

BExIS 2.5.1 Release Notes 27.11.2014

Due to a bug in handling datasets with EML metadata structures viewing and downloading of these datasets was not working properly. So this is primarily a bug fix release, although it contains two small improvements in the data planning module as well. When creating new data structures users are now able to add similar data attributes multiple times with a counter. Also, copying entire data structures is now possible.

With version 2.5.0 of the BExIS data management application the security system has been altered and extended following a number of user requests. Most importantly, the system now provides single sign-on authentication using the Lightweight Directory Access Protocol (LDAP). So users are now able to register with the system using existing credentials from their home organization, for example.

The security system is now based on the more common User/Group model instead of the User/Role model. This is primarily a name change. Upon another user request, permissions on dataset can now be specified more fine-grained (Create, View, Update, Delete, Download, Grant).

Another new feature of this release is that BExIS 2 now provides an export function for metadata. With this feature metadata structures are transformed from an internal representation into XML files conforming either to the EML or ABCD standard. This function is primarily designed for data managers and administrators who want to make selected datasets available for data discovery by external portals. This feature was developed through the German Federation for the Curation of Biological Data (GFBio).

In order to reduce the effort of setting up a new BExIS 2 instance the application is now shipped with a basic set of units and data types taken from the National Center for Biomedical Ontology (source: <http://purl.bioontology.org/ontology/UNITSONT>).

Besides these new features, extensive testing has been conducted on all components, resulting once more in number of bugs and improvements being identified and fixed.

I. Software Information

- Name: BExIS
- Version: 2.5.1
- Application Type: Web Application
- Platform: Windows

II. License Agreement

The contents included in this software are licensed to you, the end user. Your use of BExIS software is subject to the terms of an End User License Agreement (EULA) accompanying the software and located in the \License subdirectory. You must read and accept the terms of the EULA before you

access or use BExIS software. If you do not agree to the terms of the EULA, you are not authorized to use BExIS software.

Note: From version 2.5.0 onwards BExIS is released under the LGPLv3 Licence.

III. System Requirements

- Operating System: Windows Server 2003 or 2008. For personal or small installations the software is able to run on Windows 7 too.
- Application Server: IIS 7.0 +. For IIS settings see the installation manual
- DBMS:
 - IBM DB2 Express C 10 (version 10.1.2) 32 or 64 bits. BExIS has not been tested on other versions!
 - PostgreSQL: (version 9.3) 32 or 64 bits. BExIS has not been tested on other versions!
- Dependencies
 - .NET Framework 4.0, created by Microsoft
 - ASP.NET MVC 3.0, created by Microsoft
 - Lucene.Net 3.0.3, created by Lucene.Net Community
 - Lucene.Net Contrib 3.0.3, created by Lucene.Net Community
 - SharpZipLib 0.86.0, created by ICSharpCode
 - NHibernate 3.3.3.4001 created by NHibernate Community
 - Telerik MVC Extensions 2012.3.1018, created by Telerik
 - Unity Application Block 2.1.505.2, created by Microsoft
- Disk Space:
 - The Software: 20-30 MB
 - The Workspace: 10 MB and more, based on the number of modules and the amount of data

IV. Installation

To install the software, please follow the instructions in the [Install_Manual_BPP251.pdf](#) provided with the installation package.

V. System Functionality

The software consists of two main types of components; 1) core components that include Data Lifecycle Management (DLM), Web Interface, Security and some other internally used ones. 2) A collection of coherent features and use-cases bundled as modules so that each module is designed around one or a small set of steps of the DLM workflow i.e., Data Submission, Quality Control/ Assurance, Data Publication and so on.

a) Data Discovery Module

The main purpose of this module is to enable users to search for data available within the system. The module executes user defined queries against the metadata and primary data and provides advanced features such as facets, keywords, categories, filters, grouping, and sorting. Depending on individual user permissions the system provides access to full metadata records, the underlying data structure and the primary data.

Existing Features

- **Discover data using metadata dimensions (aspects, facets) and attributes** (version 2.0.0)
- **Configure search module through a user interface** (version 2.1.0)
- **Primary data search** (version 2.4.0)

New Features

- NA

Changed/ Enhanced Features

- In the Search Component Manager multiple Metadata Nodes can now be mapped to one Lucene field, thus supporting queries across multiple metadata structures

Dropped Features

- NA

Fixed Issues

- Updating of facets now works at the first Refresh Search

Breaking Changes

- NA

Known Issues

- Facets do not act as expected when names contain special characters (e.g. “-“)

b) Data Collection Module

The purpose of the data collection module is to allow users to submit data alongside with metadata. The module processes the data in accordance with its data/ metadata structure rules to find any inconsistencies and communicates with the user to overcome the issues. The module may also notify or trigger other internal services such as the search indexing engine to make the dataset ready for discovery. Based on the settings of the associated research context/ plan, datasets get versioned.

