B2SHARE Advanced search

About

User documentation about the advanced search capabilities of the EUDAT B2SHARE service.

Modified: 7 June 2021

Synopsis

B2SHARE is a web-based service for storing and publishing data sets, intended for European scientists. The service utilises other EUDAT services for reliability and data retention, while storing the data at trusted repositories with national backing, in order to provide a professionally managed and supported IT environment.

B2SHARE supports a default set of metadata fields that are defined in the root schema. Every record must provide the mandatory fields in this schema. In addition, the service allows communities to define their own set of metadata fields aggregated in the so-called community metadata schema. Every record that is published as part of these communities must also fill in the mandatory fields in these community-specific schemas. All metadata field values are indexed by B2SHARE and are therefore searchable using the web interface and REST API.

This document gives an introduction in the possiblities for searching using queries. For a complete overview of all possibilities, see the Elasticsearch documentation on search queries.

How to access the B2SHARE service

B2SHARE is available from the following URL: https://b2share.eudat.eu.

Searching in B2SHARE

Search is available directly on the front page of the web interface or after clicking on the 'Search' button on that same page. Searching for any keyword always ends up on the advanced search page of B2SHARE.

To search using the REST API, append the query normally entered in the search box to the search API URL as the value of the parameter 'q': https://b2share.eudat.eu/api/records?q=<query-string>. For more information on searching through the REST API, see the relevant section in the B2SHARE REST API documentation.

Underlying implementation

B2SHARE indexes all the metadata of each record using Elasticsearch search technology. After each record creation, versioning or metadata update, this index will be updated automatically and therefore all changes can immediately be found using the search functionality.

Elasticsearch indexes all metadata fields by not storing the exact field values but the more generic equivalents instead. In practice this means that when a singular keyword is entered, also plural equivalents are matched successfully.
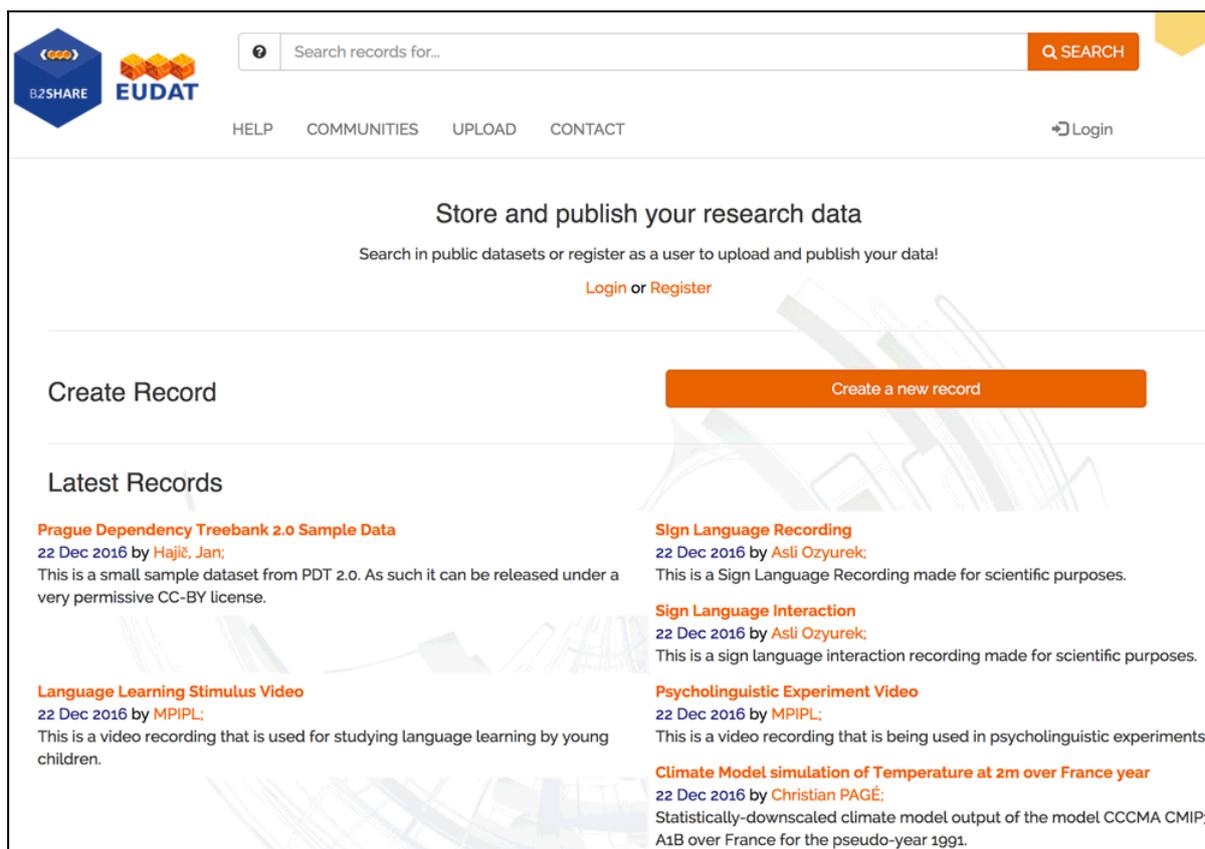
Search records

