already contains a variable with such a key please select it. Otherwise a primary key can be created by combining available variables. Please click the **Check** button to verify whether the selected combination is unique. If you go back and change something in the process of uploading, you need to check the primary key again.

## Validation

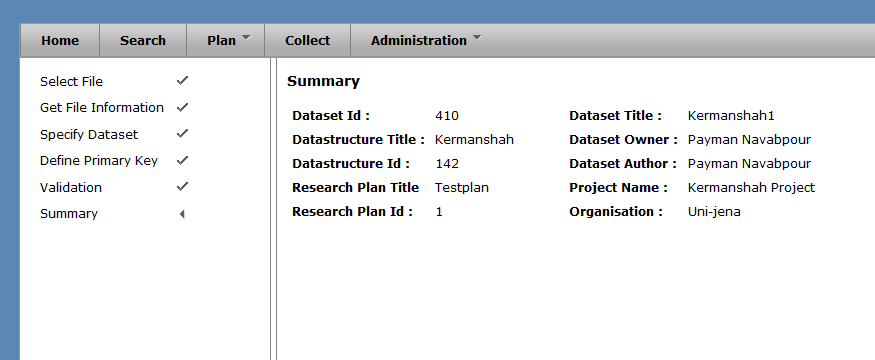


With this step, the selected data file is validated against the selected data structure. Both, the structure of the data (e.g. variable properties) and whether the data values fit to the specified structure (e.g. data type, value range) is evaluated.

Click on **Validate** button to validate the data file.

If you go back and change something in the process of uploading, you need to validate the file again.

## Summary



With this final step a summary of your uploaded data file is provided. Please check the information and click the **Finish** button to confirm and finalize the upload.

## Push big files to server

Each user has its own folder on the server where it can store files.   
On this page you can see the uploaded files can upload and delete one or several of files.

