BEFdata User Guide

V1.2

Date: 20.08.2012

Authors: Karin Nadrowski and Sophia Ratcliffe

Special Botany and Functional Biodiversity
Universität Leipzig
Institut für Biologie I
Johannisallee 21-23
04103 Leipzig
Germany

Phone: 0049 (0)341 9738576

UNIVERSITÄT LEIPZIG
AG Spezielle Botanik und funktionelle Biodiversität

Contents

1. The BEFdata portal

The BEFdata portal is an online data management tool that allows members of cooperative research projects to upload, store and validate their data in a secure environment. The portal was developed within the Biodiversity-Ecosystem Functioning (BEF-China http://www.bef-china.de/) research unit of the German Science Foundation (FOR 891).

Import of data into the portal is via a formatted MS Excel 2003 Workbook, hereafter called the workbook; the only way to upload data into the portal is by completing the workbook. A template of the workbook is available for download from the portal. Data export is organized using paper proposals, in which members can collect datasets into a cart and then issue a request to receive download rights for these datasets.

For help with the BEFdata portal, support details can be found under Imprint in the footer of the website.

This document is a guide to the website and shows how to perform common actions such as filling out the workbook with research data, uploading it to the portal, and managing categories and datagroups within the portal.

Separate instances of the BEFdata portal can have different themes or styles applied to the website, to allow projects to apply their specific style (such as headers and footers) to their website. In the following we use the FunDivEUROPE instance to describe the use of the BEFdata portal. To aid with understanding, the glossary in section 7 lists some of the common terms used in the portal and their definitions.

2. The website

2.1 Home page

Figure 1 illustrates the FunDivEUROPE Befdata portal home page. Members can log in through the boxes highlighted in the illustration. Unauthenticated users (those people who have not logged on) can navigate the portal but are not able to download any data, unless the dataset owner has allowed this level of access. They also cannot contribute data to the portal or request data via paper proposals.

Figure 1: Home page of the data portal

The data portal navigation is available on all pages across the top of the page.

2.2 Projects

The projects page lists all the sub-projects in the portal (see Figure 2). Each FunDivEUROPE task has its own project listed here, where members of that particular task can upload their data to.

Figure 2: The Projects index page of the data portal

Clicking on the project name takes you to a page that lists all the datasets for the project and any paper proposals that are linked to datasets within the project (see Figure 3). The box in the right-hand column lists all the members of the project.

Figure 3: A Project page in the data portal

2.3 Staff

The staff index page lists all members of the portal, including a portrait picture if one is available. Clicking on the name takes you to their home page, which lists those datasets and paper proposals that they are involved in, and in the right-hand column, any projects that they belong to.

2.4 Data

The data index page lists all the datasets that have been uploaded to the portal. Keywords from already uploaded datasets are displayed as a tag cloud in the right-hand column. Clicking on a keyword displays links to all the datasets tagged by the keyword. Logged on users are also able to upload datasets from this page by clicking on the **'Create new dataset'** link. The upload process is described in Section 4.

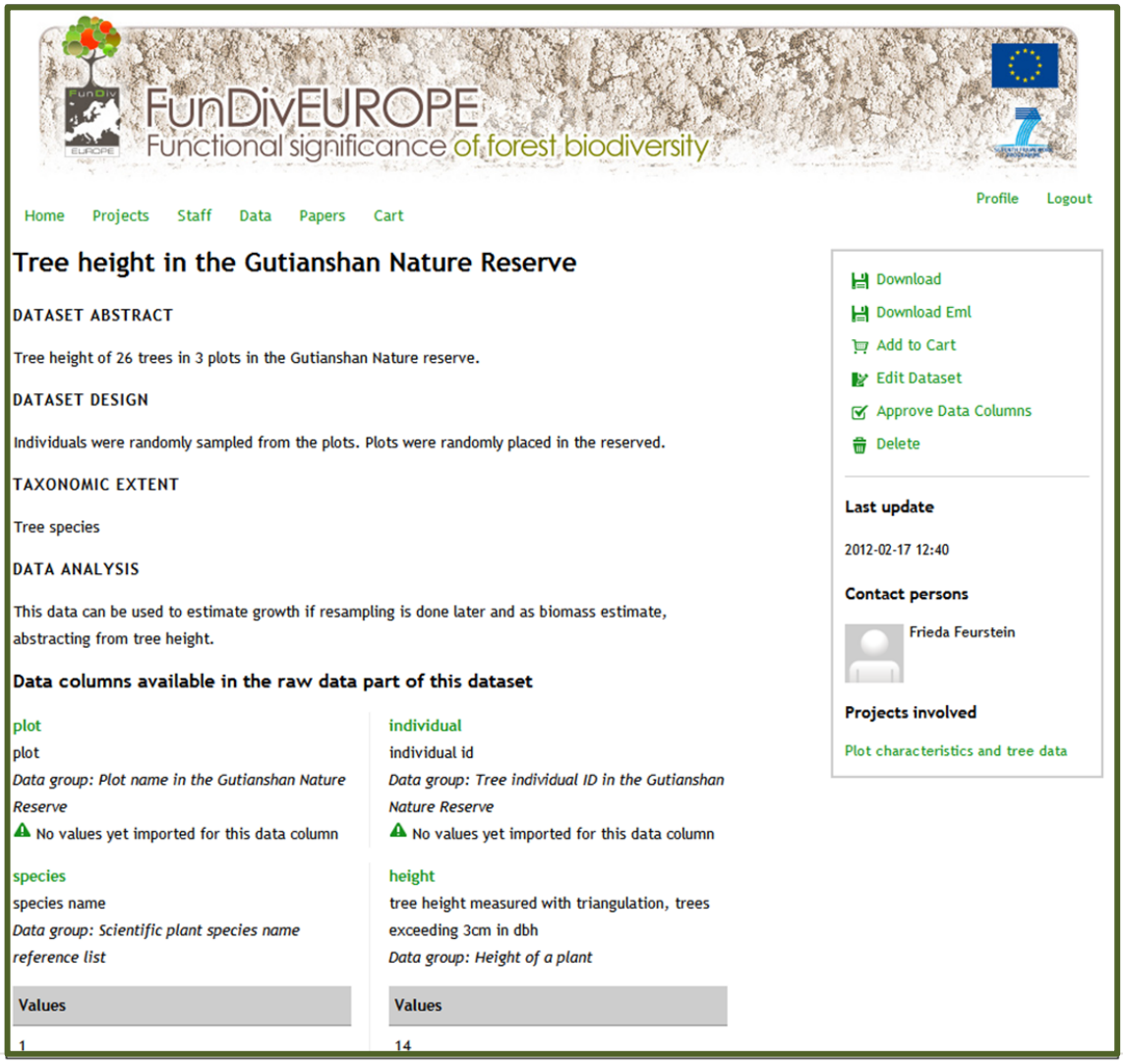
Clicking on a dataset name opens the overview page of the dataset. This page displays the metadata, datacolumn details and actions that the current user is able to perform on the dataset. All of this information comes from the original uploaded workbook which is described in Section 3.

Metadata is the information required by people, not familiar with the data, to be able to understand and utilise the dataset in their research. Underneath the metadata on the page is a box of information for each datacolumn in the dataset. The title, description and

datagroup of the datacolumn are presented, including the first five unique values from the column.

Figure 4 illustrates the dataset overview page as an owner of the dataset.

Figure 4: A dataset overview page in the data portal



The available options in the right-hand panel are different depending on the users' access rights.

Unauthenticated (not logged in users) are only able to see the date the dataset was updated and the list of members to contact about the dataset.

Members of the site who do not have download rights to the dataset and are not owners are able to:

1. Download the dataset metadata in EML format
2. Add the dataset to their Paper Proposal cart

Members of the site who have download rights to the dataset but are not the dataset owner are additionally able to:

3. Download the dataset

In addition to the items above, dataset owners are able to:

4. View the comment field
5. Edit the dataset, including re-uploading the raw data

The comment field is only available to owners and those with the ‘project board’ role. The comment also will appear in any short view of datasets in lists for the owners and project board members. The delete option is only available to administrators.

When a dataset is downloaded the name of the person and the time of download are stored. The list of who download the dataset and when is displayed in the right-hand column and is only available to data owners and users with the ‘data admin’ role.

Regenerate dataset download

Datasets are downloaded in the formatted Workbook and contain all the changes applied to dataset during the upload and validation process. For performance reasons the download file for each dataset are generated after uploading and editing, this means that the user doesn’t have to wait for the dataset to be recompiled at each download. It is possible to regenerate a download file if the most recent download file is not considered up to date.

2.5 Paper proposals

Paper proposals are data requests accompanied by formulated research ideas that specify what data is needed and whose expertise should be consulted. Members who would like to use datasets for joint publication must submit a paper proposal that will be reviewed by the project board to make sure that it is novel and doesn't compete with other activities. In a second step, it is then reviewed by all dataset owners listed in a paper proposal. Datasets can be added to a logged-in user's cart and this collection of datasets is then used as part of a paper proposal. The proponents can only gain access to the datasets once all owners have approved the proposal. At the same time, a preliminary author list is collected based on the proponents and dataset owners. Dataset owners gain a 'co-author' role in the paper proposal (not necessarily in the resulting paper), whilst members mentioned in the Acknowledgement sheet (worksheet 2) of the data workbook will be added to the acknowledgements section of the paper proposal.

