DwB
Data without Boundaries

The Data without Boundaries – DwB – project exists to support equal and easy access to official microdata for the European Research Area, within a structured framework where responsibilities and liability are equally shared.

The DwB Resource Discovery Portal (RDP)

The DwB RDP is composed of the following parts:

A Harvesting Framework with related tools to facilitate the retrieval and ingestion of metadata in various formats from a variety of participating providers.

An Administrator Dashboard with underlying management services to configure and monitor the infrastructure, and orchestrate the various tasks.

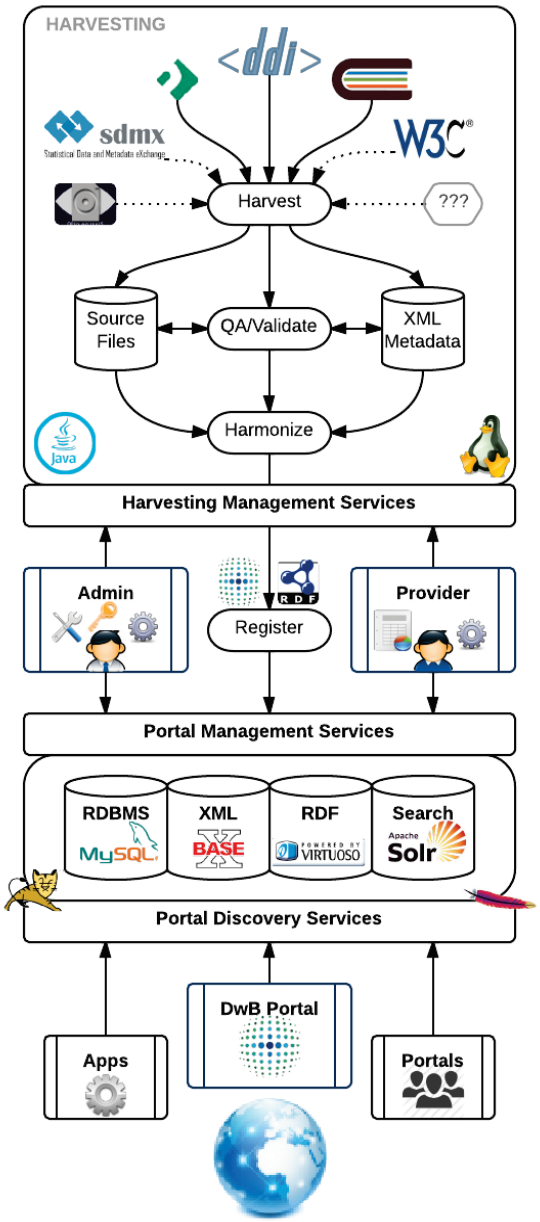
A Provider Portal enabling participating agencies to gain insights on their information present in the system. This include reports (quality assurance, usage, harvesting) and tools to control their profile, visibility of metadata, or profile information.

A Storage Framework (databases) where the information is hosted and indexed in various shapes and form (Relational, XML, RDF) in order to support the search, discovery and other services built on top of that. The metadata is based on commonly used standards and the DwB RDP Metadata Model.

Discovery Services exposing the platform and information to the outside world, enabling integration in applications or web sites, along with the Resource Discovery Portal user interface

The implementation leverages several open sources technologies and packages, chosen for their robustness. The DwB-RDP prototype implementation is woven around these in an innovative fashion to deliver an enterprise grade scalable solution.

DwB
Data without Boundaries
WORK PACKAGE 12
DATA DISCOVERY PORTAL



Discovery Model

The DwB Resource Discovery Portal is built on a metadata model starting from DDI as a representation of community best practices and standards in metadata development. This initial model has been enriched against the DDI Discovery RDF (Disco) and SKOS (Simple Knowledge Organization System), to support richer functionalities and features. The DCAT Application profile for data portals in Europe developed for describing European public sector datasets is also relevant.

Ingestion standard

To harvest and register information into the portal repository, both development lines of the DDI, Codebook and Lifecycle, are identified as core specifications. DDI is a flexible specification that can be used in different ways and, as a consequence, the XML content can vary significantly depending on the metadata producer, depending on which elements are used and how they are populated. DwB RDP therefore cannot generically support any form of DDI-C or DDI-L but rather need to focus on known tools, outputs or flavors. Known project specific flavors are supported on a case by case basis. It is expected that other flavors of harvesters will emerge as additional metadata providers are identified. The same is true for supporting other metadata standards such as SDMX, DCAT, or Dublin Core. Any format that can be mapped to DwB-Disco can potentially be ingested by the discovery portal.

Minimal Metadata

To be able to register a resource with the portal, a minimal amount of metadata is necessary to ensure that

- it can be properly indexed,
- meaningful information can be delivered to the user, and
- the user can be redirected to the resource location.

The DwB RDP presently specifies the following minimal set:

Title: the resource name, to display in search result and for text search
***Abstract*:** A minimum level abstract is needed, to give a description of the *content* of the data collection.
Geography: the spatial coverage of the resource underlying data (this is a fundamental search criterium). The minimal requirement is to provide a country.
Time: the time period the resource underlying data covers (at least a year).
Location: a url where the user can be redirected to retrieve or find additional information around gaining access to the resource.

Provider Portal

Organisations that provide metadata to the portal will get access to the portal's provider services. The Provider Portal gives access to the organization's portion of the harvested/indexed metadata holdings, and to a set of tools to help test and improve quality and consistency of the metadata, as well as re-indexing.

Portal Features

The purpose of the discovery portal is to provide mechanisms to both researchers and applications to search and retrieve metadata for the catalogued datasets. It currently allows search on study level, but will in future versions include variable search. The Portal UI consists of three separate views, the discovery portal home page, the search results page, and the study details page. The portal home page displays the search bar and information about the discovery portal client (portal UI) and its features. It explains how the search bar, facets, and search results work. In general this landing page provides users with the information they might need to use the site.

DwB-RDP: DwB Resource Discovery Portal.

URL: http://dwb-dev.nsd.uib.no/portal.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 262608