

Become a Content Provider

You're in the right place if you want to offer metadata for **searching** in the portal. If you want to provide data products for the portal **viewer** with our spatial data infrastructure, look [here](#).

To become a content provider for the portal you have to provide metadata for your datasets, articles or map projects. We basically support the defacto standardized [OAI-PMH](#) and the [OGC CSW](#) metadata harvesting interfaces. One of these standards should be accessible from the internet via HTTPS without authentication. If you have no catalog and standardized interfaces to your content until now, you might have a look at the [GeoNetwork](#) solution. Different metadata schemata and profiles are supported, e.g. [ISO19115](#), but also customized ones. For special cases, also other web-resources can be harvested, but are not preferred.

Checkpoints for basic metadata

- ☐ Title
- ☐ Publication date
- ☐ Temporal extent of data coverage (begin date, end date)
- ☐ Geographical extent of data coverage (point, bounding box, polygon or line string)
- ☐ Author(s) (lastname, firstname, email, [ORCID](#))
- ☐ Abstract or description
- ☐ Stable URL to splash page with additional information
- ☐ [DOI](#) or [Handle](#) if existing
- ☐ Type of object (data, publication, article, map)

Checkpoints for extended metadata profiles

- ☐ Platform metadata, e.g. research vessel, Polarstern
- ☐ Device metadata, e.g. Thermosalinograph (device type)
- ☐ Expedition or campaign metadata, e.g. [PS101](#)
- ☐ Event metadata, e.g. [PS101/058-1](#) (date, description, latitude, longitude)
- ☐ Measured parameter, e.g. temperature
- ☐ URL(s) reference(s) to related data services or download
- ☐ URL(s) reference(s) to further information and related content

 You provide basic metadata and support standard interfaces? Fine. Read about [Harvesting and Indexing](#).


Harvesting and Indexing


We are using an internal JSON metadata format ([schema](#), [example](#)) for harmonized indexing and full-text search against an [Elasticsearch](#) cluster. Metadata is harvested regularly once a night from configured content providers. Metadata is annotated, e.g. with information from Gazetteers to resolve geographical locations to place names, during harvesting. So users are able to search also for places names not only for geometries.

As content provider you have to map your metadata content against the given [JSON Schema](#). Best you are using international vocabularies for your metadata content. For example

- [ORCID](#) for persons
- [ROR](#) for organisations
- [PANGAEA](#) or [CF Standard Names](#) for parameters
- [SeaDataNet device categories](#) for names of instruments

 The list of vocabularies is not fixed and can be easily expanded. But you really should map wherever possible to achieve harmonization.

 There is a harvesting and indexing **sandbox** where you can test your mappings, [harvesting](#) and [search](#). Please contact o2a-support@awi.de to get access.

 To finally become a content provider contact o2a-support@awi.de.