3 The BEFdata workbook

Data, and its associated metadata, are uploaded to the portal from a formatted MS Excel 2003 spreadsheet. As long as the raw data is in a table format (where columns are the different measurements and each row is an observation) then any type of data can be uploaded. There is no size restriction, apart from those of Excel (64,000 rows), but larger datasets (those over 10,000 rows) can take some time to upload. An empty, template workbook is available for download from the site and it is recommended to download the latest workbook so that you are always using the latest version. The title of the template workbook dataset is 'How to fill out a BEFdata Workbook'.

The workbook has five sheets. The first four sheets describe the data and the fifth contains the raw data. Table 1 describes each sheet in the workbook.

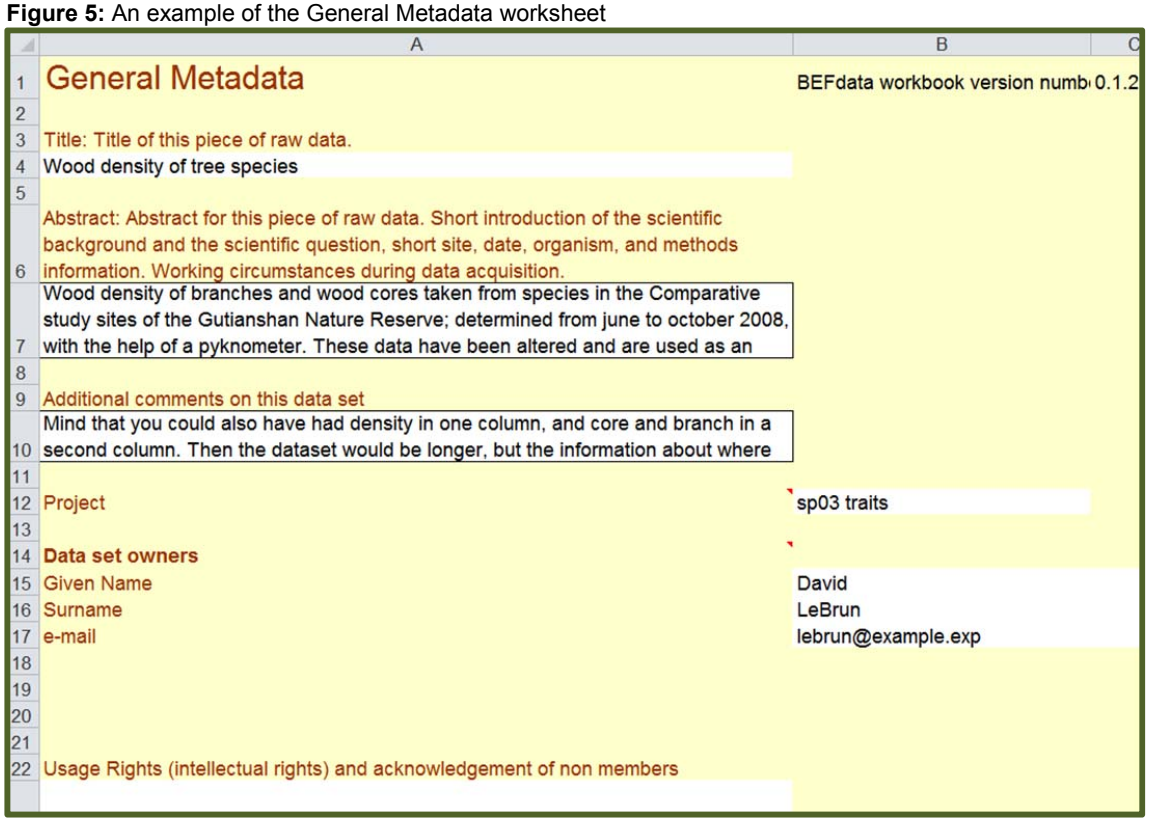
Table 1. The datasheets in the workbook

Name	Description
General metadata	Contains fields for dataset title, abstract, location, project and data ownership.
Acknowledgements	Lists users who have contributed to data capture or collection for a particular column, other than those listed in the 'General metadata' sheet.
Columns and datagroups	Lists all the columns in the 'Raw data' sheet. Including definitions, unit of measure and data type.
Categories	Lists the individual options available for category datatypes used in the 'Raw data' sheet.
Raw data	Contains the actual data, each row as a single observation.

3.1 Metadata – sheet one

The general metadata sheet captures the title, abstract, project and any additional information required to understand and interpret the data correctly. Data ownership can also be recorded in this sheet, however during upload the current user is added as the data owner. Only the dataset title is required; all other fields are optional. The title must be unique within the portal otherwise an error will be raised. See Figure 5.

Figure 5: An example of the General Metadata worksheet



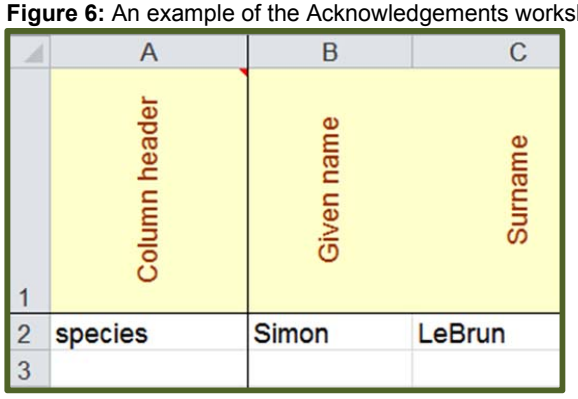
	A	B	C
1	General Metadata	BEFdata workbook version numb	0.1.2
2			
3	Title: Title of this piece of raw data.		
4	Wood density of tree species		
5			
6	Abstract: Abstract for this piece of raw data. Short introduction of the scientific background and the scientific question, short site, date, organism, and methods information. Working circumstances during data acquisition.		
7	Wood density of branches and wood cores taken from species in the Comparative study sites of the Gutianshan Nature Reserve; determined from june to october 2008, with the help of a pyknometer. These data have been altered and are used as an		
8			
9	Additional comments on this data set		
10	Mind that you could also have had density in one column, and core and branch in a second column. Then the dataset would be longer, but the information about where		
11			
12	Project	sp03 traits	
13			
14	**Data set owners**		
15	Given Name	David	
16	Surname	LeBrun	
17	e-mail	lebrun@example.exp	
18			
19			
20			
21			
22	Usage Rights (intellectual rights) and acknowledgement of non members		

The more information you are able to add to the metadata sheet the easier it will be for project members to understand and use your data.

3.2 Acknowledgements – sheet two

The acknowledgements sheet allows project members, who were responsible for the collection of data in a single datacolumn to be referenced, so that data ownership is not lost. Figure 6 illustrates an acknowledgement sheet where the data in the species column is linked to a name.

Figure 6: An example of the Acknowledgements worksheet



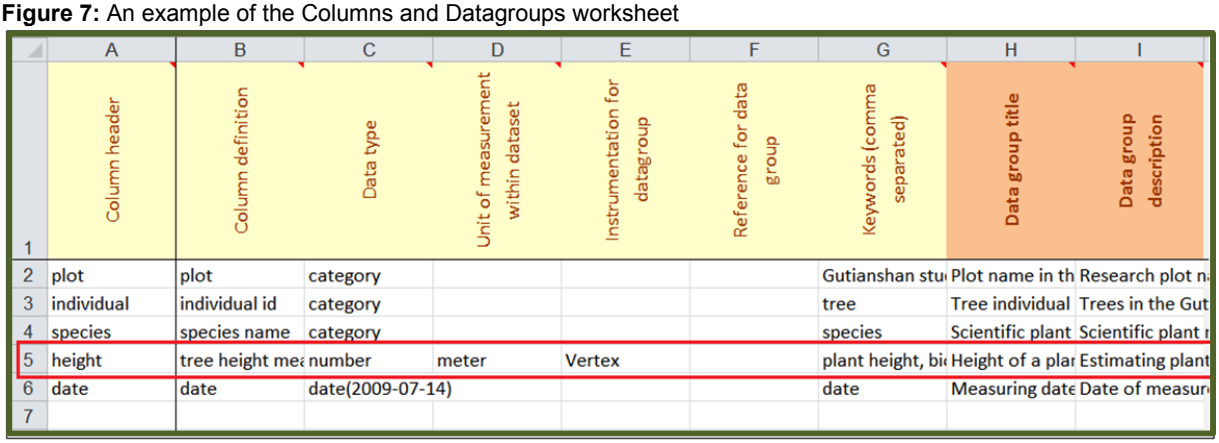
	A	B	C
1	Column header	Given name	Surname
2	species	Simon	LeBrun
3			

To reference people simply add the column title and their first and last name in a single row.

