

ANNA WAŁEK
Gdańsk University of Technology Library
e-mail: anna.walek@pg.edu.pl
ORCID 0000-0001-8782-013X

DATA STEWARDSHIP IN POLAND – CURRICULA AND CAREERS PATHS – THE VERY BEGINNING OF THE NEW PROFESSION



Dr Anna Wałek, President of IATUL – International Association of University Libraries, director of the Gdańsk University of Technology Library. An experienced library manager, an expert in the field of Open Science, and organization and management of a scientific library. She conducts scientific research in data management in various scientific disciplines, metadata for research data, and data management support services, incl. defining the roles and tasks of data stewards and librarians. A graduate of the Institute of Information and Library Science at the University of Wrocław, where she obtained a PhD in Library and Information Science in 2013. Since February 2023, expert and Coordinator of the Eastern Hub as part of the Focusing on Open, Collaboration and Useful Science (EOSC Focus) project in the structures of the National Science Centre and the main specialist in the Data Processing Laboratory at the CYFRONET Academic Computer Center of the AGH University of Science and Technology in Krakow in the EOSC FUTURE project (Integration and consolidation of the existing pan-European access mechanism to public research infrastructures and commercial services through the EOSC Portal). Since March 2023 represents Poland as part of the Council for National Open Science Coordination (CoNOSC). In the years 2007-2016 she worked at the Wrocław University of Science and Technology. Since January 2017, she has been the Director of the Library of the Gdańsk University of Technology. Member of the Board of Directors (since 2018) and President (since 2022) of IATUL (International Association of University Libraries), member of the SPARC Europe Board of Direc-

tors (2018-2023), member of Task Force Open Science (2017-2021) and Task Force Openness of Science & Technology (2022-2023) of CESAER (Conference of European Schools for Advanced Engineering Education and Research), as well as the Research Data Alliance. Representative of the Gdańsk University of Technology in the EOSC Association (2021-2023) and a member of the EOSC Association Data Stewardship Curricula and Career Paths Task Force. Originator and coordinator of the national working group of data stewards - Polish Data Stewardship Competence Centers Implementation Network (DSCC-IN PL) established in December 2019 as part of the GO FAIR initiative (DSCC-IN national chapter). Author of scientific publications, organiser and member of the program and organising committees of national and international scientific conferences.

KEYWORDS: Research Data Management. Data Stewardship. Academic libraries. data steward.

ABSTRACT: Thesis/Objective – The profession of data steward is not fully defined. In particular, in Polish scientific institutions, where even the name of this position is still hard to find, this is a new phenomenon that requires a deeper analysis. It is difficult to specify not only the definition of this profession, or rather the role, but also the scope of tasks and required competencies. All this depends on the role of the data steward in the unit, at the university or in the research team. **The aim** of the article is to present the origins of the creation of a new specialisation in Poland, which is the role of a data steward. In addition, the forms of education for data stewards in Poland are discussed, as well as initiatives and projects aimed at defining the professional profile, qualifications and career paths of data stewards. **Research methods:** The critical literature review was used to analyse the content of foreign and Polish LIS literature published in the years 2018-2022. Also, official documents and websites, project descriptions, working group's, and Task Force's working papers were analysed. In addition, a survey was carried out and analysed among the members of the Polish data stewards working group (DSCC-IN PL) on organisational structures and positions dedicated to research data management existing at Polish universities. **Results and conclusions** – The first data stewards appeared in libraries. This resulted directly from the policy of the National Science Centre (NCN) but above all also from the fact that the first unit providing support for researchers in the field of data management – the Open Science Competence Center – was established in the Library of the Gdańsk University of Technology (Gdańsk Tech). The Gdańsk Tech Library has established numerous contacts with international networks and organizations, which resulted in active cooperation and placing it in international structures. In particular as a national chapter of the GO-FAIR network (DSCC-IN PL: Data Stewardship Competence Centers Implementation Network – Poland). This initiated the creation of the Polish working group of data stewards, which is now the foundation of the nascent profession and provides a field for the exchange of knowledge and experience. The lack of educational programs at university or in any other form of vocational education has been filled in a sense through the training activities of the Competence Center and the only Data Steward School in Poland. It was only in 2022 that the increased activity of the National Science Center and the Ministry of Educa-

tion and Science began. In conjunction with the work on the national policy for open data as well as activities within the European Open Science Cloud, a coherent environment began to emerge for the development of a new specialization at the level of Polish universities and scientific institutions.

