

Dr Anna Wałek

Dr Piotr Krajewski

Gdańsk University of Technology Library

Poster proposal:

Linking research information within the knowledge portal - the Bridge of Knowledge services for researchers

In 2016, the Gdańsk University of Technology (Gdańsk Tech) launched the project titled "Multidisciplinary Open System Transferring Knowledge" (in Polish: "Multidyscyplinarny Otwarty System Transferu Wiedzy – MOST Wiedzy"). The acronym of the project (MOST Wiedzy) means "the Bridge of Knowledge". The project's main aim was to build an IT platform (mostwiedzy.pl) that supports sharing information about universities' achievements and the research potential of their researchers. The second stage of the project, the implementation of which began in 2018, is a continuation of the previous one. The "Bridge of Data" project (Bridge of Data – Multidisciplinary Open System Transferring Knowledge. Stage II Open Research Data) was established to support researchers in their activities regarding different Open Science layers. It provides both an open data repository and a range of additional services for researchers who want to share their research results openly. The project is unique in this part of Central and Eastern Europe. It involves three Pomeranian universities: Gdańsk University of Technology (leader), the University of Gdańsk, and the Medical University of Gdańsk. More universities and research institutes are now joining the consortium and will use the infrastructure that has been created.

From a technical perspective, the Bridge of Knowledge platform integrates multidisciplinary and multidimensional data from many databases. The metadata describes each piece of data to connect with other objects and to machines that understand the metadata format. In addition, there are e-services dedicated to sharing information about research output, its metrics and potential.

As a result of the Bridge of Knowledge project, the mostwiedzy.pl platform was created. It contains databases of related data and repositories and is divided into five modules.

The first one, "Research infrastructure", includes databases containing:

- research teams – contains a list of team members, their research interests, research offer and service offer (e.g. for the purposes of cooperation with business);
- research equipment – the record contain the name of the manufacturer, parameters, laboratory and the maintainer of the equipment;
- laboratories – the record contains detailed information about the laboratory equipment, people associated people, research topics being pursued, commercial offering, related research data, and other resources.

The second module is named "Ministry points" and refers to the Ministry of Education and Science scores on which the scientific evaluation of researchers and research units is based. These points are awarded to journals, conference materials and monographs and are published on the list of publishers and journals announced by the Ministry. The "scientific journals" modules create a separate extensive database of journals and publishing policies.

"Scientific activity", the third module, contains people, projects and inventions databases. One of the critical elements is the scientific profile. It has information about researchers, publications and ongoing projects, scientific achievements, organisations, didactic activities, and shared research data. The profile is completed with a bio and can be linked to profiles on social media.

The fourth module includes three repositories that are behind the implementation of the Open Science policy. These are, in turn:

- publications repository – containing bibliographic records of all publications, authored by people with a profile on the portal;
- Open Access Repository, containing publications available in full text, following the Open Access green route;
- Open Research Data repository, containing data sets.

The last fifth module includes additional databases, including events (both scientific and university life events), descriptions of online courses available on the university's e-learning platform and a virtual microscope (digital pathology).

The Open Access Repository contains records of all publications in full text. The database enables searching and sorting results, as well as refining the search according to specific criteria. A publication record contains bibliographic data, citation information, altimetric data, export to various citation styles, license information, and a full-text button. It also provides publication statistics such as the number of views and downloads and an intelligent mechanism for suggesting similar publications.

The main task of the Open Research Data repository is to provide data from the universities that make up the Bridge of Data consortium. The very process of designing the platform and its elements, from the metadata schema to the content management tools, was long and tedious. However, it resulted from the close and effective cooperation of various teams (Library, IT, researchers) at all three universities.

Since the repository collected data from several different disciplines, the challenge was to provide both appropriate functions and a universal metadata description.

The metadata should be human- and machine-readable and also compatible with commonly used metadata standards.

Selecting the standards that will be appropriate for dataset collections and fulfil the FAIR principles, it is a weighty and challenging decision. For scientific publications, mostwiedzy.pl already supported Dublin Core and Highwire Press tags. Additionally, to ensure the project's compatibility with 5 stars Open Data, each object is described by schema.org with JSON-LD formatting. The creators of the datasets are encouraged to link the datasets to other objects existing in the Bridge of Knowledge Portal, such as publications, science projects, science teams, laboratories or other datasets. Moreover, research data may be grouped into series. All the mentioned links are also represented in the descriptive, JSON-LD based metadata. These links to other data sources, combined with other features and rules, such as using non-proprietary file format, assigning unchangeable DOI and URLs to each dataset fulfill the requirements of 5 star Open Data in terms of metadata descriptions of datasets.

As metadata is essential for efficiently storing, sorting, retrieving, sharing, and linking scientific data, the team decided to use the DDI (Data Documentation Initiative) standard for the first level to ensure

the description of granular levels of resources of the metadata. Due to the wide range of disciplines covered by the project, from humanities, social sciences, technical and engineering, to medical science, the team was looking for a standard that best reflects the needs and assumptions. The DDI standard is quite general, flexible, and more accessible for all disciplines and broader communities than others. Besides this, the DDI is more interoperable than other standards, which will result in better indexing of the provided datasets in various search engines and data hubs, increasing awareness of their presence and availability. The second level of metadata was subject-specific and more constrained to ensure that scientific objects were more findable and reusable (for example, the INSPIRE standard for GIS data).