3.3 Datacolumns and datagroups – sheet three

The datacolumns and datagroups sheet describes the columns in the data. Each datacolumn in the raw data sheet is represented by a row in the sheet. An example of the third sheet is shown in Figure 7. The sheet is separated into two parts: one that is specific to the dataset (from 'Column header' to 'Keyword'); and a part that can be reused within and between datasets ('Data group', 'Data group description'). For example (see the highlight in Figure 7), the DBH of a tree can be measured in centimetres or in millimetres. In this case 'Tree size' is the data group and can be reused across datasets, whilst the dataset specific part is that tree height was measured in centimeters (stored in the 'unit' column).

Figure 7: An example of the Columns and Datagroups worksheet



	A	B	C	D	E	F	G	H	I
1	Column header	Column definition	Data type	Unit of measurement within dataset	Instrumentation for datagroup	Reference for data group	Keywords (comma separated)	Data group title	Data group description
2	plot	plot	category				Gutianshan stu	Plot name in th	Research plot n
3	individual	individual id	category				tree	Tree individual	Trees in the Gut
4	species	species name	category				species	Scientific plant	Scientific plant
5	height	tree height mea	number	meter	Vertex		plant height, bi	Height of a plar	Estimating plant
6	date	date	date(2009-07-14)				date	Measuring date	Date of measur
7									

There are 5 different datatypes that can be used in the portal and the choice of datatype is important because the data will be validated against specific rules for each datatype. If no datatype is provided, or an invalid datatype is used, then the user will be asked to select one from one of the 5 built-in datatypes during the upload process (see Section 4).

Table 2: Description of the five datatypes

Datatype	Validation rules
Text	No validation.
Date	Validated as a date. Once Excel recognizes the value as a date it will be handed over to the portal in the correct format. In this case you can choose 'date(2009-07-14)' as datatype.
Number	Validated as a numeric. Negative numbers and decimal places are allowed. If there are numerical categories predefined for the datagroup, they will be handled as categories (e.g. -9999 for missing altitude).
Year	Valitated as year (4 digit number)
Category	Validated against any categories already present in the portal for the given datagroup, and then against the list of categories included in the Categories sheet of the Workbook.

The datagroup and datagroup description are useful for describing or classifying columns of similar data that are used across several datasets. Consistent use of categories and datagroups within and between datasets is essential when searching for, and merging, datasets. The use of naming conventions is described in more detail in Section 5. To help select a suitable datagroup the list of datagroups already used in the portal can be found here; http://fundiv.befdata.biow.uni-leipzig.de/datagroups.

Keywords included for a column are added to a ‘tag cloud’ of keywords on the Data page. Users can then use the tag cloud to find datasets.

3.4 Categories – sheet four

The categories sheet details any naming conventions used in the data, for example species codes instead of their full scientific name. Categories are the individual values of any category datatypes. If the category has already been uploaded in a previous workbook it is not required to list them again, assuming that they are identical. On validation the categories are matched to existing categories from the same data group and if one is not found the user will be prompted to create a new category during the upload process. We recommend that categories are added to the sheet as it speeds up the upload and validation process, however this is not necessary. See Section 4 (Datatype approval) for more details.

Figure 8 shows an example of a completed category sheet.

Figure 8: An example of the Categories worksheet

	A	B	C	D
1	Column header	Category short	Category long	Category description
408	id	379838	379838	379838
409	id	379867	379867	379867
410	id	R02-025	R02-025	R02-025 Böhnke Wood
411	species	Acer amplum	Acer amplum	Acer amplum
412	species	Acer cordatum	Acer cordatum	Acer cordatum
413	species	Acer pubipalmat	Acer pubipalmatur	Acer pubipalmatum
414	species	Adinandra millett	Adinandra millettii	Adinandra millettii
415	species	Alangium kurzii	Alangium kurzii	Alangium kurzii

Each category should be added as a row in the spreadsheet with the column name in which the category value appears in the first column. The category value goes in the second column (Category short) and the longer category name and description are in the third and fourth columns, respectively. A long category name is required when the

purpose of the category is unclear from the short name. If no category long or category descriptions are provided then the short name is used.

3.5 Raw data – sheet five

The Raw data sheet contains the data, one row for each observation. There is no limit to the number of rows, other than MS Excel's own limit; however, over 10,000 rows can take an hour to upload. Figure 9 gives an example of a Raw data sheet.

Figure 9: An example of the Raw Data worksheet

	A	B	C	D	E
1	plot	individual	species	height	date
2	1	1	1	12	12-Jan-12
3	1	1	2	34	13-Jan-12
4	1	2	1	23	14-Jan-12
5	1	2	2	27	15-Jan-12
6	1	3	1	23	16-Jan-12
7	2	1	2	43	17-Jan-12
8	2	1	1	66	18-Jan-12
9	2	2	2	43	19-Jan-12
10	2	2	1	123	20-Jan-12
11	2	3	2	72	21-Jan-12
12					

It is essential that the column headers in the raw data sheet match those used in the acknowledgements, columns and datagroups, and categories datasheets otherwise the import process can't match the column descriptions to the data. The column headers must also be unique; the upload process will return an error if this is not the case.

4 Data upload and approval

Once the data has been added to the MS Excel 2003 Workbook and described correctly, the workbook can be uploaded to the portal.

The process of uploading the data has two steps: the first step is to save the metadata (from the first sheet in the workbook) and the initial upload of data; and the second is the

process of going through each column of the uploaded data and verifying the datatype and data group, and adding any additional column descriptors. If the workbook has been described correctly then it is possible to approve all the columns in one step.

The following are required for the upload process to run with no problems:

- Use the current version of the workbook by downloading the dataset 'How to fill out the BEFdata Workbook'
- Column headers in the raw data sheet are **unique**;
- Column headers from the raw data sheet match those in the other sheets of the workbook;
- Correct datatype used in the column and datagroup sheet (**text**, **year**, **date(2009-07-14)**, **number** or **category**);
- A datagroup and datagroup description are included for each column in the column and datagroup sheet. They do not have to exist in the portal, however, when using category datatypes an existing datagroup must be chosen;
- When category datatypes are used, their individual values are stored in the category sheet, unless you know that the categories already exist in the portal for the given datagroup.

4.1 Upload – selecting a workbook

To upload data, click on the **'Create new dataset'** link on the **Data** page, as shown in Figure 10.

Figure 10: Location of the 'Create new dataset' link

On the next page use the file upload button to select the workbook from the file system, and click on the '**Create new dataset'** button.

4.2 Upload – saving the metadata

Figure 11 illustrates the top section of the general metadata page which is the next page in the workflow. The information is taken from the General metadata sheet of the workbook and can be amended on this page. The additional metadata fields and the save button are not visible in Figure 11. There is an additional section that the user can complete that is not available in the General metadata worksheet, for user rights.

There are four checkboxes that determine the level of visibility of the dataset to users of the portal. Table 3 describes each option.

Table 3: Dataset access level options

Access level	Description
Visible for public	This is checked by default. The dataset is visible to unauthenticated users. They are not able to download the data.
Free for public	Unauthenticated users are able to download the data.
Free for members	Members of the BEFdata portal (logged in users) are able to download the data.
Free within projects	Members who belong to the project in which the dataset resides are able to download the data.

Figure 11: The top section of a metadata page

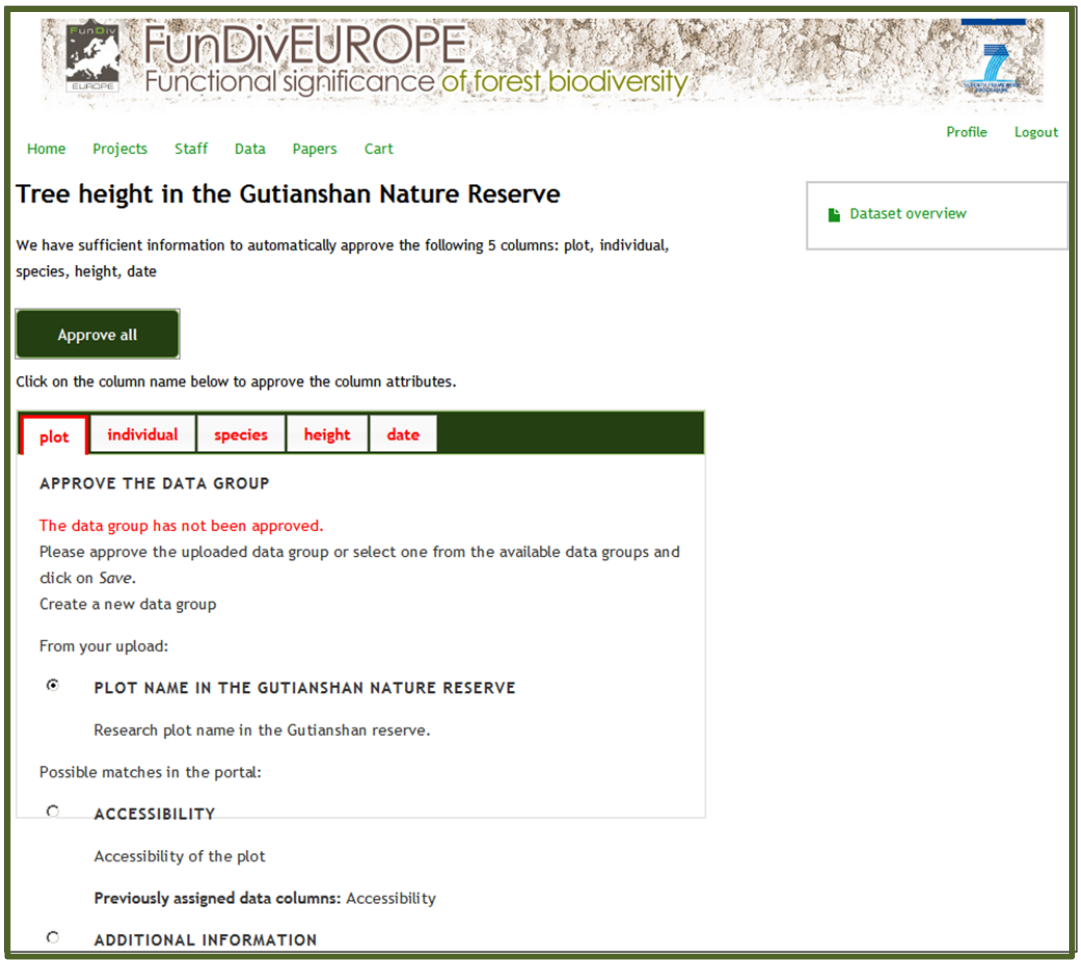
To save the metadata click on the **'Save metadata'** button at the bottom of the page.