Research data are becoming increasingly important for researchers who appreciate the benefits of sharing data or are obliged by funders of their research projects to provide open access to the research results. In addition, researchers are beginning to see the advantage of sharing and reusing data sets. More and more funding bodies have defined requirements that Data Management Plan (DMP) for managing the data produced during a research project should be included in grant proposals. These plans show how research data will be collected, organized, managed, shared and preserved during the project and after its completion. What such a plan looks like depends on the specificity and conditions of the project. In general, DMP requires a description of the data that will be produced or used, information about the formats, metadata standards that will be used to store and organize data, the “how” and “where” of the data storage and accessibility (Wałek, 2020). The raising interest in reusing and sharing research data has been confirmed by the results of the research conducted by Digital Science-Figshare. *The State of Open Data Report*, published every year by Science Digital, examined global attitudes toward Open Research Data (ORD) (Science Digital, 2022). Since research data management (RDM) has become the norm in most research projects and providing access to data is becoming more and more common, it is necessary to provide researchers with appropriately qualified support in the form of people and teams dedicated to providing services in the field of data management.

Research data services are defined as services that address the full data lifecycle, including the DMP, digital curation (selection, preservation, maintenance, and archiving), and metadata creation and conversion (Tenopir, Sandusky, Allard and Birch, 2012).

The data steward position has several different definitions, although there is still no official and uniform description of what this profession is really about. An attempt to answer this question was made by a team established within the EOSC Association as a Task Force Data Stewardship, curriculum and career paths¹. Members of the subgroup examining career paths have prepared a state-of-art document in which they summarize the findings so far in this regard (Kalová, T. et. al., 2023).

According to the findings of the Task Force, two main definitions of data steward are adopted for the purposes of this publication.

¹ Task Force Data stewardship, curriculum and career paths <https://www.eosc.eu/advisory-groups/data-stewardship-curricula-and-career-paths>

“Data steward” definition 1: A person (or a group of people) to ensures professional and careful treatment of data throughout all stages of a research process responsible for planning and executing all actions on digital data before, during and after a research project, with the aim of optimising the usability, reusability and reproducibility of the resulting data (OECD, 2020).

“Data steward” definition 2: A person responsible for keeping the quality, integrity, and access arrangements of data and metadata in a manner that is consistent with applicable law, institutional policy, and individual permissions. Data stewardship implies professional and careful treatment of data throughout all stages of a research process. A data steward aims at guaranteeing that data is appropriately treated at all stages of the research cycle (i.e., design, collection, processing, analysis, preservation, data sharing and reuse) (Jetten et al.,2021).

Describing the current state of the art in relation to the career paths of data stewards, the authors of the document highlighted the challenges and recommendations as to the next steps in relation to this issue.

The authors highlighted that the overall objective career paths work stream was to produce a report on how data steward roles map in an international, disciplinary and institutional context. The goal was to build on existing initiatives, projects and professional networks and issue recommendations for data steward career paths. However, the initial discussions have led the team to the conclusion that “the current panorama on Data Stewardship is too scattered, both in terms of training and careers. While many research institutions are hiring data stewards, there does not currently seem to be enough consensus as to how these roles map onto existing professional positions within institutions. With most organisations focusing on the training and embedding of new Data Stewards in the institution, developing sustainable long-term career paths is currently not being addressed in a systematic way. Differences in approach among various countries and types of institutions (such as universities, libraries, scientific IT services, research infrastructures) are also prevalent” (Kalová, T. et. al., 2023). Another finding of the research was that the relationship between data stewardship and research activities is also blurred. It appeared that researchers who fulfil the role of a data steward as part of their job may have different career path perspectives and issues with respect to those who are hired solely as data stewards. Finally, the authors highlighted that data stewardship roles are simply too recent in many organisations and countries to allow for a diachronic investigation of the career paths of those hired in these positions.

As the result of the team’s work was to be a report, the authors of the document indicated that the result of their work is only the beginning of further research and findings, requiring supplementation. That’s why the

document is called “state of art”. In light of the above considerations, the possibility of producing a comprehensive report on the current panorama with a focus on career paths in the opinion of the authors was deemed premature. For the long-term objectives, the document provides recommendations for (a) gathering additional information on Data Steward’s careers (b) monitoring the development of this professional role at the European level and (c) facilitating the creation of a professional network at the European level.

THE ROLE OF THE ACADEMIC LIBRARY

Academic libraries have always been pioneers in implementing openness to research and creating the necessary tools (platforms and repositories) as well as providing training for researchers. Open Access policy implementation, including both access to research publications and data, is accelerating. That is why the role of academic libraries and academic librarians has become crucial. Some new specialisations in librarianship have been developed, e.g. a data librarian who is responsible not only for academic staff training sessions on Open Research Data (ORD) but also for assistance for research teams in the field of data management and data curation. It seems more and more certain, based on past experience, that academic libraries will be responsible for coordinating the work of data stewards responsible for supporting the process of research data producing, managing and archiving at universities and in research teams (Walek, 2020).

