1st Conference on Research Data Infrastructure

Poster Presentations II

https://doi.org/10.52825/CoRDI.v1i.357

Published: 07 Sept. 2023

Introducing JuOSC

Monica Gonzalez-Marquez[1] and Ines Schmahl[1] [https://orcid.org/0009-0003-5878-3706]

[1] Forschungszentrum Jülich

Abstract. Where does one go to find information on Open Science? At the moment, the literature is scattered across the various fora that comprise scientific discourse. An inquiry about any topic in Open Science, for example, how to manage a particular type of data in chemistry or what the current stance of the DFG or the NSF might be on data on photovoltaics or the relationship between Kuhn and Open Science, invariably require a mixture of asking colleagues or conducting random searches in academic indexing services, with varying degrees of success. Part of the problem is that so much of Open Science discourse has occurred in fora made possible by the advent of the internet. Between blog posts, twitter threads, talks and podcasts, there is a wealth of information that needs to be organized and archived to be most useful to the scientific community. There are efforts to organize some of this information in various online databases, and they are laudable, but to serve as long-term tools, they require 1. orchestration to gather and curate materials as a community effort, 2. long-term infrastructure to store copies of the materials, and 3. the expertise of librarians to manage such a collection.

Keywords: Special Collection, Open Science as a discipline

1. Our solution

We have begun to build the Jülich Open Science Collection (JuOSC) as the hub for all literature related to Open Science as a discipline. The collection will include traditional publication outputs (books and papers) as well as gray literature (blog posts, preprints, theses, etc.) and ephemera (twitter threads, podcasts, etc.). Because the collection is specifically for Open Science as a discipline, by definition all materials it contains will be freely available to everyone at no cost, as well as available on demand. As such, we are ensuring that all materials are either open access or available through a cc-by license. To ensure access on demand, copies of all materials will be stored on the JuSER infrastructure at Jülich. We will also construct mirrors at Zenodo and at the Internet Archive.

2. Implementation

JuOSC will be stored on the Juelich Shared Electronic Resources (JuSER), the publications database developed for Forschungszentrum Jülich. The Central Library maintains and develops the database based on the open source software Invenio. The use of open source software ensures that modifications can be made by developers as needed by stakeholders, including the open science community, now and in the future. At present, JuSER provides a list of document types. The complete workflow covers the collection of potential literature, the curation, the importing of metadata, archiving the files in JuSER and publishing to make the open science literature accessible worldwide to everyone. In addition, archiving workflows for all materials, including audio and visual records, are being developed in collaboration with experts from the LZA-TREFF-KOELN, a network of people dedicated to addressing issues relating to long-term archiving. Because we are committed to creating a special collection that

serves the Open Science community *in collaboration with* the Open Science community, curation of materials is being done with the participation of the Open Science community. Specifically, this means that previously developed databases on Open Science literature will be integrated into JuOSC. Several such databases have already been identified, e.g. the Open Research, Open Science, Open Scholarship database developed by Erzsébet Tóth-Czifra and colleagues, and collaboration agreements are in development for others, i.e. Framework for Open and Reproducible Research Training (FORRT). We are also creating tools to facilitate input from the Open Science community such as suggestions for categories of literature and for particular items using a Google Form and hashtags for use on social media, i.e. #JuOSC. For all steps we aspire to an intensive exchange with the community in general and with any and all groups, experts and stakeholders who are willing to support our efforts. As to legal access to all materials we will archive, as stated above, all will be either open access or cc-by. Any materials that do not meet those requirements and for which an alternative is not available, i.e. a preprint instead of a paywalled paper, will be excluded. Our firm policy is that if an item is contained in our database, it will be freely available on demand, no exceptions.

3. Awareness

The purpose of JuOSC is not only to serve as information infrastructure, but to be actively used to build the open science community, both within Jülich and beyond. The same as with all disciplines, when scholars are first exposed to a new idea they begin with a literature search.The overarching goal of JuOSC is to facilitate education and exposure to new and established ideas and practices surrounding Open Science for all stakeholders. Our hope is that JuOSC will help the stakeholders at large to align their strategies and use limited time and resources more productively.

4. Conclusion

Our immediate aim is to develop a prototype by the end of the year, and to start using JuOSC to organize events and teach courses and workshops at the Forschungszentrum Jülich. Our approach will hopefully be useful to other institutions, and help strengthen the Open Science community by increasing awareness of and adoption of Open Science practices.

Competing interests

All authors declare there are no competing interests.