Registered and unregistered users can use the search field on the B2SHARE home page (Figure 1). Enter a search

query in the search field and click the "Search" button. The text entered can be part of a title, keyword, abstract or any other metadata, including values for metadata fields defined in community-specific schemas.

Unregistered users can only search for data sets that are publicly accessible. The default search mode is simple search, which provides an input box where your queries can be typed. Usually it is sufficient to just type some keywords one is interested in and hit return. Please note that only the latest version of matching records is shown in the search result.

Figure 1. The B2SHARE front page with search bar.

Multiple keyword search

You can enter multiple keywords at once to either find records that match all of the keywords or any of the keywords. To search for exact multi-keyword matches, encapsulate the keywords in double quotes. See for example the difference between the results for '"container technologies"' and 'container technologies'.

Logical search operators

To make a distinction between all or any matching keywords or to exclude a specific word, use the logical operators 'AND', 'OR' or 'NOT'. The default operator between multiple keywords is 'OR'. Always make sure to use capitals, otherwise the operator will be interpreted as a keyword. To see the difference between the operators, compare the searches between 'lofar AND singularity', 'lofar OR singularity' and 'lofar NOT singularity'.

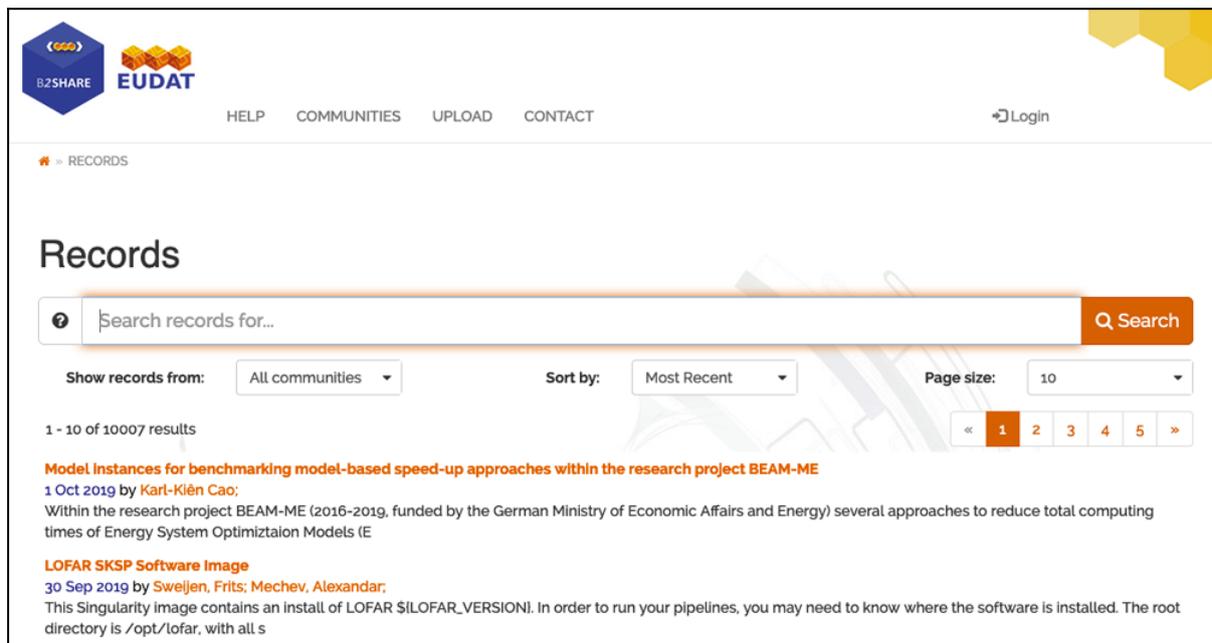It is also possible to use the character equivalent for each operator. For an overview, see Table 1.