Existing Features

- **Uploading an Excel file** (version 2.1.0)
- **Uploading a delimited file (CSV, TSV),** (version 2.2.0)
- **Adding data to an existing dataset** (version 2.2.0)
- **Create Structured Dataset** (version 2.3.0)
- **Create Unstructured Dataset** (version 2.4.0)
- **Push big files to server** (version 2.4.0)

New Features

- NA

Changed/ Enhanced Features

- Added an EML Metadata Structure
- Extended the ABCD Metadata Structure

Dropped Features

- Removed a project specific Metadata Structure (Biodiversity Exploratories)

Fixed Issues

- NA

Breaking Changes

- NA

Known Issues

- Empty rows within or at the end of Excel or CSV files lead to validation errors and should be removed before adding (i.e. uploading) data to a dataset

c) Security Module

Authentication and authorization of users against the application and modules' functionalities as well as data are main duties of the security module. Defining and managing all security related entities such as users, roles, and permissions are also part of this module.

Existing Features

- **User self-registration** (version 2.1.0)
- **Users and roles management for an administrator** (version 2.1.0)
- **Feature security** (version 2.2.0)
- **Dataset Security** (version 2.4.0)
-

New Features

- **Single Sign-On (LDAP)**: Users can register with the system using existing credentials from an external authentication provider (LDAP Service)

Changed/ Enhanced Features

- Switched from a User/Role model to a **User/Group model**
- **Dataset Security**: changed the available permission types to: Create, View, Update, Delete, Download, Grant
- Implemented **Dataset Security for Uploading Data**
- Implemented **Dataset Security for Downloading Data**
- Implemented **Dataset Security for Metadata Export**
- Renamed items in **Authorization Menu**

Dropped Features

- Spaces in user names and passwords are not allowed anymore (version 2.4.0)

Fixed Issues

- NA

Breaking Changes

- NA

Known Issues

- NA

d) Data Planning Module

The purpose of the data planning module is to manage (create, edit, delete) data and metadata structures, data attributes (variables), units, data types and their relations. The aim is to foster reuse and data integration while retaining flexibility.

A conceptual model showing the connection and relationship between the different terms and entities is available at: <http://fusion.cs.uni-jena.de/bppCM/index.htm>

Note: In previous releases this module was called Research Planning with a much broader scope. The new name now better reflects the actual purpose and the functionality available.

Existing Features

- **Data structure manager** (version 2.1.0)
- **Data attribute manager** (version 2.1.0)
- **Data type manager** (version 2.1.0)
- **Unit manager** (version 2.1.0)
- **Unstructured data structures** (version 2.4.0)

New Features

- Copying entire data structures with the new “save as” button

Changed/ Enhanced Features

- In the Data Structure Manager variables in data structures can be reordered now (version 2.4.0)
- Minor improvements in all related user interfaces (e.g. filtering added, messaging)(version 2.4.0)
- When creating a data structure a data attribute can now be added multiple times base on a counter, Data structure manager (version 2.1.0)

Dropped Features

- NA

Fixed Issues

- NA

Breaking Changes

- NA

Known Issues

- NA

e) Data Dissemination Module

The main purpose of this module is to provide users with access to primary data and metadata managed by the system. In the future this module shall also provide support for publishing, orchestrating and harvesting data of the system by external resources.

Existing Features

- **Primary data download as Microsoft Excel file** (version 2.2.0)
- **Primary data download as comma or tab separated file (CSV)** (version 2.4.0)
- **Download of unstructured data (single file, multiple files in ZIP)** (version 2.4.0)
- **Customized Download of primary data of structured data (WYSIWYG)** (version 2.4.0)

New Features

- **Export of metadata** from an internal representation into XML files conforming either to the EML or ABCD standard. This function is primarily designed for data managers and administrators who want to make selected datasets available for data discovery by external portals. This feature was developed through the German Federation for the Curation of Biological Data (GFBio).

Changed/ Enhanced Features

- NA

Dropped Features

- NA

Fixed Issues

- Downloading datasets with EML metadata structures did not work
- Viewing unstructured datasets with EML metadata structures did not work

Breaking Changes

- NA

Known Issues

- NA

VI. Contact

Help desk:

- Website: <http://fusion.cs.uni-jena.de/bexis>
- Email address: bexis-support@uni-jena.de
- Telephone: +49-(0)3641-948968
- Fax: +49-(0)3641-946302