

Data FAIRness at the University of Ljubljana, Slovenia

Abstract. The activities of the Slovenian Social Science Data Archives (ADP) are presented for the implementation of open data policies at PhD school of University of Ljubljana. The activities are part of the Archive's broader efforts to promote open science and a culture of data sharing that underpins transparency and advancement of knowledge. ADP draws on best practices and domain standards developed within CESSDA. Through the University of Ljubljana example, we aim to show the obstacles at several steps of the implementation of the policy and the benefits of the collaboration between the archive and the university.

1 Introduction

The Social Science Data Archives is Slovenia's data infrastructure for the social sciences. The primary mission of the archive is to select high-quality social science research data relevant to Slovenia for long-term preservation and to provide access to data to different audiences for re-use. The archive was established in 1997 at the Faculty of Social Sciences, University of Ljubljana. Today, the archive contains more than 600 social science research data accessible. The following subject areas are well represented: demography (population, vital statistics, and censuses), economics, education, health, history, housing and land use planning, information and communication, labour and employment, law, crime and legal systems, media, communication and language, natural environment, politics, psychology, science and technology, social stratification and groupings, social welfare policy and systems, society and culture, trade, industry and markets, transport, travel and mobility. The earliest research data went back to the 1960s. The archive holding is a rich resource for social scientists and interdisciplinary research, as well as for educational usage.

The research data package published in the ADP includes data and its' context metadata description, together with the supporting documentation, which enables understanding of the process of data creation and enable verification of analysis results. The data are available for re-use by researchers, educators, students, government departments, civil society organisations, journalists and other interested individuals, communities and organisations. The challenge ahead is how to integrate citizen science research into our work. The archive has around 1000 registered users per year, of which around 90% use the data for educational purposes and 10% for scientific purposes. On average, users use around 150 surveys per year for secondary data analysis.

2 Promoting open data

In 2010-2013, ADP implemented the Open Data project for the Slovenian Ministry of Education and Science, which resulted in two policy proposals: a draft proposal for a policy on the management of publicly funded research data and an Action Plan for the establishment of a system of open access to publicly funded research data [1]. As part of the project, ADP conducted several in-depth interviews with researchers, managers, librarians, etc. from different scientific disciplines to find out about their experience of managing data, what services and infrastructure they are familiar with, what they would need. Through this project, we have established contacts with various stakeholders in the field of science, held several consultations and workshops with science funders, chairs, research infrastructure providers, librarians, etc., and identified opportunities for the future development of services to support open data [2].

After the end of the project, we continued our activities to promote data management planning, as it became clear that researchers have many concerns about sharing data. Firstly, because they do not have the skills to prepare data in time for archiving and publishing in a suitable repository and secondly, because they are not aware of the services available to support them at every step of the data lifecycle.

Discussions with various stakeholders also revealed that older generations in particular have many concerns. While established researchers were generally reluctant to change their learned practices of working with data, many of them were in favour of introducing change for new generations just starting out on the scientific path.

2.1 FAIR principles and Data Management Plan

Since the end of the Open Data project described above, ADP has been working hard to promote the concepts of "open access", "open data", "open science", "data management plan", "FAIR principles", etc. This has been supported by the CESSDA collaboration, where the training team has been working on training materials, e.g. the DMEG guide. We deliver data training both nationally and internationally.

We work with a wide range of stakeholders. We have presented the importance of a systematic data management regime to the heads of research organisations, heads of research centres, and higher education institutions. We have delivered several training sessions for librarians on the role of data steward. We work with educators and present the ADP database in lectures. We responded to an invitation from the Faculty of Social Sciences, University of Ljubljana to present Research Data Management following FAIR principles to PhD students, already from 2015.

3 Data management plan for doctoral students at the University of Ljubljana

In 2017-2021, in cooperation with the management of the University of Ljubljana, we prepared a proposal for the inclusion of the data management planning into the Rules and Regulations for Doctoral Studies at the University of Ljubljana. The regulation made the preparation of a DMP and the publication of data prior to the defence of the doctoral dissertation mandatory for doctoral candidates [3]. The Vice-

Rector for Research was helpful during the drafting phase, expressing his understanding and sympathy for open access and open data. His research and publication experience abroad confirmed the necessity to reformulate the university's policies and update the rules for future generations. Despite some concerns from senators, the rules were finally adopted.

The Data Management Plan is mentioned in several places in the Regulation (in force from 1 October 2021). Its definition is given as follows in the Research data management Article 50:

Research data generated and collected for the needs of a doctoral dissertation must be published or otherwise accessible in such a way that allows their visibility, access, interoperability and the possibility of renewed evaluation and use.

The doctoral candidate shall submit research data to a data repository, data centre or research data archive, which shall satisfy the principle of verifiability, transparency and open science. As a priority the research data shall be sent to the sectoral national or international data centres intended for specific types of data, or to the UL Repository.

The doctoral dissertation shall state where the data are accessible and how they can be accessed. Exceptions in the sharing of data shall be justified where they involve personal or sensitive data, or where there are reasons for protecting intellectual property or for non-disclosure of vulnerable areas, groups or species. In the case of implementing justified exceptions to data sharing, the doctoral candidate shall ensure an appropriate method of protecting the data and limiting access to such data in agreement with the data centre. In this case at least freely available metadata must be generated for the catalogue of the data centre, so as to indicate clearly where and under what conditions the research data are accessible.