The upload progress is displayed in the right-hand column. A background process uploads the data and it is possible to navigate away from the page during the upload. Once the data has uploaded the status of the dataset uploaded will be 'finished' and at this point all the data is in an unvalidated form. For large datasets (>10,000) this can take take upto an hour.

4.3 Validation and approval

Once the data has been uploaded in an unvalidated form it is necessary to approve and validate the data. Click on **'Approve data columns'** and the dataset overview page will be displayed, as shown in Figure 12.

Figure 12: Dataset overview page after initial upload



The red text on the button tabs indicates that the column is unapproved. Orange text indicates partially approved (the datagroup has been approved but not the datatype), and green indicates that both the datagroup and the datatype have been approved.

In the case shown (Figure 12), all the data columns were correctly described in the workbook and so it is possible to automatically approve them in one go, rather than clicking through each of the columns. To automatically approve all of the columns click on the **'Approve all'** button.

Columns where both the datagroup and the datatype are approved, and that have no invalid values, will be green, those where the portal was able to approve the datagroup by not the datatype will be orange, and those where the portal not able to approve the datagroup will remain red. It is necessary to go through each of the red and orange columns and manually approve the datagroup and datatype.

Datagroup approval

Clicking on a 'red' column will display the form to approve a datagroup. This is also the default view, as illustrated in Figure 12.

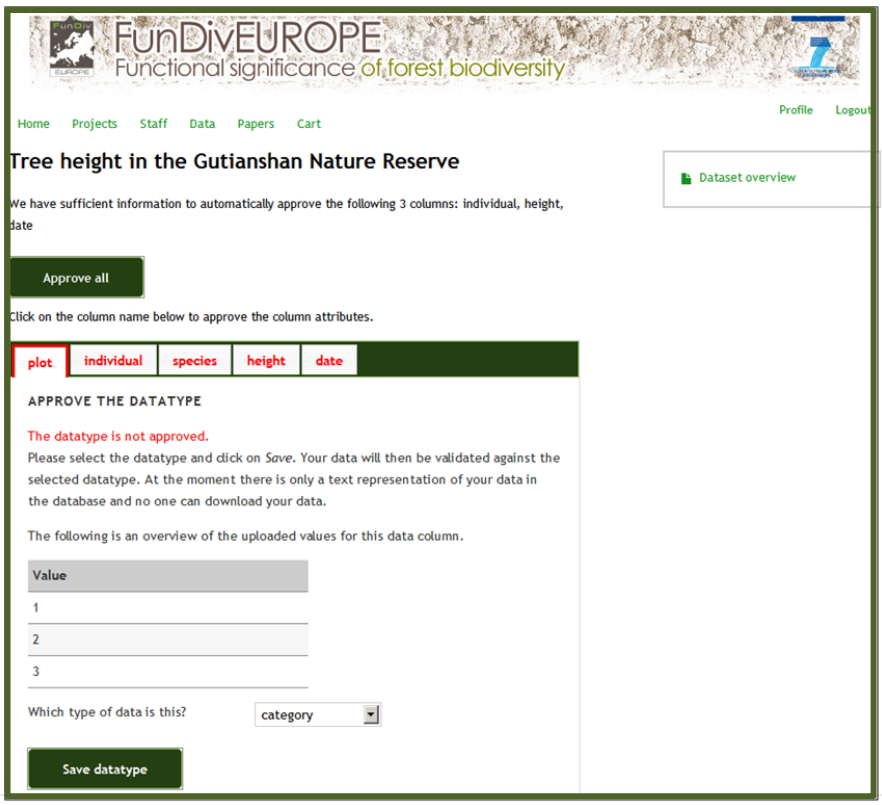
The approval page lists all of the available datagroups in the portal, plus the datagroup entered for the current column in the workbook. The datagroup from the workbook is the first in the list and is selected by default. It is possible to select another datagroup if a mistake was made in the workbook. Clicking on the **'Save'** button (not shown in Figure 12) will approve the datagroup.

It is also possible to add a new datagroup: clicking on the '**Create a new data group**' link will open up a dialogue box with a form in which the datagroup details can be entered and saved.

Datatype approval

Once the datagroup has been saved the datatype approval form will be displayed, as illustrated in Figure 13.

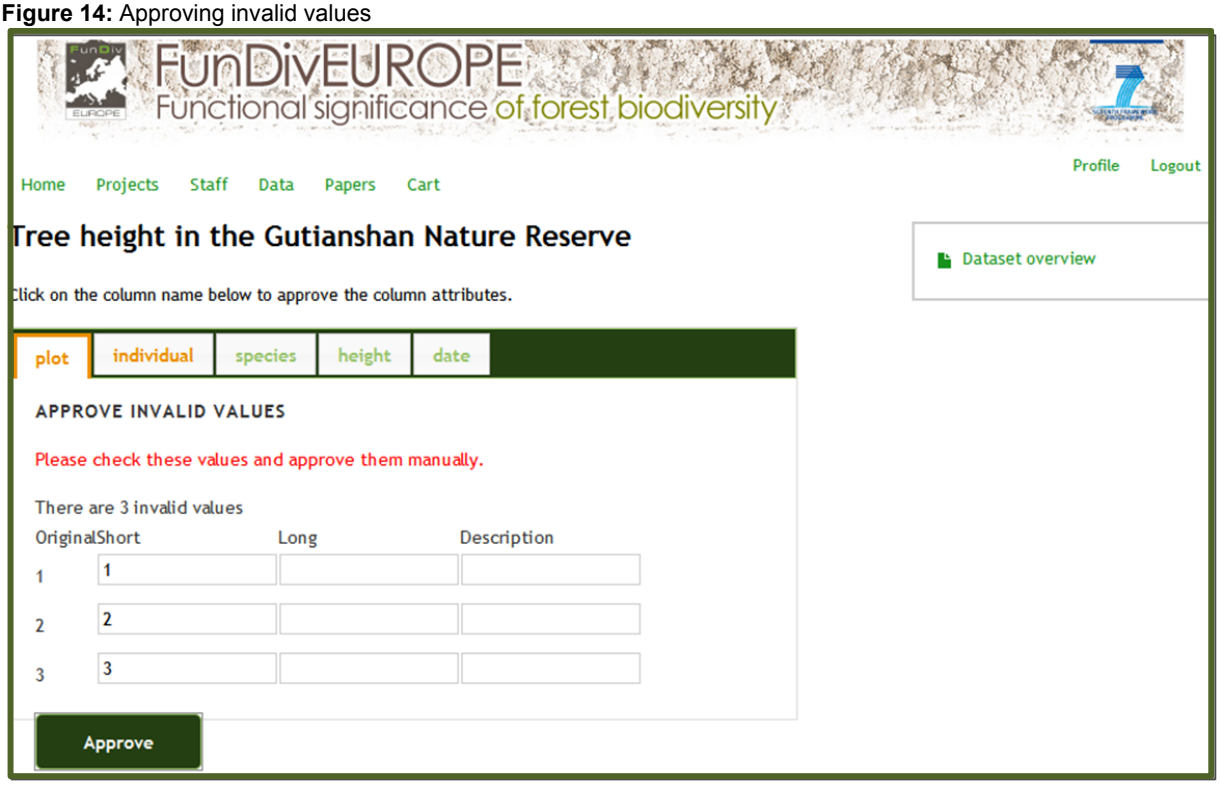
Figure 13: Datatype approval



The page displays the first 10 unique data values and a dropdown list of the five datatypes: text, date, year, number and category. Select the correct datatype and click on the '**Save datatype**' button.

The uploaded data for the column will be validated against the selected datatype. Any invalid values will be returned to the screen for the user to validate. An example of this is shown in Figure 14 using the **category** datatype where the data values did not exist in the database and they weren't included in the categories sheet of the workbook. Another example could be for the **number** datatype where <100 was entered. This is not a valid number but may be correct in the context of the data.

