

# O2A specification for PANGAEA sources (.sdi.pangaea.lst)

- [Preface](#)
- [Version 0.1](#)
  - [Requirements](#)
  - [Notes](#)
  - [Query Syntax/Formatting](#)
    - [Query Examples](#)
    - [File Examples](#)

## Preface

The [Spatial Data Infrastructure](#) can be fed with data published on [PANGAEA](#) to have it published as (part of) an OGC Web Service (OWS). How to specify which data our SDI should integrate/harvest from PANGAEA can be specified with `.sdi.pangaea.lst` files. They can hold both specific [PANGAEA DOIs](#) and [PANGAEA Search Queries](#). The latter will result in all resulting DOIs to be integrated/harvested into the SDI. All of an DOI's content which is available as PANGAEA tab files will get integrated/harvested.

## Version 0.1

### Requirements

- plain text file
- file extension: `.sdi.pangaea.lst`
- encoding: UTF-8
- one PANGAEA URL (query or direct DOI) per line
  - entries get separated by [newline](#)
  - no other separators allowed!
  - see section on [Query Syntax/Formatting](#) ⚠

### Notes

- readability features for humans
  - empty lines are fine
  - lines beginning with a hashtag (#) can be used for comments
    - no comments in lines with URLs
- a file's basename (e.g. `foobar` from `foobar.sdi.pangaea.lst`) can be used for convenient structuring but will not have any impact on SDI /OWS issues.
  - [Exchange Folder Specification](#) might apply

### Query Syntax/Formatting

#### Advanced Search

The used syntax formatting can be explored in the [PANGAEA WIKI: Advanced Search](#) webpage. But it must be used as decoded link address without '+' signs.

- **do copy** filter from PANGAEA Search Bar content (and combine it with <https://www.pangaea.de/?q=>)
- **do not copy** filter from browser location bar (URL)

https://www.pangaea.de/?q=device%3A%22Buoy%22+AND+location%3A%22Arctic+Ocean%22+AND+(project%3A%22AWI\_Sealce%22+OR+project%3A%22AWI\_BioOce%22)

PANGAEA.

ALL TOPICS

device:Buoy AND location:Arctic Ocean AND (project:AWI\_Sealce OR project:AWI\_BioOce)

89 datasets found on search for »device:Buoy AND loca...«

1. Nicolaus, M; Hoppmann, M; Lei, R et al. (2021): Snow height on sea ice, meteorological conditions and drift of sea ice from autonomous Snow Buoys during MOSAiC 2019/20

## Query Examples

- **valid** usage of **Advanced Search Syntax**:
  - Decoded URL: [https://www.pangaea.de/?q=device:Buoy AND location:Arctic Ocean AND \(project:AWI\\_Sealce OR project:AWI\\_BioOce\)](https://www.pangaea.de/?q=device:Buoy AND location:Arctic Ocean AND (project:AWI_Sealce OR project:AWI_BioOce))
- **invalid** usage of **Advanced Search Syntax**:
  - Encoded URL: [https://www.pangaea.de/?q=device%3A%22Buoy%22+AND+location%3A%22Arctic+Ocean%22+AND+\(project%3A%22AWI\\_Sealce%22+OR+project%3A%22AWI\\_BioOce%22\)](https://www.pangaea.de/?q=device%3A%22Buoy%22+AND+location%3A%22Arctic+Ocean%22+AND+(project%3A%22AWI_Sealce%22+OR+project%3A%22AWI_BioOce%22))
- **invalid**, **Facet Filtering Syntax**:
  - Decoded URL: [https://www.pangaea.de/?f.method\[\]=Buoy&f.location\[\]=Arctic+Ocean&f.project\[\]=AWI\\_Sealce&f.project\[\]=AWI\\_BioOce](https://www.pangaea.de/?f.method[]=Buoy&f.location[]=Arctic+Ocean&f.project[]=AWI_Sealce&f.project[]=AWI_BioOce)
  - Encoded URL: [https://www.pangaea.de/?f.method%5B%5D=Buoy&f.location%5B%5D=Arctic+Ocean&f.project%5B%5D=AWI\\_Sealce&f.project%5B%5D=AWI\\_BioOce](https://www.pangaea.de/?f.method%5B%5D=Buoy&f.location%5B%5D=Arctic+Ocean&f.project%5B%5D=AWI_Sealce&f.project%5B%5D=AWI_BioOce)
  - **Facet Filtering Syntax is NOT supported!**

https://www.pangaea.de/?q=project:dam

PANGAEA.

ALL TOPICS

project:dam

Advanced Search

79 datasets found on search for »project:dam«

1. Bartholomä, A; Höller, P; Wisshak, M (2020): Multibeam bathymetry raw data (EM 1002 echosounder entire dataset) of RV MARIA S. MERIAN during cruise MSM55

2. Schoening, T; Kopte, R; Mohrmann, J et al. (2020): ADCP current measurements (75 kHz) during Maria S. Merian cruise MSM56

3. Devey, CW; Le Saout, M; Petersen, S et al. (2020): Multibeam bathymetry raw data (Kongsberg EM 712 entire dataset) of RV MARIA S. MERIAN during cruise MSM75

4. Damasko, D; Becker, M (2021): Water column raw data (Kongsberg EM712 entire dataset) of RV MARIA S. MERIAN during cruise MSM97/2

5. Römer, M; Galde, S; Wolff, A-C (2021): Multibeam bathymetry raw data (Kongsberg EM 712 transit dataset) of RV MARIA S. MERIAN during cruise MSM98

6. Reicherter, K; Wolff, A-C (2020): Multibeam bathymetry raw data (Kongsberg EM 122 entire dataset) of RV METEOR during cruise M152

Facet Filtering

## File Examples

Valid File

#### valid\_example.sdi.pangaea.lst

```
# single DOI(s)
https://doi.pangaea.de/10.1594/PANGAEA.123456
https://doi.pangaea.de/10.1594/PANGAEA.234567

# Filtered by project label 'DAM' as decoded URL
https://www.pangaea.de/?q=project:label:DAM

# Filtered by method and project as decoded URL
https://www.pangaea.de/?q=method:"Multicorer with television" AND (project:Hausgarten OR project:HERMES OR
project:FRAM)
```

#### Invalid File

#### invalid\_example.sdi.pangaea.lst

```
# single DOI with invalid comment behind entry
https://doi.pangaea.de/10.1594/PANGAEA.123456 # my favourite DOI!

# single DOI with invalid address (missing https prefix)
doi.pangaea.de/10.1594/PANGAEA.234567

# Filtered by project label 'DAM' ...
# A) using Advanced Search Syntax but as encoded URL
https://www.pangaea.de/?q=project%3Alabel%3ADAM

# B) URL copied from address bar after using facet filtering
https://www.pangaea.de/?f.project[]=DAM

# URL copied from address bar after using facet filtering -- works in browsers but not for harvesting
https://www.pangaea.de/?q=method:"Multicorer+with+television"&f.project[]=Hausgarten&f.project[]=HERMES&f.
project[]=FRAM
```