OPEN SCIENCE COMPETENCE CENTER

The idea of creating the Open Science Competence Center (OSCC) at the Gdańsk University of Technology (Gdańsk Tech) Library arose from the global trend of data stewardship and previous experiences. It has been revealed that researchers’ knowledge of Open Science has gaps, especially those related to copyright and research data management issues. To resolve these practical difficulties faced by researchers, a support team that includes data specialists and librarians has been organised and managed at the Gdańsk Tech Library. Part of the team is recognised as data stewards – a relatively new position in the Polish academic landscape. Since the beginning, a data steward is usually seen as a disciplinary expert with diverse knowledge and experience in research data management practices. Another critical role played by OSCC members is that of data support librarians. Their essential characteristic is supporting researchers at multiple stages of the data life cycle, both during the research process and during the curation process (Walek, 2018).

The Open Science Competence Center at the Gdańsk Tech Library serves as a contact point that offers help and support services for sharing research outputs. The OSCC provides various training, consultancies, and other events promoting the idea of opening science and trying to popularise its benefits such as broader collaboration, increasing usage and citations of scientific articles and data, faster impact, and greater public engagement.

Different types of training are offered to scientific staff and students, such as tailored training (e.g. concerning the scientific discipline), face-to-face consultations, and online webinars focused on the overview of open research data, Data Management Plans (national and EU grant applications), legal support (data licensing, data protection, reusing data), using the data repository (depositing datasets), FAIR meta(data) and Plan S (implications and requirements). OSCC team members also validate the metadata descriptions and data formats entered into the data repository (Bridge of Knowledge²). Another main challenge for the OSCC was considering the differences between scientific disciplines and their different scholarly communication practices regarding sharing scientific output to provide complex support with RDM. Each field has its scholarly communication practices and should be treated individually concerning maintaining the regulations and procedures following the research data issues (Wałek, Nowacki, Lubomski, 2022).

The Bridge of Data project was established to support researchers in their activities regarding different Open Science layers. The project was unique in Central and Eastern Europe. It involved three Pomeranian universities: Gdańsk University of Technology (leader), the University of Gdańsk, and the Medical University of Gdańsk. The Bridge of Data project (Bridge of Data – Multidisciplinary Open System Transferring Knowledge. Stage II Open Research Data) started in October 2018 and finished in June 2022. It was focused on implementing the Open Research Data repository. It also provided system modules that support organisers of scientific conferences and editorial offices of scientific journals in their processes. The project's key objective was to create substantive support in the Open Science Competence Center at the Gdańsk Tech Library. The project was co-financed by the European Regional Development Fund within the Digital Poland Operational Program for 2014–2020.

The Competence Center was created immediately after the Bridge of Data project was launched – in the autumn of 2018. Being the only university in Poland with a team of professionals who can support research teams in the preparation of DMPs for the needs of project applications,

² Open Research Data Catalogue – Bridge of Knowledge Platform – mostwiedzy.pl

Gdańsk University of Technology has grown into the role of a leader in creating services and standards for RDM in Poland.

DATA STEWARDS AT POLISH SCIENTIFIC INSTITUTIONS

The official inauguration of the Polish Working Group (Polish Chapter) of the Data Stewardship Competence Centers Implementation Network (DSCC-IN PL), one of the implementation networks of the GO-FAIR Initiative³ took place on 3rd December 2021. The workshop was an accompanying event for the 5th edition of the Pomeranian Open Science Conference held in Gdańsk on 2nd December 2021. The conference, hosted by the Gdańsk University of Technology, gathered more than 120 attendees and 18 speakers from all over the country and guests from abroad. The event and its theme, “Services based on Collaboration”, focused on obtaining practical experience in Open Science and Research Data Management. The DSCC-IN PL workshop welcomed almost 60 attendees interested in establishing the Polish Working Group. The Working Group and its future possible role and activities were presented by the DSCC-IN Polish chapter coordinator, dr Anna Wałek, the Director of the Gdańsk Tech Library. Participants also enjoyed the pre-recorded guest talk by Erik Schultes, GO FAIR International Science Coordinator, who encouraged them to engage and collaborate. The OSCC Team, which coordinates the activities of the group, presented the dedicated website, which contains the mission of the Group, its structure proposal, tools description and a piece of information on how to become its member. The DSCC-IN PL presentation met great interest among the Polish librarians and researchers who attended the workshop (Dudek, 2021). The first Working Group official online meeting was scheduled for the end of January 2022, and since then the group is meeting every month. At the end of 2022, the team had 102 members from 42 institutions.

In 2022, 9 online meetings took place on the MS Teams platform. The meetings served to exchange information and good practices in the field of Open Science and RDM between group members. Each of the meetings was devoted to current affairs and one of the main topics.