Figure 14: Approving invalid values

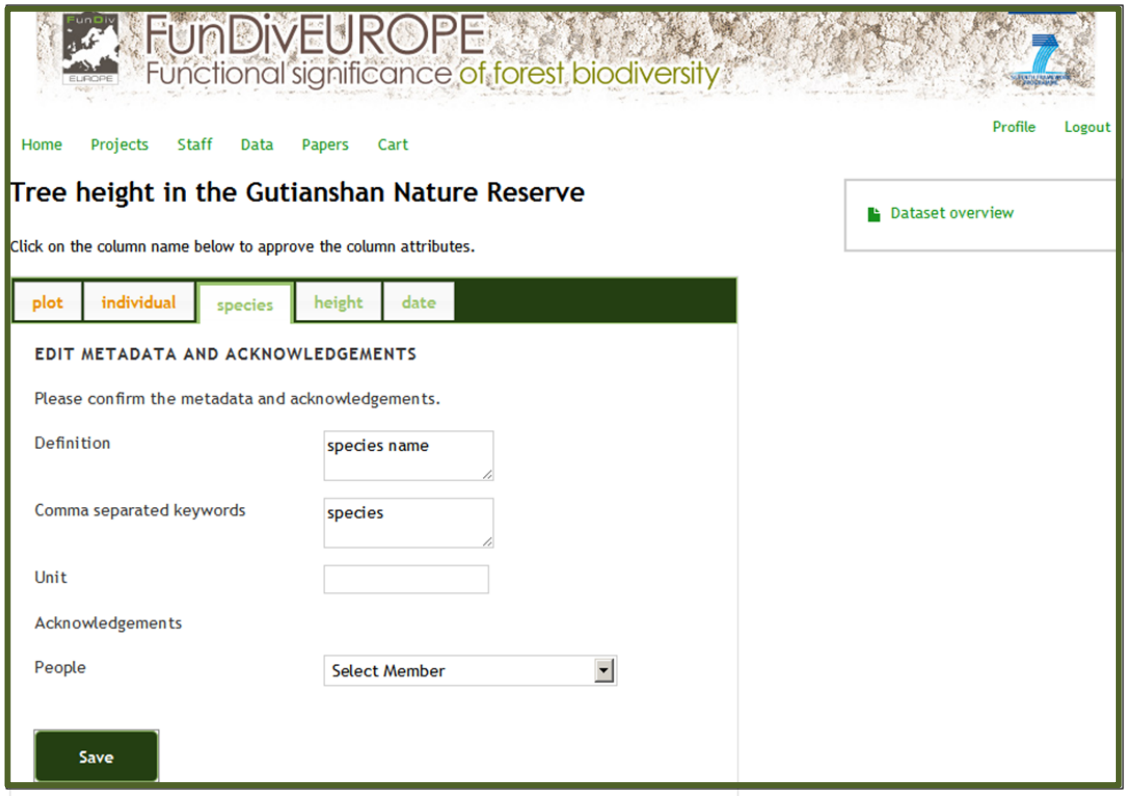


The user must then approve any invalid values and click on the '**Approve**' button. If the user doesn't provide a long name or a description for the category the short name will be used.

Column metadata and acknowledgements approval

The final page in the approval process is to confirm the metadata and acknowledgements (data provenance) for the column. Figure 15 illustrates an example of this page.

Figure 15: Saving column metadata and acknowledgements



It is possible to change the datagroup, datatype, metadata and acknowledgements of an approved column at any time.

4.4 Editing a dataset

Owners are able to edit their datasets. They can delete all the datacolumns and re-upload new values; however they are not able to completely delete the dataset. On re-uploading new data the user will have to go through the same approval process as described above.

5 Naming conventions

A central theme of the portal is naming conventions, the use of consistent names and codes for identical objects. Consistent naming standards are essential when merging and sharing data.

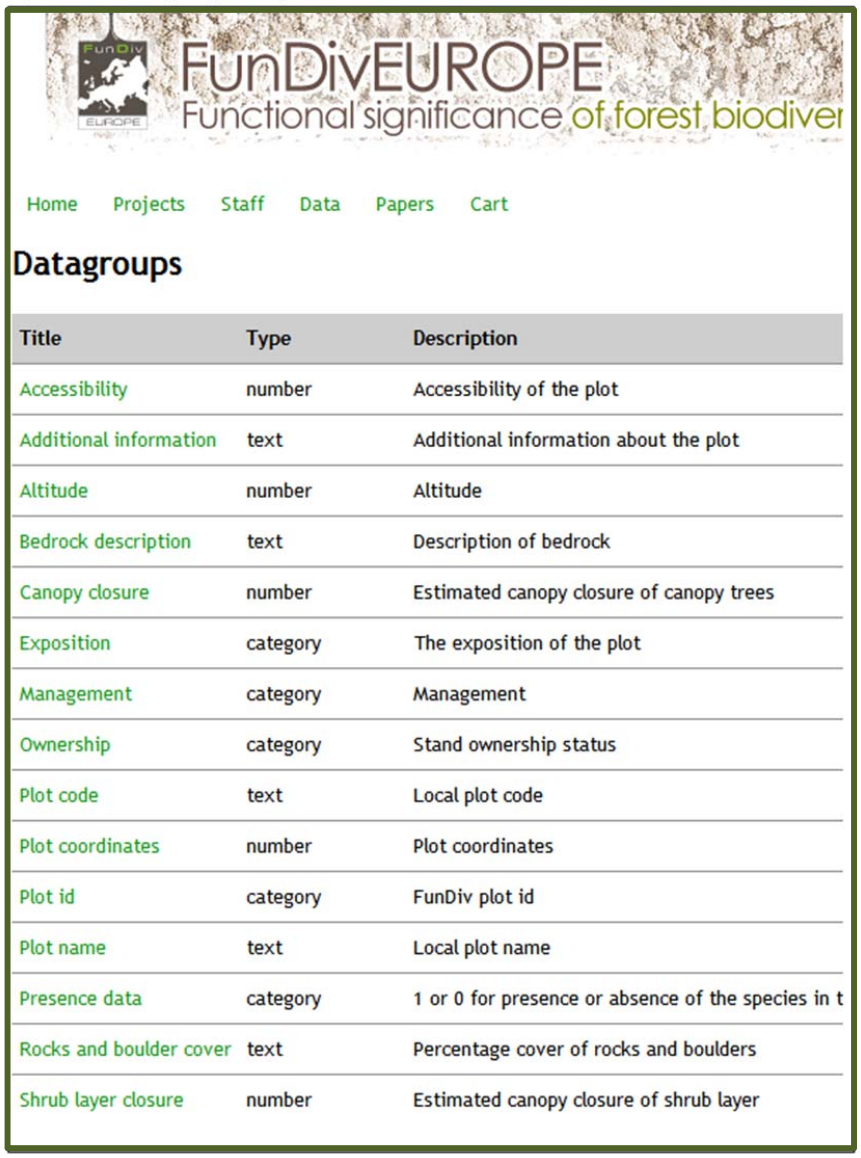
Naming conventions are managed in the BEFdata portal by datagroups and categories.

5.1 Datagroups

Datagroups are used for the semantic classification of the data. This enables similar data concepts to be grouped together within and between datasets which helps improve the efficiency of data retrieval. The list of all datagroups in the portal can be found here: http://fundiv.befdata.biow.uni-leipzig.de/datagroups, however, you must be logged in to view the page. Figure 16 illustrates the datagroups currently available on the portal. The listing includes the default datatype of the datagroup and a description. It is possible for a datagroup to have multiple datatypes (text, year, date, number, category).

As the project continues and more datasets are uploaded, the number of datagroups available will grow. Before completing the workbook it is recommended to check the datagroup list page to see if there is an existing datagroup that should be used.