It also defines roles and responsibilities and sets out the process. Students are required to prepare a first draft of the research data management plan already for the doctoral dissertation proposal (Article 36). At the Presentation of the results of research work before completion of the doctoral dissertation candidate presents also the updated research data management plan (Article 43). A doctoral dissertation shall include: the final draft of the research data management plan (Article 45).

Supervisors and co-supervisors shall advise the doctoral candidate in the planning of research and research data management (Article 33). In its assessment, the DSC shall also determine the suitability of the draft of the research data management plan (Article 40).

3.1 Implementation of the Regulations into practice - Supporting the university

Following the adoption of the Regulations, the Head of the Doctoral School invited the ADP to prepare a template RDM plan to be used by doctoral students, their

supervisors and the Committee. The need for such a template is particularly in scientific fields where data services are not yet available in Slovenia and researchers are not familiar with the services available internationally. As it was necessary to draw up a template for all disciplines,

ADP proposed to use the Science Europe template. ADP used the following criteria to argument the proposal. The template has been developed internationally by experts from different disciplines with the aim of making it suitable for a diverse and heterogeneous research landscape. The template is internationally recognised and recommended by Science Europe in order to standardise the way researchers plan the handling of their research data, avoiding a flood of different templates that may be tailored to specific needs but are not aligned with international common standards. In addition to the RDM Guidance for Researchers, the developers have made available the RDM Guidance for Organisations and RDM Guidance for Reviewers. This is particularly important for areas where practices are less developed and data experts are unavailable or scarce. ADP has translated the Science Europe's Template into Slovenian. The university, after consideration, rejected the proposal and prepared a shortened version of their own template. The basic guideline was to keep the template as short as possible and as burdensome as possible for the stakeholders.

The university asked ADP, as data expert, for support in translating the template into English. ADP is also providing training support. The first one at university level is organised in spring 2022, followed by a series of events on the following topics: Fundamentals of research data management for PhD students, FAIR research data, workshop on Preparing a Research Data Management Plan for PhD students at UL, Collection and preservation of research data - in the course of the research and in the long term (file formats, licensing, repositories, metadata), Workshop on examples of research data management in the context of doctoral studies at UL, Ethics and integrity in research [4].

3.2 Implementation of the Regulations into practice - support for PhD students

Since 2018, the ADP has been contacted by individual PhD students seeking assistance in preparing a data management plan. Until the implementation of the Regulations, these were young researchers who had participated in H2020 projects and had committed themselves to preparing a DMP. They first sought help at the university, from where they were referred to the ADP. Advice for students from other fields, e.g. wood science and technology, has been a particular challenge. In such cases, ADP explained the terminology of the DMP and asked the questions needed to better understand the process of working with the data itself, but was not able to advise on domain-specific issues, as our expertise is in the social sciences.

We have also advised several PhD students in the social sciences. It can be concluded that young researchers show a willingness to prepare a DMP and a satisfaction that support and advice is available. In addition to individual counselling, once a year we also hold a lecture for PhD students in which we present the basics of DMPs in the social sciences and the DMP template we have developed at CESSDA, which is tailored for the social sciences. Students are advised to use the domain template as this will help them to meet the domain standards and good practices that

are established internationally. To support this, the CESSDA Data Management Expert Guide includes 15 hours of online content and is intended as a Self-study for researchers [5].

4 Conclusion

Bringing together the different stakeholders in the research community is key to achieving the goal of open data. In our case, the decisive factor was the sympathy of the university management, which strongly advocated the importance of enforcing research data management policies and introduced them for early career researchers. Although there were many concerns, in particular about the additional commitment and effort and the underdeveloped support in certain scientific fields, the policy was enforced.

A similar experience can be seen at the Faculty of Social Sciences, University of Ljubljana, which went a step further and in 2020 introduced a data management rules for all employees who collect data in research projects. Again, there have been many concerns about the support that researchers will need. The fact that the ADP is based at the Faculty and will be available to help researchers in any way was seen as an advantage.

References

1. Štebe, J., Bezjak, S., & Lužar, S. (2013). Odprti podatki. URN:NBN:SI:DOC-US3XRRB2 from <http://www.dlib.si>.
2. Project description Odprti podatki (2010 - 2013). Retrieved from: <https://www.adp.fdv.uni-lj.si/spoznaj/projekti/pretekli/>.
3. Doctoral School. University of Ljubljana. Rules and Procedures. Retrieved from: https://www.uni-lj.si/study/doctoral_school/rules/.
4. Serija dogodkov Doktorske šole UL o ravnanju z raziskovalnimi podatki doktorandov v okviru doktorskega študija na UL. Retrieved from: https://www.uni-lj.si/raziskovalno_in_razvojno_delo/odprta_znanost/prihodnji_dogodek/.
5. CESSDA Training Team (2017 - 2020). CESSDA Data Management Expert Guide. Bergen, Norway: CESSDA ERIC. Retrieved from <https://www.cessda.eu/DMGuide>.