Main topics of meetings in 2022:

- February – organizational meeting
- March – Meeting with Michał Nowacki (CUI, Gdańsk Tech) regarding the MOST Wiedzy platform
- April – Support for scientists in the management of research data at the Lublin University of Technology (dr Katarzyna Weinper)
- May – Subject related to Open Science at Doctoral Schools

³ GO-FAIR Initiative <https://www.go-fair.org/>

- June - Meeting with Wojciech Fenrich (ICM UW) on RepOD
- September – Archiving and sharing research data
- October - The role of Rector’s representatives for Open Science (Dr. Anna Wałek, Prof. Anna Cedro, Dr. Leszek Szafranski)
- November - Publishing in Open Access
- December - Data Governance Plans: Ethical Sharing of Human Research Data (Dr Laura Bandura-Morgan)

During the meetings, notes are kept and all participants of the group can have access to them.

In 2022, there were also two physical meetings of the group:

On May 30-31, 2022, the Bridge of Data seminar was organized in Sopot – on “how to effectively motivate scientists to share research data? Meeting of the DSCC-IN PL working group”. The topics of the papers oscillated around issues related to research data, the profession of data stewards and Open Science, and a number of speeches were presented by members of the group. Participants also had the opportunity to take part in workshops.

On December 1st, 2022, as part of the 6th Pomeranian Open Science Conference, another meeting of the Polish Working Group was held in stationary form. The main topic of the meeting was to discuss legal issues related to research data. Organizational issues were also discussed, including meeting plans for the year 2023.

The basic communication channel is the Google group, where information about subsequent meetings and more important events and resources that may be of interest to the stewards is provided. For information and communication purposes, the team also used the Slack platform (until October 2022), as well as the Trello platform. Since May 2022, the group’s Facebook page has also been in operation: Data Stewardship Competence Centers PL - Polska Grupa Robocza.

In April 2023, members of the Working Group took part in a short survey aimed at determining the state of Open Research Data support services at Polish universities. Representatives of 28 institutions responded to the survey⁴.

⁴ There was no response from: Koszalin University of Technology, Maritime University of Technology in Szczecin, Poznan University of Technology, Main Business School in Warsaw, Cracow University of Economics, University of Adam Mickiewicz in Poznań, Jan Kochanowski University in Kielce, Maria Curie-Skłodowska University, University of Opole, The Pontifical University of John Paul II in Krakow, SWPS University, University of Warmia and Mazury in Olsztyn, West Pomeranian University of Technology in Szczecin.

Other institutions represented in the DSCC-IN PL Working Group: Institute of Agrophysics of the Polish Academy of Sciences; Institute of Bioorganic Chemistry of the Polish Academy of Sciences, Institute of Nuclear Physics Henryk Niewodzański, Polish Academy of Sciences, Institute of Animal Reproduction and Food Research Polish Academy of Sciences, International Institute of Molecular and Cell Biology, Information Processing Center - National Research Institute.

University	Competence Centre/ dedicated team	Library involved	Other units involved	Data steward/s	Librarians/Data librarians
AGH University of Science and Technology in Krakow	no	yes	yes	no	yes
Medical University of Gdańsk	yes	yes	yes	no	yes
Bydgoszcz University of Technology	no	no	yes	no	no
Gdańsk University of Technology	yes	yes	no	no / yes in the past	yes
Lublin University of Technology	yes (informal)	yes	no	no	yes
Warsaw University of Technology	no	yes	no	no	yes
Pomeranian Medical University in Szczecin	no	yes	no	no	yes
Warsaw University of Life Sciences	no	yes	no	no	yes
Wrocław University of Economics	yes	yes	yes	no	yes
University of Gdansk	yes	yes	no	no	yes
Jagiellonian University	yes (informal)	yes	no	yes	yes
Kazimierz Wielki University in Bydgoszcz	no	yes	no	no	yes

University of Lodz	no	yes	no	no	yes
Medical University of Silesian Piasts in Wrocław	no	yes	no	no	yes
Medical University of Lodz	no	yes	no	no	yes
Nicolaus Copernicus University in Toruń	no	yes	no	no	yes
University of Life Sciences in Lublin	yes (informal)	yes	no	no	yes
Nature University of Poznan	no	yes	no	yes	yes
University of Szczecin	yes (informal)	yes	yes	yes	yes
University of Silesia in Katowice	no	no	yes	yes	no
University of Warsaw	no	yes	no	no	yes
University of Wrocław	yes	yes	yes	no	yes
Military Technical Academy	yes	yes	no	no	yes

The information provided shows that a special team dedicated to Open Science and Research Data has not been established in most institutions (6 formal teams have been created, 4 informal). The information provided shows that a special team dedicated to Open Science and Research Data has not been established in most institutions (5 teams have been created). In seven universities, other units are involved. In five institutions data steward was appointed or employed. At the same time, the analysis shows that this is not a separate position indicated in the employment contract, but rather an additional function. Regardless of whether this person is employed in a library, administrative, technical or research position. Apart from one institution, all of them show the emergence of a new speciality in the competence of librarians. For the purposes of this article, we call all library roles with the original job names retained as data librarians. however, such a function does not appear in official nomenclature and documents.

