

# BEXIS 2.2.0

## 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](mailto:bexis-support@uni-jena.de)

Phone: +49-(0)3641-948968

#### **Acknowledgement**

The development of the BEXIS 2.2.0 software would not be possible without the German Science 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**

Overview .....3

Start .....3

Select File .....3

Get File Information.....4

Specify Dataset .....5

Define Primary Key .....5

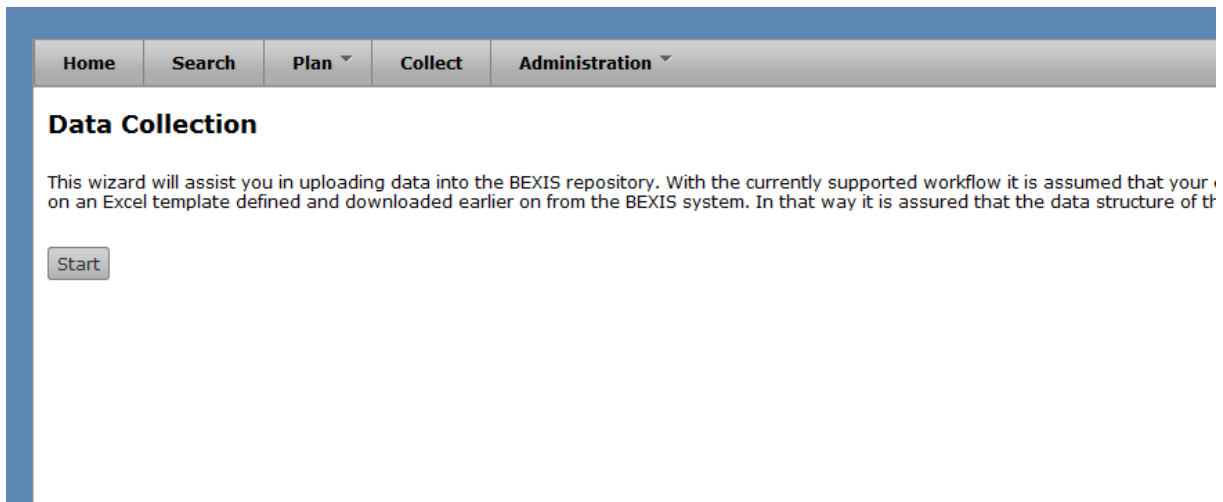
Validation .....6

Summary .....6

## Overview

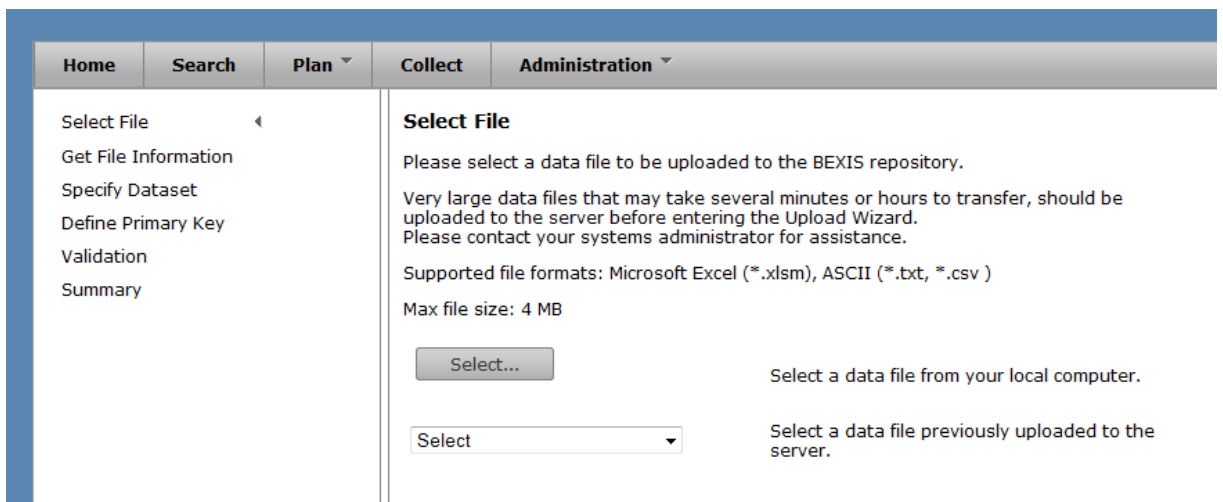
To upload a data set, you have to start in **Collect** tab. This wizard will assist you in uploading data into the BEXIS++ repository. With the currently supported workflow it is assumed that your data file is based on an Excel template defined and downloaded earlier on from the BEXIS system. In that way it is assured that the data structure of the data file and the data structure of the repository are identical.

## Start



To upload your data, please go to the **Collect** tab and click on **Start** button.

## 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 (\*.xism) 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

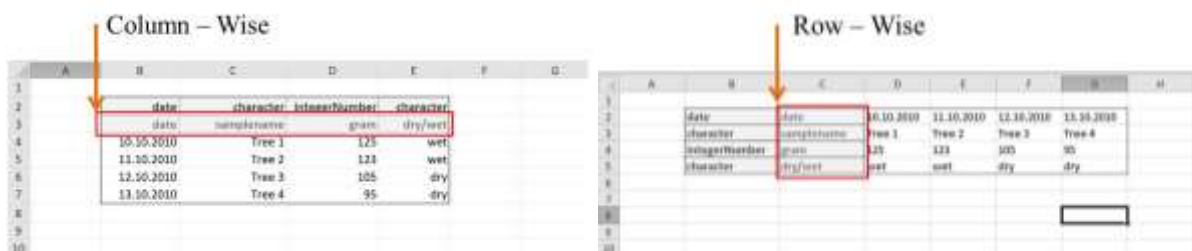
For all Microsoft Excel files using a BEXIS 2 template the file information and data structure is automatically extracted and this step is omitted. Please refer to the Data Planning User Guide for more details on how to create such a template.

