



B2FIND Usage

About

Documentation about the usage of the EUDAT metadata service B2FIND.

Modified: 25 October 2017

Synopsis

Scientific work largely depends on the availability of published research data for further analysis and re-use. Therefore, the process of finding and easily retrieving relevant data and their metadata has become increasingly important. This demands user-friendly technologies and practices for searching and retrieving data collections within an overarching framework of inter-disciplinary and international archives.

EUDAT provides the [B2FIND](#) service which allows users to find data collections within a pan-European and inter-disciplinary scope. In order to achieve this, EUDAT has developed a comprehensive metadata catalogue and a user-friendly discovery Portal, which exposes the CKAN functionality and API.

The B2FIND repository collects diverse metadata from heterogeneous sources inside EUDAT and presents them in a consistent form. The homogenisation of community-specific data models and vocabularies enables not only the unique presentation of these datasets as tables of field-value pairs, but also faceted search in the B2FIND metadata portal or via an easy-to-use command line tool.

The service provides advanced search functionalities, such as filtering by location and time; or selecting e.g. a specific author or discipline. Furthermore, the service provides transparent access to the scientific data objects through the given references in the retrieved datasets.

How to find and access data using B2FIND

B2FIND is available through a [web-based portal](#), through an [API](#), or through a [command-line tool](#).

Use of the B2FIND Portal

The B2FIND portal can be found at <http://b2find.eudat.eu> and is a graphical, web-based tool which is designed to be self-explanatory. The only required prerequisite to use the web interface is that JavaScript is enabled in the user's browser.

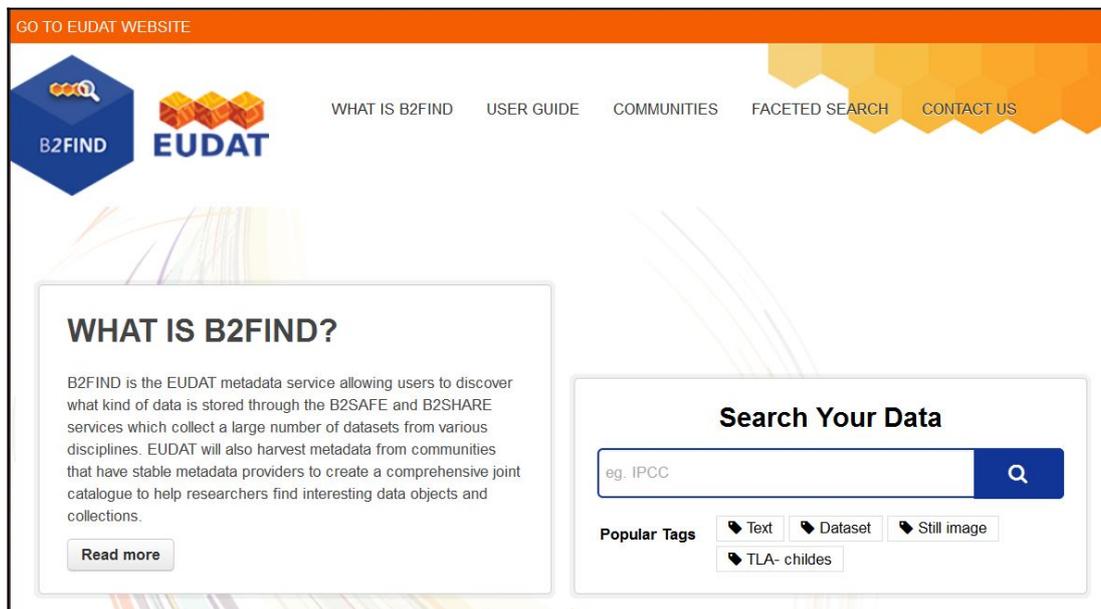


Figure 1. The B2FIND metadata search service.

