

## RDA/WDS Publishing Data Services WG

### What is the problem?

When research data lives in isolation, it can be difficult to find, access, and make sense of it. By creating bi-directional links between data and the literature, the visibility of both data and publications is increased, and data can be interpreted in the right context much more easily. While there are a number of organizations that track and expose such links, this is usually done in isolation and through bilateral arrangements. The problem is that there is no overarching standard or service to combine links from different sources into a common resource, resulting in a very patchy landscape of isolated collections of links, which very much limits their value and usefulness for researchers.

### What is the goal?

Building on pre-existing components and international initiatives, the goal of this Working Group is to deliver a (prototype) universal interlinking service between data and the scientific literature. By combining inputs from different organizations, we want to move from a plethora of (mostly) bilateral arrangements to a common “one-for-all” service model infrastructure. While the realization of an interlinking service is the primary objective, we’re also interested to learn from our experiences and explore if there are other processes or workflows in the research data landscape that can be lifted from a set of bilateral arrangements to a common standard.

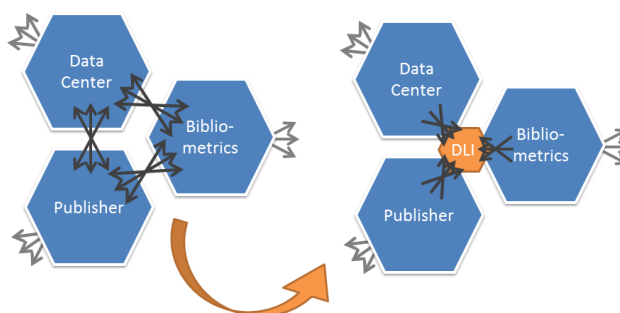


Figure 1- Moving from an array of bilateral arrangements towards common standards and one-for-all services

For more information, please also see our Case Statement at <https://www.rda-alliance.org/filedepot?cid=114&fid=239>

### What is the solution?

The WG is working to create a universal interlinking service by working with a broad range of stakeholders – including data centers, publishers, providers of bibliographical information, and research



institutes. The WG – which has a dual mandate from RDA and ICSU World Data System – operates in close coordination and cooperation with CrossRef, DataCite, ORCID, OpenAIRE, the International Association of STM Publishers, the National Data Service, and RMap. Getting all those stakeholders on board is essential to collect a broad and representative set of data-literature links, expose the links in a way that matches real-life use cases, and ensure broad adoption further down the line.

The WG currently has close to 1 million links from various contributors. In a synergistic collaboration with the OpenAIRE2020 project, these links have been loaded, normalized, and completed in a single database using a custom-made, flexible metadata schema. As we aim to further expand coverage of links, and create service API's for users to consume the information, the metadata schema can be further refined on the basis of new learnings and feedback from the community.

## **What is the impact?**

The data-literature interlinking service will bring benefits to all key stakeholders in the data publishing landscape – including data repositories, journal publishers, providers of bibliographic services, funding bodies, research institutes, and researchers themselves. For many organizations, using literature-data links will become a much more efficient (scalable), reliable (quality), and rewarding process. This will allow them to extend their offerings, or develop new ones, for the benefit of research and researchers.

## **When can we use this?**

Working in close collaboration with the Publishing Data Services WG, the OpenAIRE team has developed a prototype portal that allows you to inspect the corpus of links currently in the system. This portal is available at <http://dliservice.research-infrastructures.eu/>. A full-fledged demonstrator tool is expected to be delivered at the 7<sup>th</sup> RDA Plenary meeting in spring 2016. After that work will be continuing to further develop this into an industry-strength, reliable, comprehensive, and high-quality service.

## **What is RDA?**

To address the growing global need for data infrastructure, the Research Data Alliance (RDA) was planned and launched in March 2013 by the European Commission, the US National Science Foundation and the Australian Government through the Australian National Data Service (ANDS) as an international, community-powered organization. RDA's vision is of researchers and innovators openly sharing data across technologies, disciplines, and countries to address the grand challenges of society. Experts work together, in a self-forming and completely voluntary manner, in focused Working groups or exploratory Interest groups. Information on [www.rd-alliance.org](http://www.rd-alliance.org)

## **What is WDS?**

The ICSU World Data System (ICSU-WDS) builds on the 50-year legacy of the World Data Centres



and Federation of Astronomical and Geophysical data analysis Services established by the International Council of Science (ICSU) to manage data generated by the International Geophysical Year (1957–58). These bodies were later disbanded by the ICSU General Assembly in 2008 and replaced by ICSU-WDS in 2009. ICSU-WDS aims at facilitating scientific research endeavours by coordinating trusted data services for the provision, use, and preservation of relevant datasets.

To fulfil its remit, ICSU-WDS is striving to build worldwide ‘communities of excellence’ for scientific data services by certifying Member Organizations—holders and providers of data or data products—from wide ranging fields using internationally recognized standards. WDS Members are then the building blocks of a searchable common infrastructure with which to form a data system that is both interoperable and distributed. [www.icsu-wds.org](http://www.icsu-wds.org)