Supporting Early-Career Researchers with Open Science

BME Horizon Hot Topics - Webinars, 21 Mar 2022, online Budapest University of Technology and Economics
eva.hnatkova@techlib.cz
About me

● Open Science Coordinator at
  ○ National Library of Technology
  ○ UCT Prague
● Eurodoc former president
● PhD in process engineering

@EvaHnatkova
Overview

- Open Science practices
- Benefits and challenges for ECRs
- Open Science in Horizon Europe
- Open Research Europe
- Data Stewardship Course (DocEnhance)
- Reform on Research Assessment
Open Science: an umbrella term

- Open Data
- Open Access
- Open Source
- Citizen Science
- Open Peer Review
- Open Licencing
- Open Educational Resources
Benefits - Why do Open Science?

- opens access to research
- increases discoverability
- increases (social) impact
- facilitates reproducibility
- shares & saves resources
- speeds & helps innovation

Access to research is considered important to help solving global challenges
Challenges

- Researchers are **unaware** to what Open Science is and how to do it
- Lack of **training courses** on Open Science or Data Management Plans (DMP)
- Lack of **support** at institutions (no guidelines, no specialists)
- Lack of a clear **Open Science policy** at institutions
- Lack of **strategic plans** to implement Open Science at national level

Open Science in Horizon Europe

Mainstreaming of open science practices for improved quality and efficiency of R&I, and active engagement of society

Mandatory immediate Open Access to publications: beneficiaries must retain sufficient IPRs to comply with open access requirements;

Data sharing as ‘open as possible, as closed as necessary’: mandatory Data Management Plan for FAIR (Findable, Accessible, Interoperable, Reusable) research data

- Work Programmes may incentivize or oblige to adhere to open science practices such as involvement of citizens, or to use the European Open Science Cloud
- Assessment of open science practices through the award criteria for proposal evaluation
- Dedicated support to open science policy actions
- Open Research Europe publishing platform

Source: Horizon Europe - Investing to shape our future | European Commission
Open Research Europe (ORE)

OA Publishing Platform for H2020 & HEU research results

- Initiative of EC
- By F1000 Research
- Immediate OA
- Open/FAIR Data support
- Open Peer Review
- NO author fee
- All disciplines
- Article-level metrics

Official launch 24 March 2021

https://open-research-europe.ec.europa.eu/
ORE Publishing process

1. Article Submission
2. Publication & Data Deposition
3. UNDERGOING PEER REVIEW
   - Open Peer Review & Article Revision
4. PASSED PEER REVIEW
   - Send to Indexers & Repositories

Reviewer provides peer review and status:
- Approved
- Approved with reservation
- Not approved

NOTE: authors may continue to publish new versions, even once peer review passed.
ORE: Various types of articles

- Research Articles
- Brief Reports
- Data Notes
- Method Articles
- Software Tool Articles
- Study Protocols
- Clinical Practice Articles
- Case Studies & Reports
- & more!

With Open Research Europe you can share all your work through our range of non-traditional article types, including data notes, clinical trials, study protocols, systematic reviews and more.

Open Research Europe

See full list for the subject area: Article Guidelines
ORE: Open Peer Review example

Open Peer Review

Alongside their report, reviewers assign a status to the article:

- **APPROVED**
  - The paper is scientifically sound in its current form and only minor, if any, improvements are suggested

- **APPROVED WITH RESERVATIONS**
  - Key revisions are required to address specific details and make the paper fully scientifically sound

- **NOT APPROVED**
  - Fundamental flaws in the paper seriously undermine the findings and conclusions

Visibility & credit for reviewers:
- Co-reviewing
- ORCID ids
- DOIs for reports
Data Stewardship Course
DocEnhace

- 3 years (2020 - 2022)
- EU H2020 project
- 19 partners

UiT The Arctic University of Norway

Training resources available at DocEnhace Platform

https://docenhance.eu/
- under licence CC-BY

Course available by Dec 2022
Aims of the Data stewardship course

- To introduce research data management to PhD candidates
- to walk through all the steps of research data management, from finding relevant data to publishing an own dataset
- to help with the PhD thesis, while at the same time to make PhD candidates feel better prepared for a job outside academia after graduation
- to combine practical exercises with theoretical content
Data Stewardship Course

Developed by a group of RDM specialists at UiT, involved in RDM teaching at the institution

UiT The Arctic University of Norway

Three-person working group in charge of it all

- Helene N. Andreassen
- Henning Hansen
- Leif Longva

21/03/2022
Pilot courses

University of Chemistry and Technology in Prague, Czechia

~ 25 Doctoral candidates
~ 8 teachers

Pilot - financial remuneration for active participation and subsequent feedback

Karlstad University, Sweden

~ 25 Doctoral candidates

https://phd.vscht.cz/home/pro-doktorandy/docenhance-data-stewardship-course
## Three modules

