

BExIS 2.7.0

Data Discovery Module

User Guide

Authors

Nafiseh Navabpour, Roman Gerlach, David Blaa, Michael Owonibi

Contact

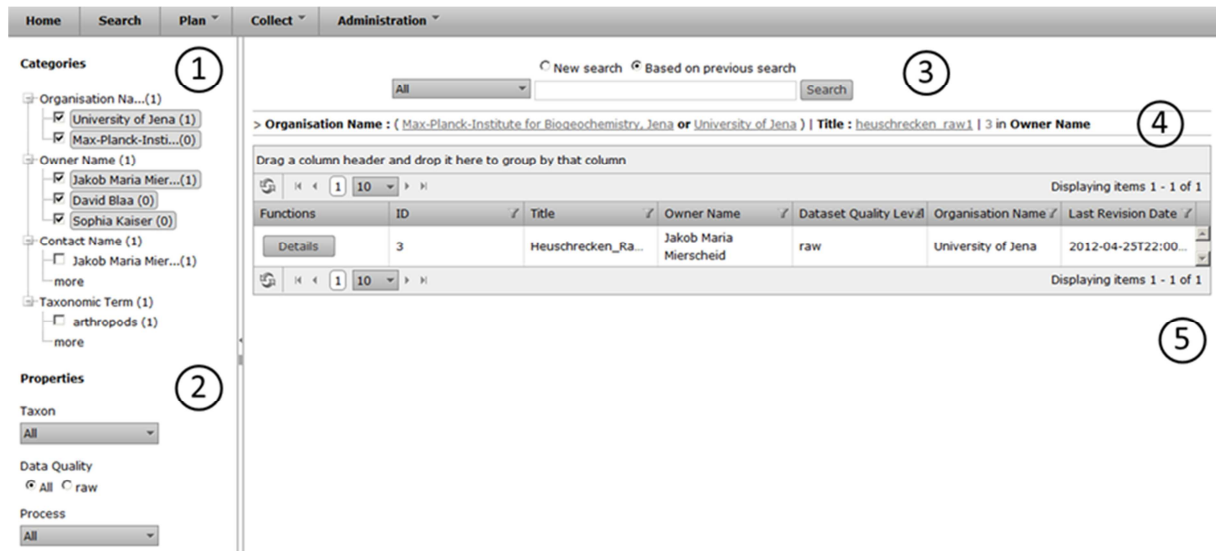
Website: <http://fusion.cs.uni-jena.de/bexis>
Email: bexis-support@uni-jena.de
Phone: +49-(0)3641-948968

Content

- 1. Search UI..... 3
 - 1.1. Categories..... 3
 - 1.2. Properties..... 3
 - 1.3. Free text search with Autocomplete..... 3
 - 1.4. Selected Filter..... 4
 - 1.5. Results 4
- 2. Data Details 4
 - 2.1. Meta Data..... 4
 - 2.2. Data Structure 5
 - 2.3. Primary Data..... 5
- 3. Search Manager..... 6

1. Search UI

Search UI contains some parts to make search easier. In this way, you are able to look for all details of datasets, but not for the uploaded data. You can control the search UI via search manager explained in section 3.



1.1. Categories

Categories are defined by number 1 in the image above. The main nodes like “Project Name” are based on nodes in the metadata. The elements are values in the metadata from a main node and can be used to restrict the current selection. The numbers next to the elements and main nodes show the number of existing data sets in the database. This list changes according to the current selection. After selection the results and the facets updated. With the “more” button it is possible to select more than one element at the same time.

1.2. Properties

Number 2 in the image above is related with properties. In this section there are predefined UI Components like dropdown, radio button or slider to filter the data. There is only one possible choice for every component. After selection the results and the facets are updated accordingly.

1.3. Free text search with Autocomplete

The free text search, number 3 in the image above, works as in any common search engine. It includes Autocomplete where words or phrases are being predicted once three letters or figures are entered. It supports also different languages.

1.4. Selected Filter

You can find selected filter part by number 4 in the image above. Every filter applied from section 1.1-1.3 is displayed here. Filters can be deselected just by clicking on it. In point 1 by the categories you are able to select more than one option to a category. If there are more than 2 you will see only the category. Click on the category and a window will open to define your selection for this category.