From the entry page (see Figure 1) you can list all communities that provide metadata to B2FIND by clicking 'COMMUNITIES'. By entering text in the provided free text field 'Search Your Data', you can perform a full text search over the whole catalogue directly from the entry page. This leads you to a result page where all retrieved datasets are listed on the panel on the right.

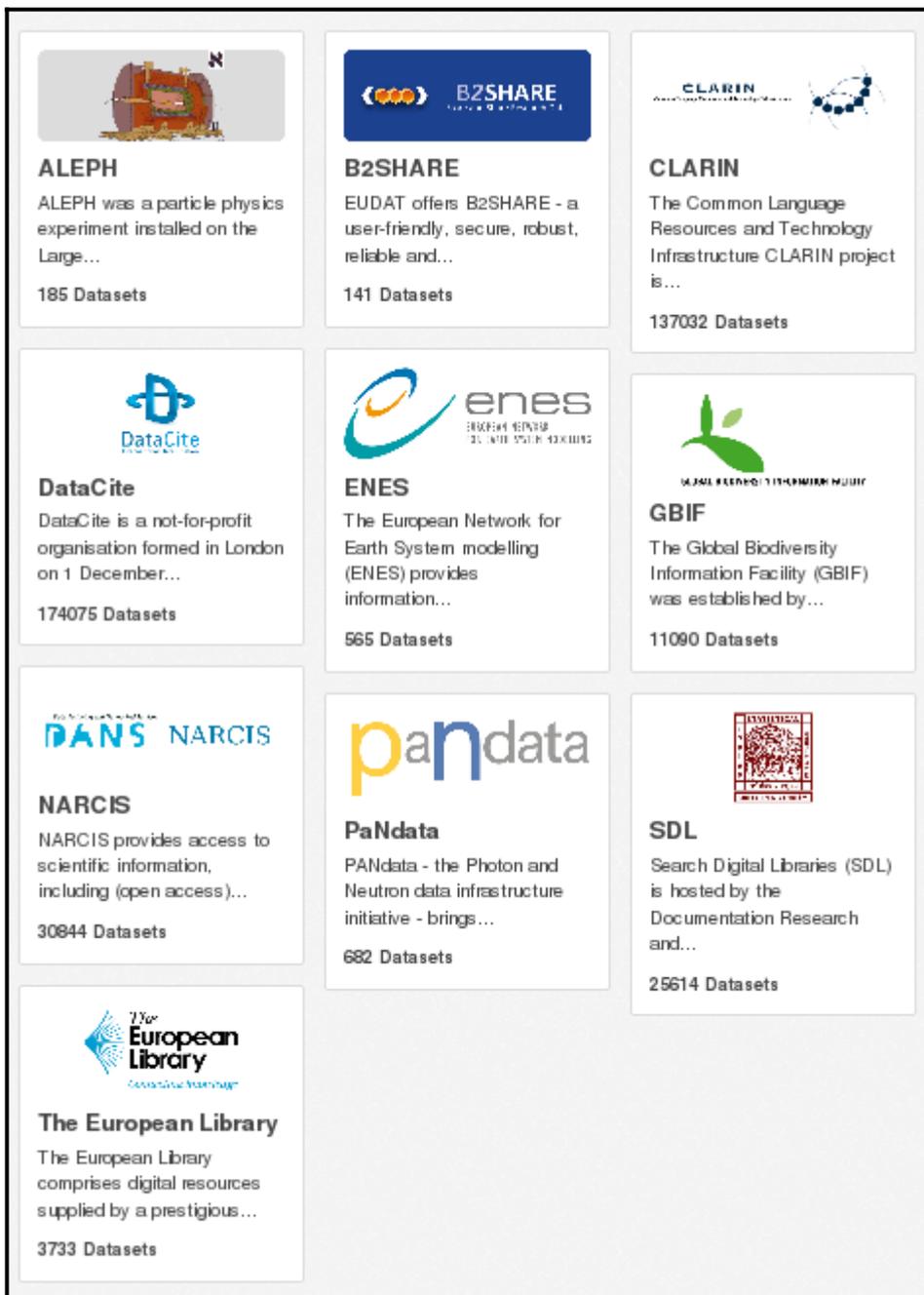
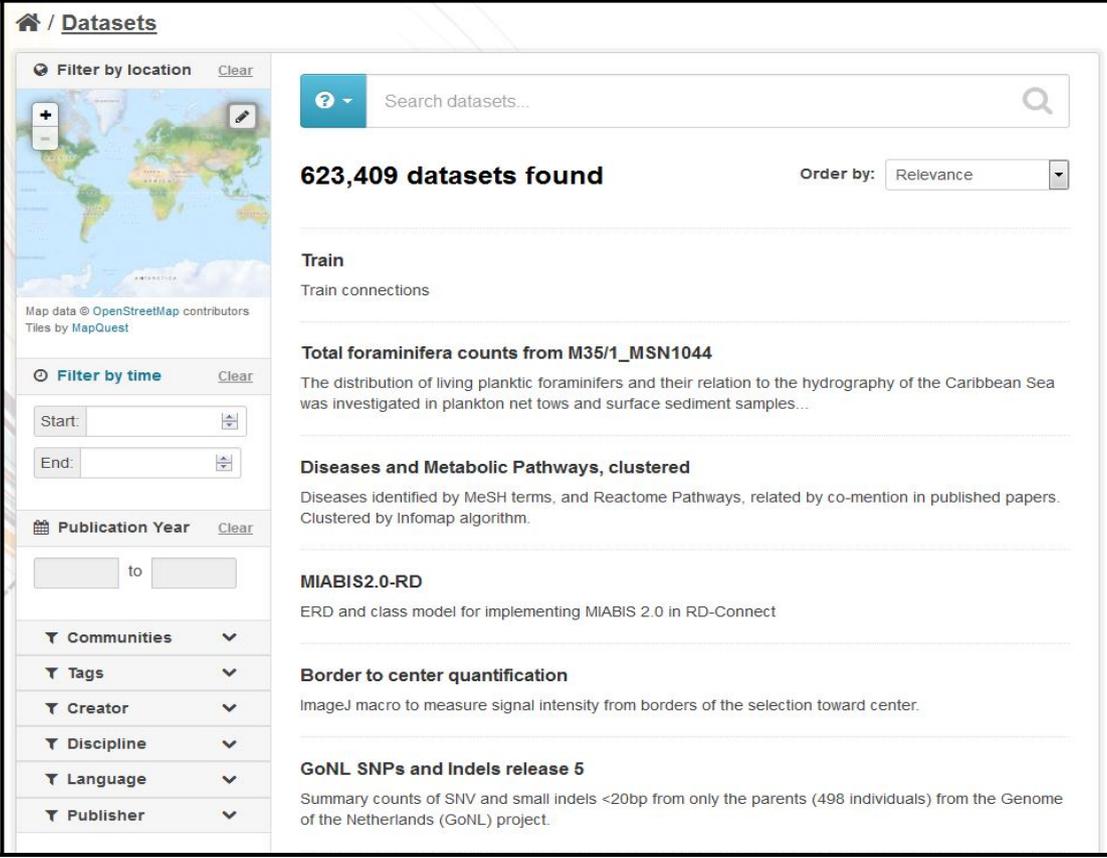