Due to organizational changes at the Gdańsk University of Technology as well as establishing closer cooperation with the Ministry of Education and Science and the National Science Centre, in 2023 there will be organizational changes in the conduct of the working group. Also, focusing the group's activities on more formalized activities and strengthening technical competencies, in particular in the field of EOSC, resulted in deepening cooperation with ACK CYFRONET AGH, which is a leader in the implementation of tools and solutions related to EOSC.

POLISH OPEN DATA POLICY AND THE NATIONAL SCIENCE CENTRE ACTIVITIES

The national policy regarding Open Research Data in Poland is under development. This is due to the provisions of the EU directive (Directive no 2019/1024 of the EP and of the Council (EU) of 20 June 2019 on open data and the re-use of public sector information), as well as the Polish act, which was subsequently introduced in 2021 (Ustawa, 2021). However, in 2019, the National Science Centre (NCN), set up the obligation to attach the short Data Management Plan to all grant applications starting from September 2019. It was the first step to implementing the Open Research Data Policy.

The policy of the National Science Centre regarding the sharing of research data was implemented in three stages:

- 2019: Data Management Plan a mandatory element of each research project (in accordance with the guidelines and recommendations of Science Europe),
- 2020: NCN policy on open access to scientific publications based on Ordinances No. 38 and 40 of the NCN Director
- 2021: data related to publications are made available on a mandatory basis (National Science Centre, 2020).

Currently, the agency is the leading institution implementing the Open Science policy, while fulfilling other important roles in this context. The NCN is implementing a grant from the Ministry of Education and Science, aimed at preparing the Polish scientific community for the implementation of the open data policy and assumptions of the European Open Science Cloud (EOSC). The NCN is also a mandated organization appointed by the Ministry to the EOSC Association. It also implements a project with the acronym EOSC Focus (Supporting the development and coordination of activities of the EOSC Partnership), the purpose of which is to implement the EOSC assumptions and engage scientific units in activities for FAIR Data and Open Data.

The NCN is also carrying out a commission from the Ministry, by means of which it intensifies the implementation of the openness policy.

As a result, several MOOC courses are being prepared, which will then be made available on the Navoica platform⁵. They will be addressed to both data stewards and researchers. In addition, domain-specific webinars are carried out. The aim of each webinar is to provide participants with knowledge about research data management in a given discipline and competencies allowing for the preparation and implementation of Data Management Plans. Each webinar covers the following topics: Why is research data management important? FAIR Principles, Research Data Opening Policy, Data Management Plan General Information, Description of Data and Acquisition or Reuse of Available Data, Research Storage and Backup, Legal Requirements, Codes of Conduct, Data Sharing and Long-term Storage.

The above activities show that on the scale of the typeface, the development of data stewards' competencies is seriously considered, and the appropriate tools and regulations require preparation and implementation. Thanks to this, Poland will follow the international trend and will be ready to provide researchers with the appropriate level of support. This can only succeed if systemic solutions are actually implemented, not just bottom-up initiatives.

INTERNATIONAL PROJECTS RELATED TO DATA STEWARDS' COMPETENCIES

One of the first and most important projects dedicated to the development of data stewards' competences was the FAIRsFAIR project⁶.

"FAIR Competence Framework for Higher Education (Data Stewardship Professional Competence Framework)" is the third deliverable from Work Package 7 "FAIR Data Science and Professionalisation" of the FAIRsFAIR project.

The report presents a proposed FAIR Competence Framework for Higher Education (FAIR4HE) that is defined as a part of the general Data Stewardship Professional Competence Framework (CF-DSP) presented in the deliverable. The proposed CF-DSP defines the set of competencies that extend the competencies initially defined in the EDISON Data Science Framework (EDSF). The proposed competence framework is defined based on a recent job market analysis for the Data Steward and related professions. The presented CF-DSP has been validated against existing Data Stewardship competence frameworks defined primarily for the research community or practitioners. CF-DSP provides the competencies definition structure that allows easy mapping to a Body of Knowledge and

⁵ Navoica <https://navoica.pl/>

⁶ FAIRsFAIR project www.fairsfair.eu

a set of Learning Outcomes that can be used for defining academic curricula. The presented CF-DSP has been discussed with and incorporated feedback from, several community events organised by the FAIRsFAIR project (Demchenko et. al. 2021).

Issues related to the creation of definitions and a catalogue of skills for positions related to open science are also the subject of work in the Skills4EOSC project⁷.