1.5. Results

The matching results are displayed in a table or in the list view. It marks by number 5 in the image above. Basic functions like sorting, filtering, paging are available in the table header. Additionally you are able to group columns by dragging a column to the top of the table. By right clicking on the header, you can change visibility of columns.

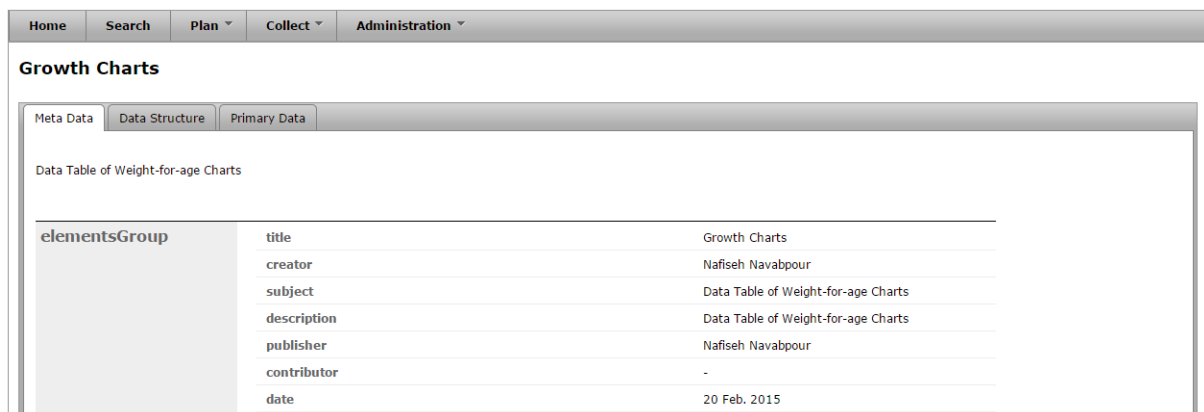
The details button opens the detailed view of the selected dataset.

2. Data Details

About each dataset created in the BEXIS, you can find a detail contains Meta Data, Data Structure and Primary Data explain later.

2.1. Meta Data

On this tab you have access to the Meta data of the selected datasets.



The screenshot shows the BEXIS interface with a navigation bar at the top containing 'Home', 'Search', 'Plan', 'Collect', and 'Administration'. Below this, the 'Growth Charts' dataset is selected, and the 'Meta Data' tab is active. The main content area displays the 'Data Table of Weight-for-age Charts' with a table of metadata.

elementsGroup	title	value
	title	Growth Charts
	creator	Nafiseh Navabpour
	subject	Data Table of Weight-for-age Charts
	description	Data Table of Weight-for-age Charts
	publisher	Nafiseh Navabpour
	contributor	-
	date	20 Feb. 2015

2.2. Data Structure

On this tab you have access to the data structure of the selected datasets.

Home Search Plan Collect Administration

LiDAR_HAI_100_0.75_Sommer

Meta Data Data Structure Primary Data

LiDAR

A LiDAR structure

Variable Name	Unit	Description
EP PlotID	None	Name for experimental plots
Ecke	None	corner of the EP which is in center, if -1, then around original plot center
Roughness average	None	surface/ground ratio
Ground hits	None	Fraction of echos hitting ground
Roughness standard derivation	None	surface/ground ratio standard derivation

Displaying items 1 - 5 of 5

2.3. Primary Data

On this tab you have access to the primary data of the selected dataset. You can download selected data in an Excel file, comma separated in a CSV file, or tab separated in a TXT file.

Home Search Plan Collect Administration

LiDAR_HAI_100_0.75_Sommer

Meta Data Data Structure Primary Data

Downloading primary data follows the What-You-See-Is-What-You-Get-Principle. So you are able to download a subset of the data by specifying filters on columns (e.g. Plot = 123), apply sorting, or select specific columns from the data table (right-click the column header). Paging does not effect the subset.

Download Excel Download Comma Separated Download Tab Separated

Drag a column header and drop it here to group by that column

EP PlotID	Ecke	Roughness average	Ground hits	Roughness standard derivation
HEW31	-1.00	412	15	279
HEW27	-1.00	29	1	158
HEW375	-1.00	331	2	211
HEW34	-1.00	115	7	32
HEW35	-1.00	353	11	265
HEW68	-1.00	256	4	222
HEW78	-1.00	28	9	229
HEW18	-1.00	266	3	182
HEW29	-1.00	314	8	269
HEW110	-1.00	172	2	92
HEW111	-1.00	214	2	128
HEW127	-1.00	186	1	102
HEW137	-1.00	289	14	196

Displaying items 1 - 13 of 13

3. Search Manager

With the help of the search manager, you can make search UI more operative.