Figure 2. The communities integrated with B2FIND.

By clicking the 'Faceted Search' button, you get directly to a page with all datasets of the metadata catalogue listed (see Figure 3). From here you can use more advanced search functionalities provided in the left navigation bar.



The screenshot displays the EUDAT Datasets search results page. At the top, there is a search bar with a question mark icon and the text 'Search datasets...'. Below the search bar, it indicates '623,409 datasets found' and an 'Order by: Relevance' dropdown menu. The page is divided into several sections, each representing a dataset:

- Train**: Train connections
- Total foraminifera counts from M35/1_MSN1044**: The distribution of living planktic foraminifers and their relation to the hydrography of the Caribbean Sea was investigated in plankton net tows and surface sediment samples...
- Diseases and Metabolic Pathways, clustered**: Diseases identified by MeSH terms, and Reactome Pathways, related by co-mention in published papers. Clustered by Infomap algorithm.
- MIABIS2.0-RD**: ERD and class model for implementing MIABIS 2.0 in RD-Connect
- Border to center quantification**: ImageJ macro to measure signal intensity from borders of the selection toward center.
- GoNL SNPs and Indels release 5**: Summary counts of SNV and small indels <20bp from only the parents (498 individuals) from the Genome of the Netherlands (GoNL) project.

On the left side, there are filter options: 'Filter by location' (with a map), 'Filter by time' (with 'Start' and 'End' date pickers), and 'Publication Year' (with a range selector). Below these are dropdown menus for 'Communities', 'Tags', 'Creator', 'Discipline', 'Language', and 'Publisher'.

Figure 3. Metadata search results page.

There are several ways to find metadata records in the service, and powerful options to browse and filter search results are provided. These include the following:

1. Free text search over the full text body ('Search Your Data', Figure 1)
2. Faceted search, i.e. selecting values of a metadata field or property ('Faceted Search')
3. Filter by location or time, i.e. search for all datasets, that cover a chosen region or time period (visible on search results page Figure 3)

These search requests can be combined and executed in one go by using the advanced search options.

Most of the navigation and filter functionalities are self-explanatory, but an extensive search guide is available by clicking the question mark in the 'Faceted Search' page (Figure 3) or directly from the B2FIND [search guide](#).

Display datasets

By clicking on one of the listed datasets the metadata are displayed (see Figure 4).

Home / Datasets / Microplankton abundance in ...

Dataset extent



Map data © OpenStreetMap contributors
Tiles by MapQuest

Social

- Google+
- Twitter
- Facebook

Dataset Communities Activity Stream Related

Microplankton abundance in the euphotic zone of the Aegean and Levantine Seas in February 1989

Gotsis-Skretas, Olympia; Pagou, Kalliopi (2008): Microplankton abundance in the euphotic zone of the Aegean and Levantine Seas in February 1989, doi:10.1594/PANGAEA.692080

Data and Resources

This dataset has no data

Dinoflagellates

Additional Info

Field	Value
Source	http://dx.doi.org/10.1594/PANGAEA.692080
Author	Pagou, Kalliopi (2008)
Discipline	Biology
GeographicCoverage	SouthernEurope,WesternAsia
MetadataAccess	http://metadata.gbif.org/catalogue/OAIHandler?verb=GetRecord&metadataPrefix=eml&identifier=oai:metadata.gbif.org:eml/portal/oai:metadata.gbif.org:eml/portal/5979.xml
Origin	PANGAEA - Publishing Network for Geoscientific and Environmental Data
PublicationYear	2008
TempCoverageBegin	62739100800
TempCoverageEnd	62739187199
TemporalCoverage:BeginDate	1989-02-15T00:00:00Z
TemporalCoverage:EndDate	1989-02-15T23:59:59Z
oai_identifier	oai:metadata.gbif.org:eml/portal/5979.xml

Figure 4. Example of data set view.