The Skills4EOSC project brings together experiences of national, regional, institutional, and thematic Open Science and Data Competence Centres from 18 European countries with the goal of unifying the current training landscape into a common and trusted pan-European ecosystem. Its goal is to accelerate the upskilling of European researchers and data professionals in FAIR and Open Data. Competence Centres (CC) are defined as centres of gravity of OS and EOSC activities in their countries. CCs will pool expertise within research institutions, universities, and thematic and cross-discipline research infrastructures. They should offer training and support, empowerment, lifelong learning, professionalisation, and resources to a variety of stakeholders, including not only researchers and data stewards and funders but also decision-makers and industry.

The Skills4EOSC CC network has the ambition to drive the co-creation of harmonised trainer accreditation pathways, academic and professional curricula and skills quality assurance, recognition frameworks, and learning material creation methodologies (CMCC, 2022).

The assumptions of the project are therefore consistent with the results already achieved in Poland by the OSCC at the Gdańsk Tech Library. What is interesting, however, when applying for the project, Polish partners (PSNC – Instytut Chemii Bioorganicznej Polskiej Akademii Nauk) completely omitted the achievements of the Polish data stewards network and the Gdańsk Tech. This shows that in many places and centres, work in this area is carried out without particular coordination. The first catalogue of such international initiatives was carried out by the EOSC Task Force, which invited, among others, the Skills4EOSC project to cooperate. OSCC, on the other hand, invited Polish representatives of the project to present its assumptions during the Pomeranian Open Science Conference in Gdańsk in December 2022.

One of the deliverables of the project is the so-called Minimum Viable Skillset (MVS). The proposed Minimum Viable Skillsets for the European Open Science Cloud (EOSC) were designed within the Skills4EOSC project for defining the Open Science mission, activities, or outcomes expected of relevant roles. MVS's task is to synthesise and profile each role as an aid to developing new curricula, career paths and courses. A simple MVS format

⁷ Skills4EOSC project <https://www.skills4eosc.eu>

is proposed to articulate key skills and competencies. It describes essential skills and concepts required to deliver Open Science outcomes for communities and organisations. An MVS may be associated with a career profile describing disciplinary variations in skills and competencies needed, or guidance on the improvement in proficiency levels needed for career path progression. The initial set includes roles such as Data Steward, Data Curator, Data Librarian, and Data Manager (Giroletti, et. al., 2023).

DATA STEWARDS EDUCATION IN POLAND

For the purposes of this article, the educational offer of Polish public universities was analysed, in particular, those educating in the broadly understood library and information science. None of the universities offers a course or studies that meet the needs of data stewards education.

The first and so far the only comprehensive training program for data stewards in Poland is the Data Steward School (DSS) initiative. The initiative was implemented by Visnea sp. z o.o. (Visnea) in cooperation with GO-FAIR for the first time in the period from September 2020 to April 2021. 13 participants took part in the first pilot edition of DSS. Librarians accounted for 75% of participants. The total number of theoretical training hours was 65 and practical classes 30 (Pawłowska, Wachowicz, 2021).

Participation in the Data Steward School Training Program is aimed at preparing participants to act as data stewards in their home research unit and allows them to gain new professional qualifications and unique competencies in the field of data management.

One of the objectives of the DSS was also to establish the best practices related to the professional role of the data steward, taking into account the specificities of scientific units. Therefore, the Data Steward School Training Programme was planned as a comprehensive, 7-month programme implemented by Visnea in cooperation with the GO-FAIR Initiative.

The Data Steward School was run in the train the trainers formula, which means that graduates obtained the appropriate competencies to train further persons in their institutions. Such a model was designed to enable the most effective implementation of international best practices in Poland. Program participants take part in the Foundation Level training module, where they learn about key issues related to the processing, collection and archiving of scientific data, and then choose one of the following specializations: strategic management, IT/technologies, scientific information, scientific data management in a research project. The final part of the course is mentoring with the RDM specialists.

The second edition took place in autumn 2021 (Pawłowska, 2021) and the third edition took place in autumn 2022 when the Strategic Partner of

the Data Steward School Training Program was Digital Intelligence Offensive from Austria, a European leader in the use of data in science and business⁸. Each edition of the course allowed for the education of 12-14 new stewards. Taking into account the growing needs and the dynamic development of the data management area, this number is only a fraction of the needs. Nevertheless, DSS lecturers and graduates are the backbones of the growing network of data stewards in Poland.

STUDIES AND COURSES

The unit that has been involved in the development of data stewards' competencies in Poland from the beginning is the Library of the Gdańsk University of Technology. It conducted a series of stationary workshops and webinars, in particular in the field of DMP preparation, which were attended by a total of almost 700 people – employees of libraries, university administration and researchers.

Gdańsk Tech Library was also involved in partnerships aimed at preparing materials and courses for people related to data management and Open Science.