You need to click on the Refresh Search button, to make the search result effective.

Home Search Plan Collect Administration

Search Component Manager

With this interface you are able to customize the search component and define what fields of the metadata are searchable and how they appear in the application. Changes made through the interface will alter the Lucene configuration file on the server. For a detailed description of each parameter please refer to the Lucene User Manual provided with the installation package.

Add Item include primary data Refresh Search Reset to last working setup

Actions	Display Name	Search Component Type	Data Type	Header Item	Default Header Item	Multi Value	Analyzed	
	Title	Category	String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Metadata/Description/Description/T... Metadata/Description/DescriptionE...
	Owner Name	Category	String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Metadata/Owner/Owner/FullName/...
	Contact Name	Category	String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Metadata/ContentContact/Person/... Metadata/Contact/PersonEML/Give...
	Dataset Description	Category	String	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Metadata/Description/Description/... Metadata/Description/DescriptionE...
	Dataset Status	Category	String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Metadata/DataDescription/DataDes...
	Dataset Quality Level	Category	String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Metadata/DataDescription/DataDes...
	Taxonomic Term	Category	String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Metadata/Scope/Scope/Taxonomic...
	GeoEcological Term	Category	String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Metadata/Scope/Scope/GeoEcologic...
	Organisation Name	Category	String	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Metadata/Owner/Owner/Organisati...
	Last Revision Date	Category	DateTime	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Metadata/Description/Description/...
	Owner Name	Facet	String	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Metadata/Owner/Owner/FullName/... Metadata/Contact/PersonEML/Surn...
	Contact Name	Facet	String	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Metadata/ContentContact/Person/... Metadata/Contact/PersonEML/Give...
	Taxonomic Term	Facet	String	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Metadata/Scope/Scope/Taxonomic...
	Taxon	Property	String	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Metadata/Scope/Scope/Taxonomic...

Reset to last working setup helps you to make changes back, if you are not happy with them.

You can use for edit and for delete a search component.

To add a new search component, click on the Add Item button. The configuration files consist of one element – the field element, and several attributes. The element represent each lucene field and its attribute are used to configure indexing, searching, and display. Take mouse over each question mark near by a field and you can see the help information.

Edit

New Search Attribute

General

id 14

Display Name ?

Source Name ?

Metadata Node + ?

Header Item ?

Default Header Item ?

Type

Search Component Type ?

Data Type ?

Index Parameters

Store ?

Multi Value ?

save

We, therefore, go through each of the elements attributes of the configuration file.

Display Name: This is the name which is displayed in the search UI for the field.

Source Name: This is the name of the field in the lucene index.

Metadata Node: Add one or more xpaths from the metadata elements to be mapped against the lucene field.

Search Component Type: This specifies the search pattern that will be should be used against this field (as discussed in the introduction). Therefore, the value of this attribute can be any of “Category” for creating a category-based search field, “Facet” for creating faceted search field, “Property” for creating a property search field, and “General” which creates an indexed field which is not displayed in the UI, but however, searched.

Data Type: This specifies the primitive data type of the value to be indexed. E.g. string, integer, double, date.

Store: This specifies if the field value should be stored. If the field is not stored, you can only search against the terms in the field, however, you cannot retrieve the value. For minimal display of the search result, it is recommended that some fields be store. This value of the store attribute can be either yes or no.

Multi Value: This specifies if there are several values of a metadata element in a given field. E.g. if a dataset can have several owners in a metadata, then, the owners field in the index in this case, will be multi-valued

Analyzed: This specifies if the field should be analyzed or not. Only analyzed field can be searched.

Norm: This can have a value of yes or no and it is used to specify if a norm should be created for the field. Norms can be used for similarities search between documents. They can also significantly increase index sizes. So, you must take care to of what field should contain norms

Boost: This specifies the importance or weight of a field relative to others in a search. E.g. you may want the terms in “title” of a document to carry more weight than the content of the “footnotes” while indexing and searching.