Action	Operator	Equivalent
Include	AND	&&
Optional	OR	||
Exclude	NOT	!

Table 1. Search operators in B2SHARE.

Advanced search

To get to the advanced search page, hit the search button on the front page and select the additional options in the new form below the search text field (Figure 2). Using the form on this page, the search can be narrowed down to a specific community, sorted by most recent or best matching records and the page size or number of records returned per page can be adjusted.

Figure 2. The B2SHARE advanced search page.

Field-specific searches

Next to simply entering one or more keywords in the search bar's text field, it is also possible to make your search specific to a metadata field in the B2SHARE root schema or one of the communities metadata schemas. Searching for specific metadata fields requires you to prepend your keyword with the metadata field's internal name and structure followed by a colon (':'). Thus if you want to limit your search to the title(s) of records only, use the 'titles.title' prefix. See for example the difference between 'titles.title:container' and 'container'.

To find the corresponding internal field names and structure, you can use the information provided for the schema of a community, e.g. the EUDAT community. This information is also available through the B2SHARE REST API, see

for example the schema definition of the EUDAT community. A non-exhaustive overview of all root schema fields is given in the Table 2.

Field name	Internal name
Title	titles.title
Creator	creators.creator_name
Description	descriptions.description
Resource type	resource_types.resource_type
Language	language
Contributor	contributors.contributor_name

Table 2. Some examples of metadata fields and their internal name and structure equivalents.

If you want to use logical operators within a specific field search, encapsulate the value for the field in parentheses, for example 'titles.title:(container NOT technologies)' returns records containing the keyword 'container' in the title field, but not 'technologies'.

Community-specific metadata field searches

Searching for values in metadata fields defined by communities specifically works similarly to searching for field-specific values.

Finding the community's metadata schema identifier

One important difference is that the community metadata schema's internal identifier is required. This so-called 'block schema identifier' value can be found on the community landing page in the section for the community-specific metadata fields. Use the copy button to copy the identifier to your clipboard.

Searching for a community-specific metadata field value

Using the identifier value of the community's block schema of InGrid 'fccd46c7-db79-460b-ad34-abf078d194a3', it is possible to search specifically for one of the metadata fields, e.g. 'Timespan'. This works as follows: prepend the value with 'community_specific', followed by the schema identifier, followed by the field's internal name and closed by a colon following the value. For the value '2016' this leads to 'community_specific.fccd46c7-db79-460b-ad34-abf078d194a3.time_span:2016'. Naturally, since this field is only defined for the InGrid community, all found records are part of this community.

Searching with date ranges

B2SHARE allows to filter your search by using ranges. Using the example of the previous section to find all records having a timespan after 2019, use 'community_specific.fccd46c7-db79-460b-ad34-abf078d194a3.time_span:(>2019)'.

Filtering by creation or modification date

To find all records between two specific dates or for a given year one can use the record's fields that represent the creation and modification dates. See Table 3.

Field name	Internal name
Creation date	_created
Modification date	_updated

Table 3. Technical metadata fields and their internal name for doing date-related queries.

Using this information one can select all records created in 2020 using the following query: '_created:(>2020 AND <2021)'. To select all created records for a given quarter, say Q1 2021, use '_created:(>2021 AND <2021-04)'.

Support

Please visit our training site on GitHub for [B2SHARE](#) and other hands-on training material.

Our [B2SHARE presentations](#) offer training material for the service.

Support for B2SHARE is available via the EUDAT ticketing system through the [webform](#).

If you have comments on this page, please submit them through the [EUDAT support request system](#).

Document Data

Version: 1.1.1

Authors:

Hans van Piggelen (SURF)

[Read more](#)