

Important aspects for the EDP harvester

- Metadata standard: CKAN / INSPIRE / DCAT-AP
- Metadata representation: XML / JSON
- API type: CKAN / OAI-PMH / dump file
- Controlled vocabulary is used for the categorization, consisting of 13 categories. The harvested portal should use the same categorisation or map their categories against the 13 provided
- The European Data Portal harvester is configured manually per portal by the EDP team using the information provided via the API.

How to be harvested by us?

1. Learn more about how the EDP harvester works on the next slide
2. Go to <http://www.europeandataportal.eu/en/content/providing-data/how-to-be-harvested-by-us> and learn how to be harvested

Manual configuration of the harvester

1. Create a repository for the new portal in our harvester component and type in some information
 - Name of the Portal
 - URL of the Portal
 - URL of the metadata sources (API, ZIP, etc.)
 - Language of the portal
 - Publisher name and mail address
 - Is the harvester incremental or not
 - Shall the harvester be visible to the public or not
 - A description of the portal
 - The European countries, covered by the metadata
2. Create a new harvester and select the new repository as source and the European Data Portal as destination for the harvesting
3. Configure the harvester
 - Name
 - Description
 - Visible or not
 - Harvesting frequency: e.g. daily or manually
 - Begin of harvesting (first harvesting)
4. Write transformation script to convert source portal metadata to destination portal metadata
 - Help from source portal providers is highly appreciated