From the 6th to 10th of September 2021 the CODATA-RDA Schools instructors together with the FAIRsFAIR project, EOSC-Pillar, EOSC Synergy and the Gdańsk Tech Library, delivered a three-day train-the-trainer workshop on Data Stewardship. The key aim of the workshop was to empower a network of peers where best practices are exchanged and where those with more experience can share their knowledge with those just getting started.

The workshop combined a series of theoretical and practical sessions, where participants were given the opportunity to interact and get to know each other and exchange experience in supporting researchers with RDM. The workshop aimed to introduce participants as data stewards to the key concepts and drivers for Open Science, RDM and FAIR principles, to enable the resulting network of practitioners in peer institutions who can collaborate, work and learn together. Participants were also introduced to key concepts of pedagogy and were able to put these into practice with a brief course design and development activity. In addition, attendees were asked to reflect on their own role as data stewards and got to evaluate the existing training offer and the status of RDM services in their respective institutions in the perspective of training their members in the future (Dudek, Krajewski, 2021).

The pre-and post-course evaluation showed that most of the participants had been in their data-related role (often managing data stewards'

⁸ Data Steward School 2022 <https://www.visnea.org/data-steward-school>

tasks besides regular duties) for less than 6 months and worked on their own or in teams of less than 5 people. The workshop met the expectations of the participants, who joined with an interest to learn more about topics like the range of skills and knowledge associated with data stewardship, political drivers for RDM, FAIR and Open Data, how to explain the difference between FAIR and Open Data to researchers, how to provide and develop effective training using open learning resources, how to implement RDM services and advance FAIR data within their organization. However, attendees identified also gaps and areas of improvement to support research activities in their respective institutions like lack of the required data management skills to effectively provide data services, lack of sufficient time to handle the requirements of data management, lack of advisory services and training on RDM, DMPs as well as they confirmed the need of developing local peer network. The event gathered more than 30 participants from the emerging RDM support staff community in the country. It was conducted in English although there was a chance to speak Polish during all practical sessions as each breakout room had the local moderator, a member of the Open Science Competence Center at Gdańsk Tech Library, to feedback in English. The purpose of the workshop, therefore, was to provide a better understanding of data stewardship, and the research data management process and support the development of data stewardship skills among staff in universities and other research institutions in Poland (Dudek, Krajewski, 2021).

The first comprehensive postgraduate studies were launched by the University of Vienna⁹. It is worth looking at the thematic scope to see that the skills that constitute the professional profile of a data steward are both technical (IT and data science) and closely related to research practice and knowledge of the scientific discipline. The creators of the course list the following areas: Basics of Research Data Management and Open Science, Basics of IT and Data Science, FAIR Research Data in the Life Cycle, Research Data Management Support and Data Stewardship in Practice. However, it is still hard to find similar courses or studies at most universities, not only in Poland.

SUMMARY

The first roles of data stewards in Polish scientific institutions appeared in libraries. This resulted directly from the policy of the National Science Centre but above all also from the fact that the first unit providing sup-

⁹ Data steward. University of Vienna Postgraduate Center <https://www.postgraduatecenter.at/en/programs/communication-media/data-steward/content-and-objectives/>

port for researchers in the field of data management – the Open Science Competence Center – was established in the Gdańsk Tech Library. The Gdańsk Tech Library worked with international networks and organizations, which resulted in active cooperation and placing it in international structures. In particular as a national chapter of the GO-FAIR network (DSCC-IN PL: Data Stewardship Competence Centers Implementation Network – Poland) and several working groups related to EOSC. This initiated the creation of the Polish working group of data stewards with more than 100 participants, exchanging knowledge and experience on a regular basis. The lack of educational programs at university or in any other form of vocational education has been filled in a sense through the training activities of the OSCC and the only Data Steward School in Poland.

It was only in 2022 that the increased activity of the National Science Center and the Ministry of Education and Science began. In conjunction with the work on the national policy for open data as well as activities within the European Open Science Cloud, a coherent environment began to emerge for the development of a new specialization at the level of Polish universities and other scientific institutions.

Not only in Poland but all around the world, depending on the institution, data stewards are located in libraries, at infrastructure providers, in research projects, and at RDM competence centers or support desks. Despite efforts to work out competencies for data stewards, there is no final, demand-oriented, formal model of data stewards and the associated roles.

Nevertheless, libraries' services provided to researchers and research institutions in order to assist them with the research data management process are evolving. The new duties of libraries differ from the previous tasks they have been responsible for so far, and libraries must adapt to the new conditions to be up to date with services provided in data management. Librarians have become a part of the research process from the very beginning of the data lifecycle. Their knowledge and engagement are essential from the very early moment of DMP creation through collecting, describing, preserving, and curating data. Additionally, librarians are usually also responsible for ensuring that data sets prepared by research teams are properly used and cited. The future role of the academic library may be to coordinate the work of such multi-task teams responsible for creating repository tools. The role will also include developing institutional research data management services and providing research data management advisory and training services. These new tasks and activities will entail the need to work collaboratively with other stakeholders to ensure that future research is supported effectively. However, despite the fact that libraries have been trying to meet the challenges so far, it is necessary to introduce systemic solutions, educate and employ qualified data stewards in scientific units who will be able to support researchers

at every stage of research work and its settlement, as well as making the results public.