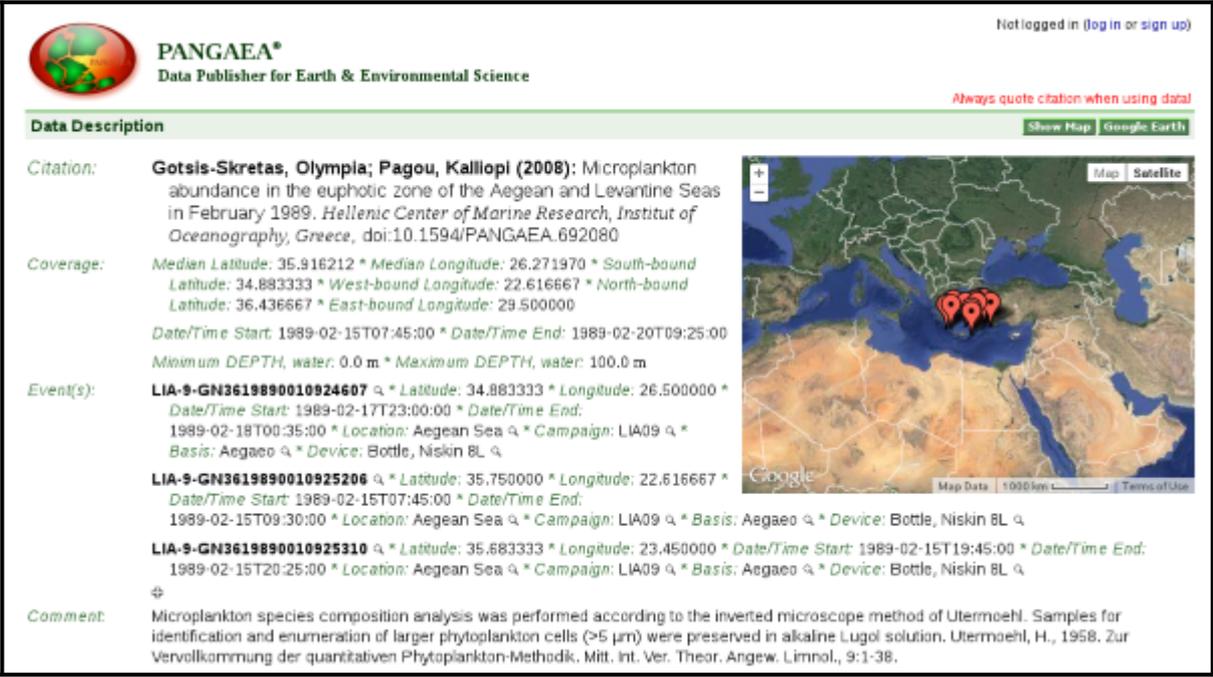
The page has the following format:

- Spatial extent is displayed in the top left corner
- Title and description are shown in the upper range
- All mapped textual B2FIND fields and their values are displayed in the "Additional Info" table. Among them, the following two references provide access to data resources:
 - Source: link to the data object the metadata refer to

- `MetadataAccess`: link to the original harvested metadata XML record (via the [OAI GetRecord](#) request)

Access data resources

The link given in the 'Source' field links to the URL of the data resource behind the metadata. In some cases this leads to the data resource itself or to a "landing page", or it may not be resolvable, e.g. if it is only machine-readable.



PANGAEA*
Data Publisher for Earth & Environmental Science

Not logged in ([log in](#) or [sign up](#))

Always quote citation when using data!

Data Description [Show Map](#) [Google Earth](#)

Citation: **Gotsis-Skretas, Olympia; Pagou, Kallopi (2008):** Microplankton abundance in the euphotic zone of the Aegean and Levantine Seas in February 1989. *Hellenic Center of Marine Research, Institut of Oceanography, Greece*, doi:10.1594/PANGAEA.692080

Coverage: Median Latitude: 35.916212 * Median Longitude: 26.271970 * South-bound Latitude: 34.883333 * West-bound Longitude: 22.616667 * North-bound Latitude: 36.436667 * East-bound Longitude: 29.500000
Date/Time Start: 1989-02-15T07:45:00 * Date/Time End: 1989-02-20T09:25:00
Minimum DEPTH, water: 0.0 m * Maximum DEPTH, water: 100.0 m