The data repository has a hierarchical structure that allows, e.g. research teams to assign a specific collection of datasets to particular projects and then a sub-collection to different research objects such as individual scholars, publications, software or images.

As research data may result from various types of projects, be linked into groups or series, and be related to publications, all functionalities have been added to the repository that allows for linking and relating individual resources to each other. It is also possible to create dataset versions. All these functionalities make it possible to adjust the dataset collection to individual needs and improve their searchability.

Based on the experience of the Competence Center and comments from users, the repository is constantly being improved. Recently added, among others citation styles and their export, test version of the bibliography, the possibility of thanking the author for sharing the file with the publication and linking the publication with other resources (funding source, project, research data, etc.).

The Database of Copyright and Open Access Self-Archiving Policies of Polish Scientific Journals collects and analyses the publishing policy in Open Access, copyright management, and the content of publications by both authors and users. According to the creators' concept, the database was to fulfil a function similar to SHERPA / RoMEO. Besides this, the database of publishing policies substantially complements the information in SHERPA / RoMEO, comprehensively collecting information about Polish journals, both those indexed in it and those that are not yet registered there. Moreover, the database extends the information relevant to Polish users by providing data on the current scoring and publication model. The collected information is presented in the form of a journal profile, the description of which consists of several sections.

The database recipients are primarily members of the research community: researchers, doctoral students, and students. However, a database is also a convenient tool that allows academic librarians, especially those involved in developing institutional repositories, to define the Green Open Access rules.

Under the Bridge of Data project, the team from the Medical University of Gdańsk undertook the creation of the Digital Tissue and Cell Atlas and the Virtual Microscope with the cooperation of the Gdańsk University of Technology and CI TASK IT Center. Experts in pathomorphology gathered carefully selected histological and cytological specimens. Samples were processed, stained and scanned in the MUG's Department of Medical Laboratory Diagnostics (DMLD). Virtual Microscope is an application made available as part of the platform. Its role is to provide comfortable and quick access to a collection of digital microscopic images stored in a data repository. The shared images are characterised by very high quality and resolution.

With the help of a virtual microscope, you can view the deposited images, which were created due to the preparation and digitisation of thousands of tissues and cells. It is an innovative way of presenting and broader use of research data collected in the repository.

The main research problem of the poster is presenting the functionality of the mostwiedzy.pl platform, as well as comparing it with other solutions of this type

References:

Gdańsk University of Technology (2021) MOST Wiedzy - your knowledge portal. Available at: <https://mostwiedzy.pl/en/> (Accessed: 8 March 2022).

Kokot-Kanikuła, K., Szymik, A. (2021). Zadania i kompetencje bibliotekarzy w zespole ds. polityk wydawniczych polskich czasopism naukowych, Biuletyn EBIB, pp. 1–8. DOI: 10.5281/zenodo.5004729.

Krawczyk, H., Lubomski, P. (2017) Multidisciplinary Open System Transferring Knowledge for R2B Development, TASK Quarterly, 21(4), pp. 299–308. DOI: 10.17466/tq2017/21.4/c.

Hausenblas M. (2015) 5-star Open Data Available at: (Accessed: 8 March 2022).

RDF Working Group (2014) Resource Description Framework (RDF) Available at: <https://www.w3.org/RDF/> (Accessed: 8 March 2022).

Science D, Hahnel M, McIntosh LD, et al. (2020) The State of Open Data 2020. Digital Science. DOI: 10.6084/m9.figshare.13227875.v2.

Wałek, A. (2018). Is data management a new 'digitisation'? A change of the role of librarians in the context of changing academic libraries' tasks, IFLA WLIC 2018 – Kuala Lumpur, Malaysia – Transform Libraries, Transform Societies, pp. 1–11. Available at: https://mostwiedzy.pl/pl/publication/download/1/is-data-management-a-new-digitisation-a-change-of-the-role-of-librarians-in-the-context-of-changing-_30244.pdf

Wałek, A., Kokot, K.. (2020). Untying a tight knot - Database of Copyright and Open Access Self-Archiving Policies of Polish Scientific Journals. Open-Access-Tage 2020 (OAT2020), Bielefeld, Germany. Zenodo. DOI: 10.5281/zenodo.4018894

Wałek, A., Lubomski, P. (2017). The Bridge to Knowledge – Open Access to Scientific Research Results on Multidisciplinary Open System Transferring Knowledge Platform, TASK Quarterly, 21(4), pp. 333–342. DOI: 10.17466/tq2017/21.4/e.

Wałek, A., Szuflińska-Żurawska, M. (2017). The Bridge of Knowledge – defining and implementing national Open Access policy by Polish universities. Gdansk University of Technology perspective. IFLA World Library and Information Congress, 83rd IFLA General Conference and Assembly : Libraries. Solidarity. Society, 1-8 Available at: https://mostwiedzy.pl/pl/publication/download/1/the-bridge-of-knowledge-defining-and-implementing-national-open-access-policy-by-polish-universities_16116.pdf

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>