Figure 16: Datagroup listing



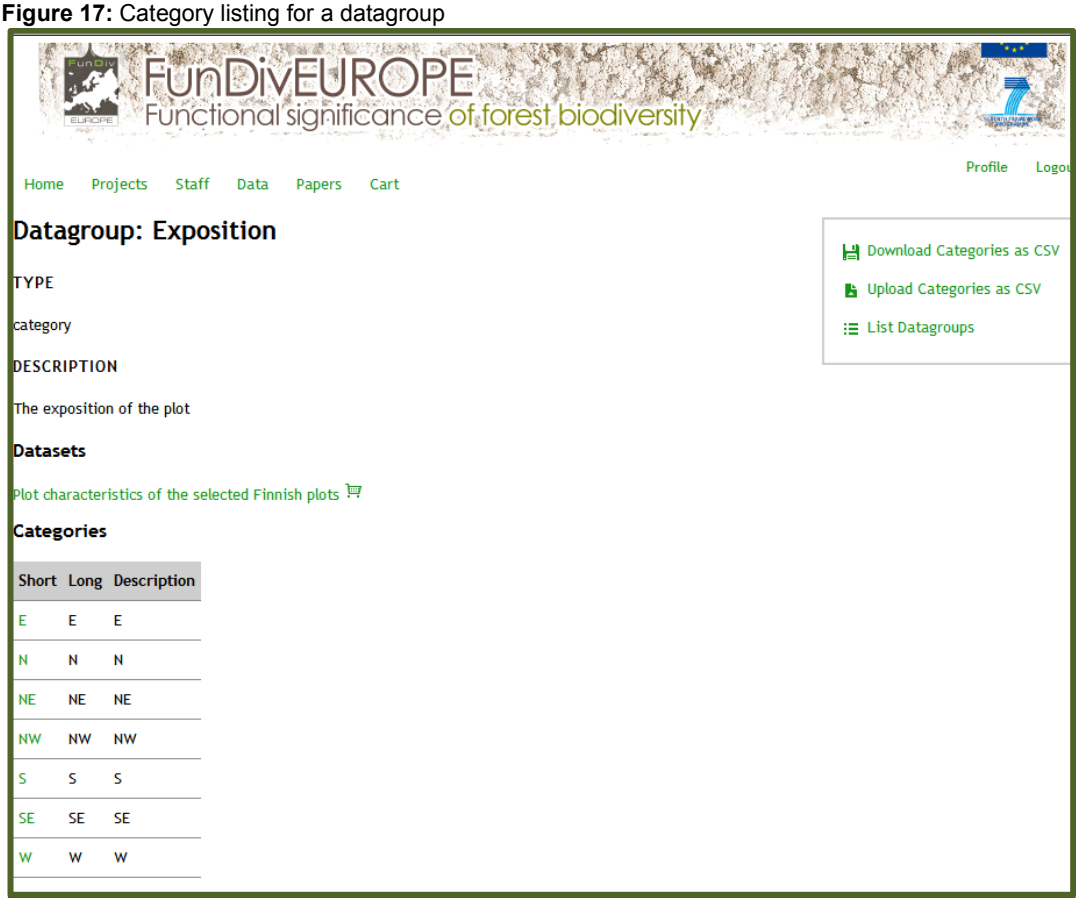
5.2 Categories

Categories enable identical objects to be referenced with the same name throughout all datasets. A good example of this is the use of codes for species names: it is essential that species have the same code throughout all the datasets in the project.

Clicking on the datagroup name displays the categories available under the datagroup. The portal ensures that category names are unique within each datagroup. Figure 17

shows an example of this page for the datagroup 'Exposition' which has 7 different category values, one for each of the points of a compass.

Figure 17: Category listing for a datagroup

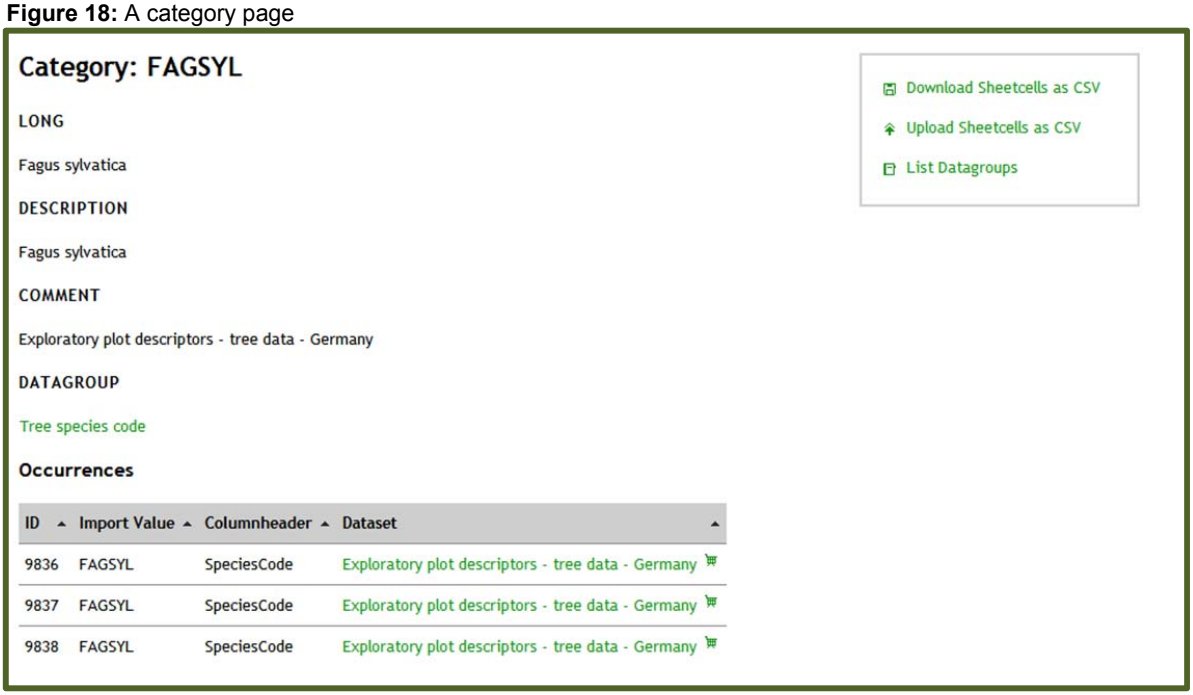


From a datagroup page it is possible to download all the categories in the datagroup to a .csv file by clicking on '**Download Categories as CSV**'.

Each category also has its own page in the data portal which is accessible by clicking on the category name. The category page lists all the sheetcells (individual raw data values) that are linked to the category.

Figure 18 shows an example of a category page.

Figure 18: A category page



From a category page it is possible to download all the sheetcells that are linked to the category by clicking on '**Download Sheetcells as CSV**'.

5.2.1 Category management

It is possible to add and merge categories on the datagroup page and split categories on the category page.

Only data owners and users with the 'data admin' role have permission to manage the categories.

Adding new categories

Add the new category names to the downloaded .csv file, ensuring that there are no duplicate names, click on **'Upload categories as CSV'**, and select the updated .csv file. The upload will throw an error if there are duplicate categories in the .csv file.

Merging categories

Categories within a datagroup can be merged using the downloaded .csv file, if, for example, two different categories actually refer to the same object. To merge categories, download the category list, add a fourth column to the spreadsheet called MERGE ID, and against the category that you would like to merge add the ID of the category that you would like to merge it into.

See Figure 19 for an example: the category with an ID of 4 will be merged into the category with ID 186. The updated .csv file can then be uploaded by clicking on **'Upload categories as CSV'.** Any data that was linked to the merged category (ID 4) will automatically be updated to point to the new category (ID 186).

Figure 19: Merging categories in the .csv file

	A	B	C	D	E
1	ID	SHORT	LONG	DESCRIPTION	MERGE ID
2	186	1	CSP01	Comparative Study Plot 01	
3	187	10	CSP02	Comparative Study Plot 02	
4	188	11	CSP03	Comparative Study Plot 03	
5	189	12	CSP04	Comparative Study Plot 04	
6	190	13	CSP05	Comparative Study Plot 05	
7	4	Plot 1	Plot 1	Plot 1	186
8	5	Plot 10	Plot 10	Plot 10	187
9	6	Plot 11	Plot 11	Plot 11	188
10	7	Plot 12	Plot 12	Plot 12	189
11	8	Plot 13	Plot 13	Plot 13	190
12					

Splitting categories

Categories can be split by downloading the sheetcells that are attached to the category by clicking on **'Download Sheetcells as CSV**', on the category page. In the final column of the downloaded spreadsheet add the name of the category that the sheetcell should be linked to. It can either be a new category name or one that already exists in the

datagroup. The .csv file must then be uploaded to the portal by clicking on **'Upload Sheetcells as CSV**' on the same category page.

Figure 20 illustrates a downloaded .csv file from the category page. A new category called 101 has been added against the first sheetcell. When the .csv file is uploaded a new category called 101 will be created and the sheetcell with ID 1263 will be linked to the new category. The import value field displays the raw data value from the original data upload; this is never deleted and is always available for reference.

Figure 20: Splitting categories in the .csv file

1	ID	IMPORT VALUE	COLUMN HEADER	DATASET	NEW CATEGORY SHORT
2	1263	1	CSP	Wood density of tree species	101
3	1167	1	CSP	Wood density of tree species	
4	1185	1	CSP	Wood density of tree species	
5	1285	1	CSP	Wood density of tree species	
6					

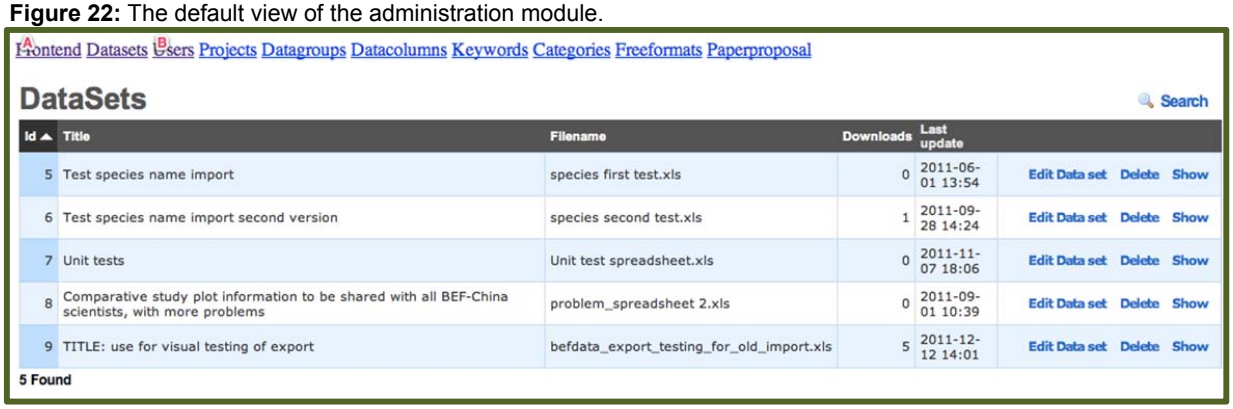
6 The Admin interface

BEFdata portal administrators can add new users and projects, and add users to projects. Administrators have an additional **'Admin'** tab in the main navigation (see Figure 21)

Figure 21: The location of the Admin tab.