BIBLIOGRAPHY

- CMCC (2022). Skills4EOSC Skills for the European Open Science Commons: Creating a Training Ecosystem for Open and FAIR Science <https://www.cmcc.it/projects/skills4eosc-skills-for-the-european-open-science-commons-creating-a-training-ecosystem-for-open-and-fair-science>
- Dudek, K. (2021). Polish Working Group of the DSCC IN established. GO FAIR newsletter. 15 December <https://www.go-fair.org/2021/12/15/polish-working-group-of-the-dscc-in-established/>
- Dudek, K., Krajewski, P. (2021). Data Stewardship Training – Gdańsk, Poland. CODATA-RDA Schools for Research Data Science newsletter. October 2021. <https://mailchi.mp/4bf1a23cdf93/newsletter-december-2020-codata-rda-schools-for-research-data-science-6294878>
- Giroletti, J., Budroni, P., Prnjat, O., Galeazzi, F., Schirru, L., Drees, T., Lazzeri, E., & Di Giorgio, S.. (2023). D8.1 MoU or Letter of Intent on key cooperation topics (1.0). Zenodo. <https://doi.org/10.5281/zenodo.7684500>
- Jetten, Mijke, Marjan Grootveld, Annemie Mordant, Mascha Jansen, Margreet Bloemers, Margriet Miedema, and Celia W. G. van Gelder. 2021. "Professionalising Data Stewardship in the Netherlands. Competences, Training and Education. Dutch Roadmap towards National Implementation of FAIR Data Stewardship." Zenodo. <https://doi.org/10.5281/zenodo.4623713>
- Kalová, T., Frontini, F., Bracco, L., Blümel, I., Legat, D., Galik, M., Meeus, J., Wałek, A., Szuflita-Żurawska, M., (2023) Data Stewardship Career Paths: State-of-the-Art Report and Recommendations. Outcome of the Data Stewardship Curricula and Career Paths EOSC Task Force [unpublished]
- Narodowe Centrum Nauki, Wprowadzenie Polityki Otwartego Dostępu w Narodowym Centrum Nauki. <https://www.ncn.gov.pl/aktualnosci/2020-06-03-wprowadzenie-polityki-otwartego-dostepu>
- OECD (2020). "Building Digital Workforce Capacity and Skills for Data-Intensive Science." Paris: OECD. <https://doi.org/10.1787/e08aa3bb-en>.
- Pawłowska M. M., Wachowicz, M., (2021). First Polish Data Steward School, TASK QUARTERLY vol. 25, No 4, 2021, pp. 511–519, <https://doi.org/10.34808/tq2021/25.4/t>
- Pawłowska, M.M., Polish Data Steward School 2021 – 2nd Edition, GO-FAIR News (Posted on 1 October 2021), <https://www.go-fair.org/2021/10/01/polish-data-steward-school-2021-2nd-edition/>
- Science, Digital, Goodey, Gregory; Hahnel, Mark; Zhou, Yuanchun; Jiang, Lulu; Chandramouliswaran, Ishwar; et al. (2022): The State of Open Data 2022. Digital Science. Report. <https://doi.org/10.6084/m9.figshare.21276984.v5>
- Tenopir, C., Sandusky, R. J., Allard, S., & Birch, B. (2012). Academic librarians and research data services: Preparation and attitudes (Session 116). Paper presented at IFLA World Library and Information Congress, Helsinki, Finland, <http://conference.ifla.org/past-wlic/2012/116-tenopir-en.pdf>

- Ustawa z dnia 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystaniu informacji sektora publicznego.
- Wałek, A. (2019). Data librarian and data steward – new tasks and responsibilities of academic libraries in the context of Open Research Data implementation in Poland. *Przegląd Biblioteczny* (z. 4), 497-512. <https://doi.org/10.36702/pb.634>
- Wałek, A., Helling, P., Rau, F., Linne, M., Dierkes, J., Jetten, M., Szuflińska-Zurawska, M. (2022). The Importance of Demand and Environment for Defining and Establishing the Role of Data Stewards. <https://doi.org/10.5281/zenodo.6511184>.
- Yuri Demchenko, Lennart Stoy, Claudia Engelhardt, & Vinciane Gaillard. (2021). D7.3 FAIR Competence Framework for Higher Education (Data Stewardship Professional Competence Framework) (1.0 DRAFT). Zenodo.<https://doi.org/10.5281/zenodo.4562089>