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).



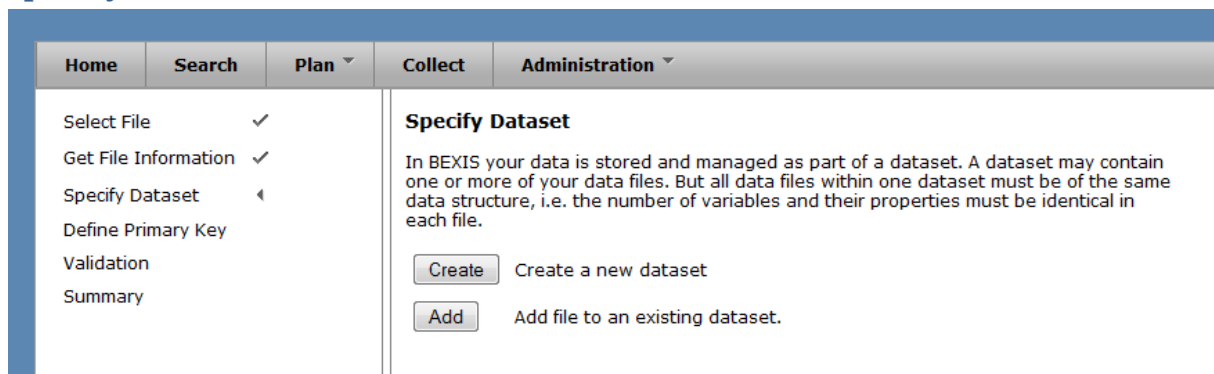
Data sets may contain empty rows or columns on top or to the left before the header and the actual data values start. Please specify this offset in number of columns or rows.

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).

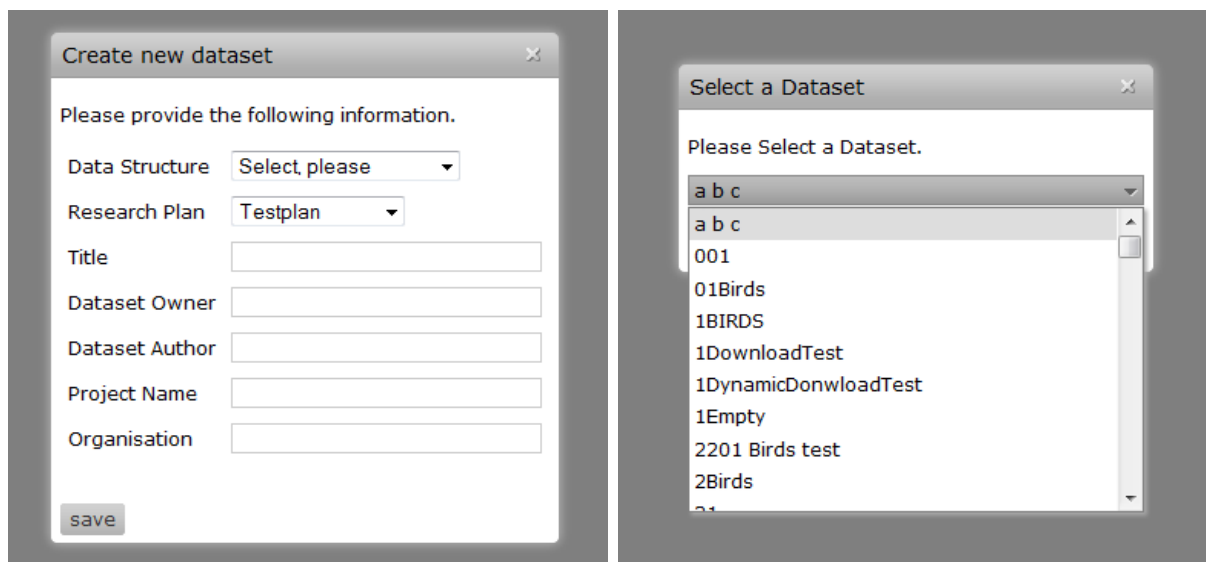


Finally, the row/column where the actual data values start needs to be specified.

## 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. You can either create a new dataset or add your data to an existing one (see figure below).



## Define Primary Key



This step is omitted if you created a new dataset in the previous step. If you added data to an existing dataset you need to define 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

The screenshot shows a software interface with a navigation menu on the left and a main content area. The navigation menu includes: Home, Search, Plan (with a dropdown arrow), Collect, and Administration (with a dropdown arrow). Below the menu, a list of steps is shown: Select File (checked), Get File Information (checked), Specify Dataset (checked), Define Primary Key (checked), Validation (selected), and Summary. The main content area is titled "Validation" and contains the following text: "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." Below this text is a "Validate" button and the instruction "Validate selected data file against a given data structure." In the top right corner of the main content area, the word "Validated!" is displayed in green.

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

The screenshot shows the same software interface as the previous step, but now the "Summary" step is selected in the navigation menu. The main content area is titled "Summary" and displays the following information in two columns:

<b>Dataset Id :</b>	410	<b>Dataset Title :</b>	Kermanshah1
<b>Datastructure Title :</b>	Kermanshah	<b>Dataset Owner :</b>	Payman Navabpour
<b>Datastructure Id :</b>	142	<b>Dataset Author :</b>	Payman Navabpour
<b>Research Plan Title</b>	Testplan	<b>Project Name :</b>	Kermanshah Project
<b>Research Plan Id :</b>	1	<b>Organisation :</b>	Uni-jena

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.