<table>
<thead>
<tr>
<th>MODULE 1</th>
<th>MODULE 2</th>
<th>MODULE 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>online lectures</td>
<td>workshops &amp; assignments in local working teams</td>
<td>workshops &amp; regional assignments with employers</td>
</tr>
</tbody>
</table>
| ● general introduction  
● basics of how to handle research data in a PhD thesis | ● learning by doing  
● practical assignments  
● working in small groups  
● discussions  
● local resources  
● meet local demands | ● practical tasks in cooperation with business / public service entity  
● help to prepare PhD candidates for a career outside academia  
● putting into practice the theory & practical skills |

21/03/2022
Module 1: 10 thematic sections

Focus on motivation, theory and concepts (~ 16 hours)

1. Introduction
2. About research data
3. How to search & cite research data
4. Research contracts
5. How to structure & document research data
6. How to store research data
7. How to visualise research data
8. How to archive research data
9. Rights and licenses for research data
10. How to write a research DMP

➢ Video, suggested readings
➢ Interaction via quizzes and open « food for thought »
5.3 The Readme-file

What is meant by documenting a dataset? And why is documenting a dataset important? In this video you will learn why good documentation of a dataset is crucial in order to avoid misinterpretation of the dataset.

How to write a Readme-file

What should be included?

- Contact information to the dataset author
  - Include author’s name and contact information
- What the data set is about
  - Write a short description
    - Include information on project relevance
- File structure and naming syntax
- Where to find which data
- Overview of the files
- Information on updates to your data set

Transcript of video “How to structure .... Documenting your data - The Readme-file”

Lessons learned:

- To document a dataset is to give a human readable introduction and explanation of what information the dataset holds.
- You document your dataset by creating a separate file usually called a Readme file.
- The information in the Readme file should make sure anyone is able to understand and interpret your dataset correctly, both now and also many years from now.
- You should start entering information into your Readme file early, and update the file as new information is obtained.
- The Readme file should supplement the information found in the dataset’s metadata.
- The Readme file should be in plain a preferred format, either plain text with UTF8, or PDF/A.

Food for thought

Think through your own PhD project and the data you have collected, or plan to collect. What do you see as essential to include and explain in a Readme file, to make sure your dataset is understood correctly by outsiders?
Module 1: exam quiz & certification

Final exam quiz
● If you get 80% or 24 correct answers

Certification M1
● generated by the Moodle
Module 2 - workshops & assignments

- Teacher material on the Moodle platform
- The local teachers decide which content and activities to use
- 6 online 2 - 3 hours workshops
- Break out rooms - working in small groups

<table>
<thead>
<tr>
<th>Description</th>
<th>24 Monday</th>
<th>25 Tuesday</th>
<th>26 Wednesday</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>8:00-09:45</td>
<td>8:00-09:45</td>
<td>8:00-09:45</td>
</tr>
<tr>
<td>Session</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main lecturer</td>
<td>The Module 2 starts at 10:00 on Monday</td>
<td>#3: Structuring, documenting &amp; metadata</td>
<td>#4: Data storage, file formats &amp; FAIR principles</td>
</tr>
<tr>
<td>Lecturer 2</td>
<td>Jan Vrbá</td>
<td>Marcela Dendisová</td>
<td>Eva Hnatkova</td>
</tr>
<tr>
<td></td>
<td>10:00-12:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Session</td>
<td>#1: The importance of RDM &amp; open archiving</td>
<td>Czech Open Science Day @ National Technical Library</td>
<td>#6: Visualising, reusing/searching &amp; citing research data</td>
</tr>
<tr>
<td>Main lecturer</td>
<td>Eva Hnatkova</td>
<td>P. Doleček</td>
<td>Radek Cibulka</td>
</tr>
<tr>
<td>Lecturer 2</td>
<td></td>
<td>Please use the following link to connect to this event:</td>
<td>Eva Svobodová</td>
</tr>
<tr>
<td></td>
<td>13:00-15:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Session</td>
<td>#2: The data management plan</td>
<td>#5: Sharing &amp; licensing research data</td>
<td>#7: Wrap Up</td>
</tr>
<tr>
<td>Main lecturer</td>
<td>Eva Hnatkova</td>
<td>Jiří Jiřát</td>
<td>Prof. Pavel Matějka</td>
</tr>
<tr>
<td>Lecturer 2</td>
<td>Robert Pergl</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example DMP