Event(s):
LIA-9-GN3619890010924607 * Latitude: 34.883333 * Longitude: 26.500000 * Date/Time Start: 1989-02-17T23:00:00 * Date/Time End: 1989-02-18T00:35:00 * Location: Aegean Sea * Campaign: LIA09 * Basis: Aegean * Device: Bottle, Niskin 8L
LIA-9-GN3619890010925206 * Latitude: 35.750000 * Longitude: 22.616667 * Date/Time Start: 1989-02-15T07:45:00 * Date/Time End: 1989-02-15T09:30:00 * Location: Aegean Sea * Campaign: LIA09 * Basis: Aegean * Device: Bottle, Niskin 8L
LIA-9-GN3619890010925310 * Latitude: 35.683333 * Longitude: 23.450000 * Date/Time Start: 1989-02-15T19:45:00 * Date/Time End: 1989-02-15T20:25:00 * Location: Aegean Sea * Campaign: LIA09 * Basis: Aegean * Device: Bottle, Niskin 8L

Comment: Microplankton species composition analysis was performed according to the inverted microscope method of Utermoehl. Samples for identification and enumeration of larger phytoplankton cells (>5 µm) were preserved in alkaline Lugol solution. Utermoehl, H., 1958. Zur Vervollkommung der quantitativen Phytoplankton-Methodik. Mitt. Int. Ver. Theor. Angew. Limnol., 9:1-38.

Figure 5. Example of data access via the link given in the "Source" field.

Application Programming Interface

The B2FIND Portal is based on the [CKAN](#) platform. B2FIND exposes the full CKAN API, version 2.2. For full documentation [see the CKAN website](#). EUDAT has built a client tool using this API, [see below](#). Feel free to use the B2FIND API to develop your own clients to access the B2FIND service.

Command line usage

The Python script `searchB2FIND.py` - available from the [B2FIND git repository](#) - uses the B2FIND API to allow submission of search requests directly from the command line.

The usage of the script is shown by calling `./searchB2FIND.py -h`, as follows:

```
usage: searchB2FIND.py [-h] [--ckan IP/URL] [--output STRING] [--community STRING] [--ids [IDS [IDS ...]]]
[PATTERN [PATTERN ...]]
```

Description: List identifiers of datasets that fulfill given search criteria

positional arguments:



PATTERN B2FIND search pattern, i.e. (a list of) field:value terms.

optional arguments:

-h, --help show this help message and exit
--ckan IP/URL CKAN portal address, to which search requests are submitted (default is b2find.eudat.eu)
--output STRING, -o STRING Output file name and format. Format is given by the extension, supported are 'txt' (plain
ascii file) or 'hd5' file.
--community STRING, -c STRING Community where you want to search in
--ids [IDS [IDS ...]], -i [IDS [IDS ...]] Identifiers of found records outputed. Default is 'id'. Additional 'Source', 'PID' and
'DOI' are supported.

Examples:

1. `>./searchB2FIND.py -c aleph tags:LEP`
searchs for all datasets of community ALEPH with tag "LEP" in b2find.eudat.eu.
2. `>./searchB2FIND.py author:"Jones*" AND Discipline:"Crystal?Structure" --ckan eudat-b1.dkrz.de`
searchs in eudat-b1.dkrz.de for all datasets having an author starting with "Jones" and belongs to
the discipline "Crystal Structure"
3. `>./searchB2FIND.py -c narcis DOI: '*' --ids DOI`
returns the list of id's and DOI's for all records in community "NARCIS" that have a DOI

Support

You can find B2FIND *training presentations* on the [EUDAT website](#).

You can also find *hands-on training material* on B2FIND on our [github repository](#); section 5 deals with using the service.

Support for B2FIND is available via the EUDAT ticketing system through the [webform](#).

If you have *comments* on this page, please submit them though the [EUDAT ticketing system](#).

Document Data

DocVersion: 1.3

Authors:

Heinrich Widmann, widmann@dkrz.de

Editors:

Hans van Piggelen, hans.vanpiggelen@surfsara.nl

Sara Ramezani, sara.ramezani@surfsara.nl

Kostas Kavoussanakis, k.kavoussanakis@epcc.ed.ac.uk

[Read more](#)