The admin interface has a page for each object in the portal which are accessible by links at the top of the page (see Figure 22). The default page is the dataset list and the first link in the navigation takes the user back to the main website, and out of the admin module.

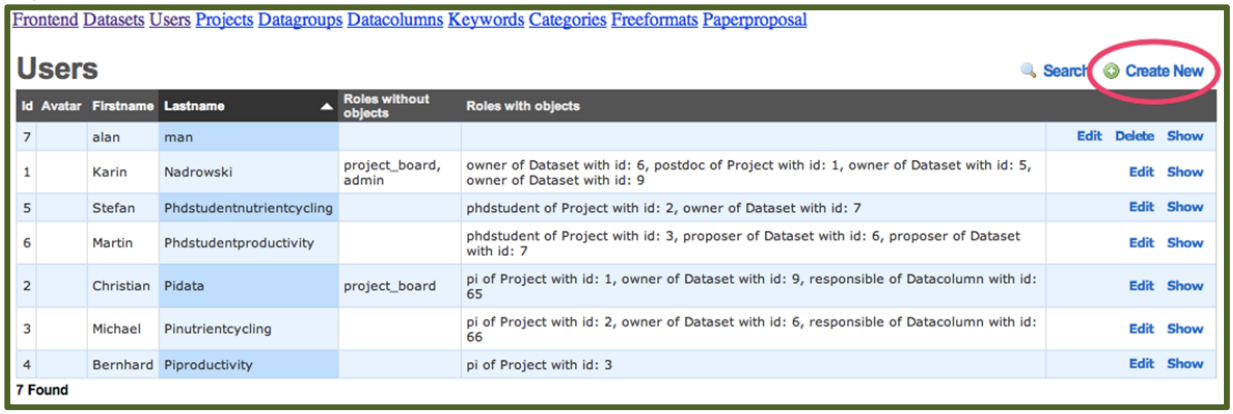
Figure 22: The default view of the administration module.



6.1 Adding users

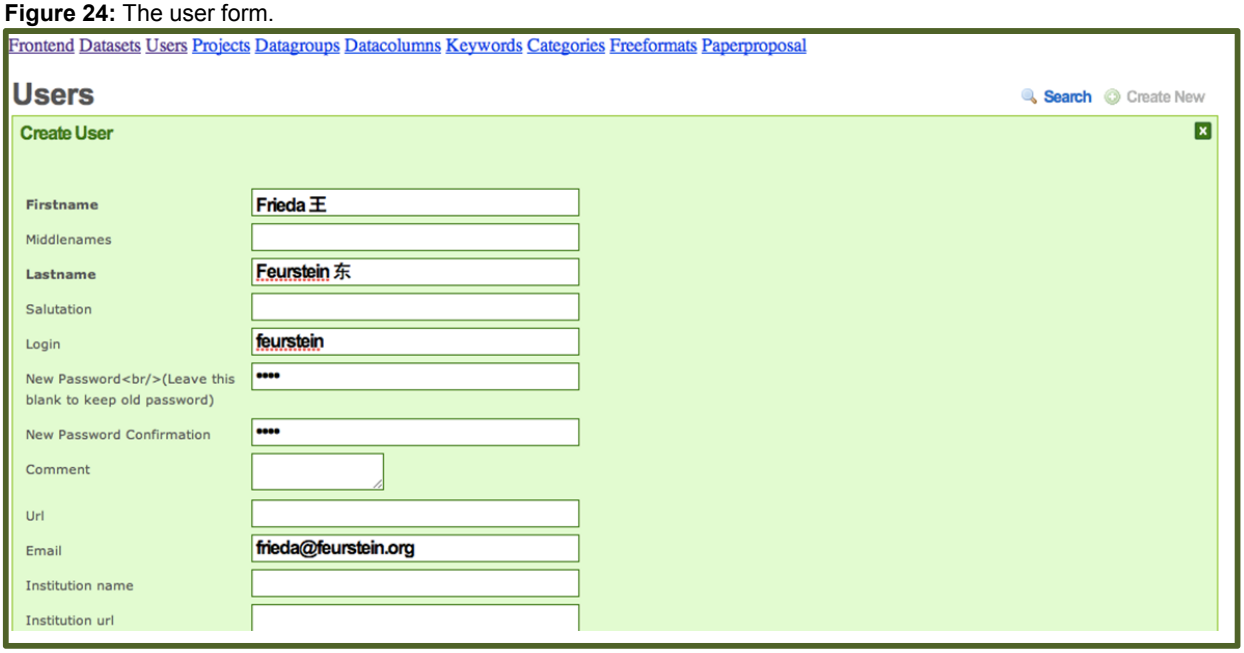
To add new users click on **'Users'**, you will be presented with a list of all the users in the portal, then click on **'Create New'** (see Figure 23)

Figure 23: The location of the 'Create New' user link.



To create a new user the first name, lastname and email address must be provided. To allow the user to log on they must be given a login and password. The password can be changed later by the user using the profile tab in the frontend.

Figure 24: The user form.

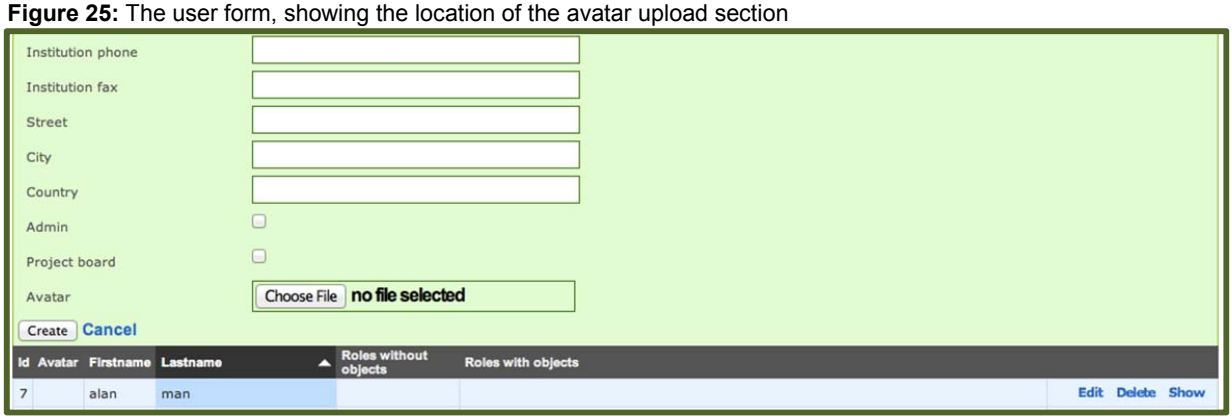


The user can be given an admin, project board and/or data admin role. See Table 4 for the list of roles and their permissions. Members of the project boards have responsibilities

for the overall contents of the data and how it is used, including approval of paper proposals. Project board members also can view the admin comments of a dataset.

The user can be given a picture (avatar, see Figure 25).

Figure 25: The user form, showing the location of the avatar upload section



After clicking the **'Create'** button, the new user can log on. The user now appears in the member list.