DSW online tool

<table>
<thead>
<tr>
<th>Feature</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Template Selection</td>
<td>✔️</td>
</tr>
<tr>
<td>Detailed Export Settings</td>
<td>✔️</td>
</tr>
<tr>
<td>Template Creation</td>
<td>✔️</td>
</tr>
<tr>
<td>Template Styling</td>
<td>✔️</td>
</tr>
<tr>
<td>Project Phases</td>
<td>✔️</td>
</tr>
<tr>
<td>Suitable for Worldwide Use</td>
<td>✔️</td>
</tr>
</tbody>
</table>

Doctoral training running on UCT’s own DSW instance: one of the first joint endeavors of CTU and UCT in promoting #datastewardship in the Czech Republic!
Module 3

- Purpose: Employ RDM knowledge in cooperation with non-academic sector
- Focus on application of knowledge from Modules 1 and 2
- Teacher material on the Moodle platform
- Advice from the non-academic sector, assignment ideas, community-driven resource bank

The content of this module very much depends on the local stakeholders involved
Module 3 - workshops & assignments

One day workshop

Session 1 “Europe and Czech Republic – public sector” (11:00 – 12:30 CEST)

Topic: European, national and institutional level and their approach to administration and sharing of the research data.

Lecturers:

Gareth O’Neill (EOSC Policy & Strategy Leader, EOSC Future)

Principal consultant on Open Science at Technopolis Group, Policy and strategy leader at EOSC Future. Adviser at EOSC Association for EOSC Secretariat. Ambassador for Open Access and Plan S at cOAlition S. Doctoral candidate in linguistics at Leiden University. Irishman between Amsterdam and Brussels.

Martin Svoboda (Director of Czech National Library of Technology in Prague – NTK)

He graduated at Czech Technical University in computer engineering in 1965 and carried out research and development in text processing and information retrieval systems. As a director of NTK he negotiated “Big Deals” with key publishers for the entire Czech scholarly and R&D community, which resulted in the nation-wide CzechELib (National Centre for Electronic Resources) project and currently the “National Centre for Information Support of Research, Development, and Innovation” project. In 2020, he was nominated as a chief negotiator of transformative agreements for access to and publication of electronic information resources for Czech research. He served as a member of the LIBER Executive Board (2014-2020), currently is a representative of the Czech Republic in the SCOAP3 project, a member of LIBER Architecture Group, ELAG Programme Committee and Journal of Documentation Editorial Board.

Luděk Matyska (Director, Institute of Computer Science; Head of ICS Management and Secretariat Centre CERIT-SC; Director of CERIT SC)

Session 2 “ORLEN Unipetrol” (13:00 – 14:30 CEST)

Topics:
1. DCS systems, data collection, archiving and protection
2. PI Osi Soft application - on-line visualization, collection and data mining Michal
3. Example of using operational data to evaluate the efficiency and performance parameters of catalysts
4. Discussion

Lecturers:

Tomáš Herink (ORLEN Unipetrol – Member of Board)
Michal Zbuzek (ORLEN Unipetrol)
Arnošt Krížan (ORLEN Unipetrol)

Session 3 “ZENTIVA” (14:30 – 16:00 CEST)

Topic: The third session will be devoted to a non-academic view - industrial “practice”: How Zentiva works with research data, what is difference from academia, how research data management is connected with the overall concept of development and research in Zentiva.

Lecturers:

Marcela Tkalčecová
Obor fyzikální chemie na VŠCHT Praha
PhD z VŠCHT Praha
Dlouhé roky na ústavu analytické chemie v laboratoři NMR spolu s J.H. Posledních více než 10 let NMR ve vývoji v Zentivě a k tomu part time učení na ÚFCH.

Jaroslav Havlíček
Tomáš Pekárek

https://phd.vscht.cz/home/pro-doktorandy/docenhance-data-stewardship-course
Implementation

- The local university decides whether PhD candidates receive credits for taking the course.
- All three modules come with an assignment that together may function as examination.
DocEnhance platform

A dynamic, adaptable & sustainable platform for **online collaboration** (under development)

- Platform will be integrated into the existing European PhD Hub
- **Training resources** (courses, guidelines, reports, studies) will be freely available to all

The aim is supporting innovative career oriented PhD training

https://phdhub.eu/
Research Assessment

BUT to practice Open Science we need **Incentives & rewards**

**Current system incentives**

- only publications with high IF
- not other results/outputs of research
- it encourages individuals rather than collaboration
Process towards an agreement on reforming research assessment

The Commission has called for organisations to express their interest in being part of a coalition on reforming research assessment.

The coalition will bring together research funding organisations, research performing organisations, national/regional assessment authorities or agencies, associations of research funders, of research performers, of researchers, as well as, learned societies and other relevant organisations, all willing to participate in this process.

Open call for expression of interest | List of organisations