Table 4: Role permission

Roles	Permissions
Project board	View comments Review paper proposals
Data admin	Manage categories Download all datasets View download log View comments
Data owner (for a dataset)	Edit and reupload dataset View download log View comments

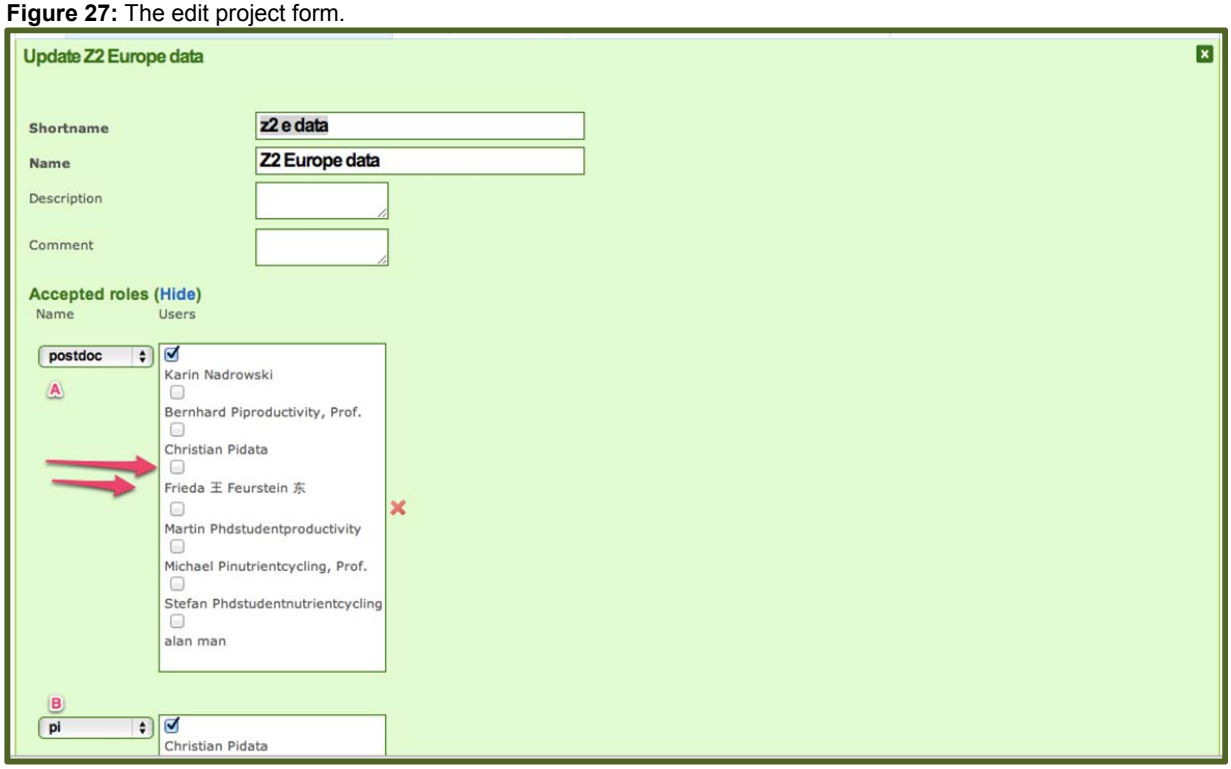
6.2 Adding a user to a project

To add a user to an existing project go to the **'Projects'**, so that they appear on the projects page, link on the top of the admin interface and select the project by clicking on the **'Edit'** link. We will add Frieda to the **'z2 e data'** project.

Figure 26: The administration project list.

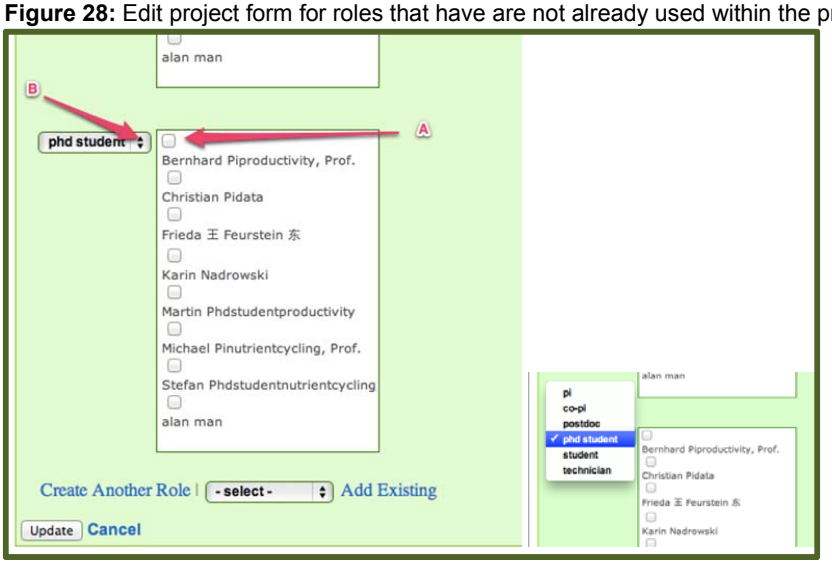
The form to edit projects includes list of all portal users for every available role (see Figure 27, A and B). To add the user to a role already used in the project select the checkbox **above** above the users' name.

Figure 27: The edit project form.



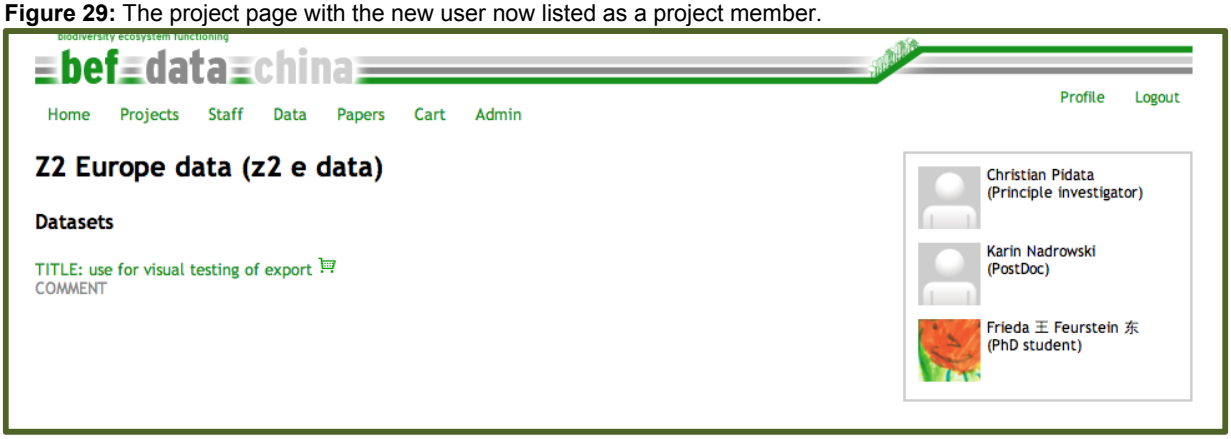
To add the user to a role not already used in the project, select the role from the dropdown list at the bottom of the form. A new list of users will be displayed. The user can then be selected ticking the checkbox above the users' name.

Figure 28: Edit project form for roles that have are not already used within the project



The project page for 'z2 e data' now lists the new user as PHD student in the website.

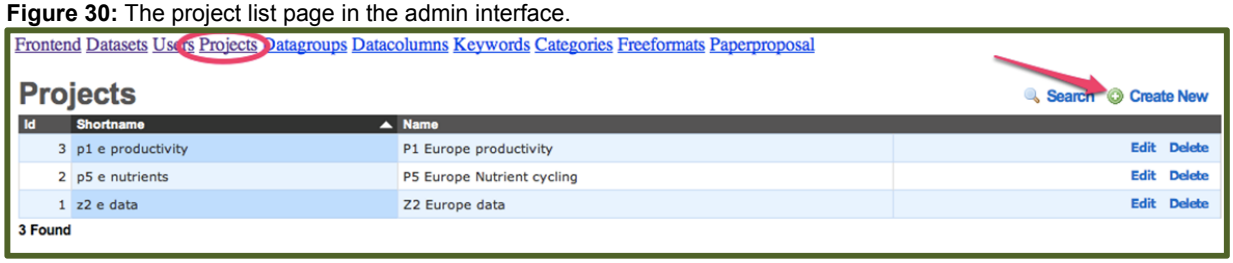
Figure 29: The project page with the new user now listed as a project member.



6.3 Adding a project

To add a new project click on the **'Projects'** link at the top of the admin module, and then **'Create New'** (see Figure 30).

Figure 30: The project list page in the admin interface.



The project only requires a code and a name. When deciding on the name for a project it is worth considering that the project list in the website are sorted by their titles in alphabetical order.

7 Glossary

Name	Description
Workbook	MS Excel 2003 Workbook formatted to work with the BEFdata portal.
Worksheet	A sheet in the Workbook. There are five sheets that need to be completed to upload data to the portal from the workbook.
Dataset	The complete set of data, including the metadata and raw data.
Datacolumn	A column of data either in the Workbook or in the Dataset.
Datagroup	Sematic classification of the type of data in the Datacolumn.
Datatype	The type of data in the datacolumn: text, number, year, date or category.
Category	Data values of Category datatypes that can be reused between Datasets.
Paper proposal	Data requests accompanied by formulated research ideas
EML	Ecological Metadata Language. It is implemented as a series of XML files.
Cart	Web page on the BEFdata portal that stores the datasets a member has selected for inclusion in a paper proposal.

8 Acknowledgements

The BEFdata portal was developed within the Biodiversity-Ecosystem Functioning (BEF-China http://www.bef-china.de/) research unit of the German Science Foundation (FOR 891).

The BEFdata portal development team is:

Biology Institute I - Special Botany and Functional Biodiversity, University of Leipzig
Karin Nadrowski
Sophia Ratcliffe
Claas Thido Pfaff

Topic Maps Lab, Natural Language Processing Group, University of Leipzig
Michael Prilop
Karl Welter
Daniel Seifarth
Sven Windisch

Institute of Botany, Chinese Academy of Sciences
Xingxing Man