# DocEnhance
# Pilot Data Stewardship Course

LIBER Annual Conference - Workshop FAIR data enabling through European research libraries, 25 June 2021, online

eva.hnatkova@techlib.cz

# Overview

- Background information
- DocEnhance
- Data Stewardship Course
- Module 1, 2, 3
- DocEnhance Platform

# About me

@EvaHnatkova

- Open Science Coordinator at
  - National Library of Technology
  - UCT Prague
- Eurodoc former president
- DocEnhance Advisory Board member
- PhD in process engineering

**NTK**
50°6'14.083"N, 14°23'26.365"E
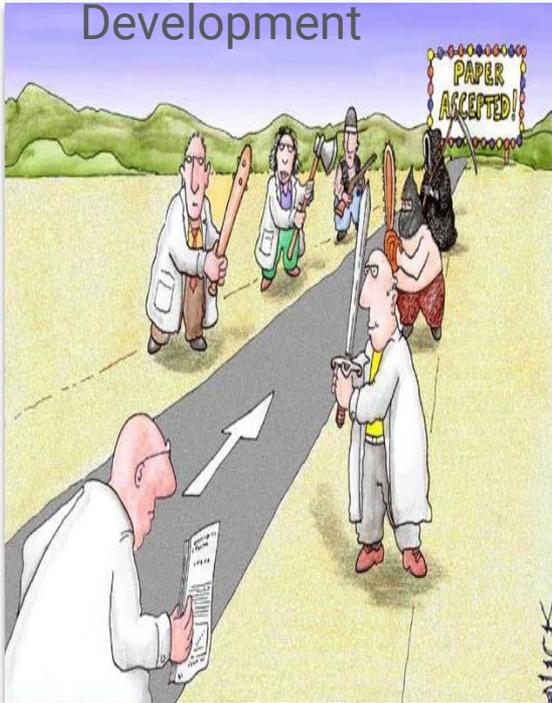Národní technická knihovna
National Library of Technology

**UNIVERSITY OF CHEMISTRY AND TECHNOLOGY PRAGUE**

**eurodoc**
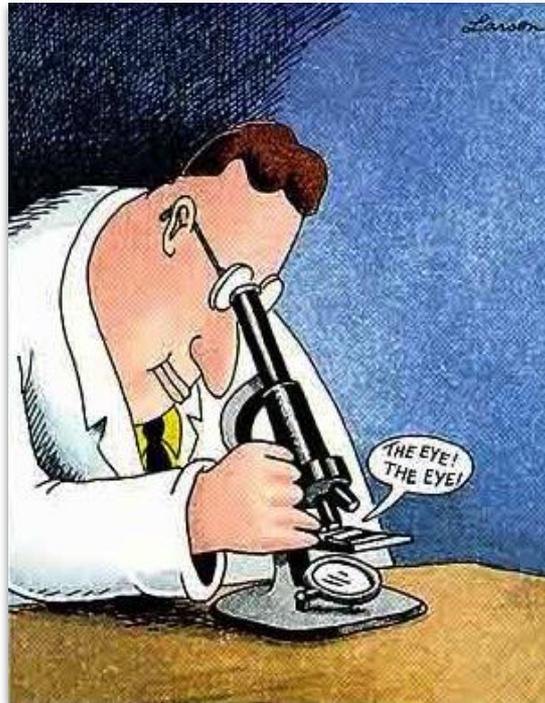The European Council of Doctoral Candidates and Junior Researchers

# Eurodoc = early career researchers

**Members** **Observers** **Individual observers**

http://www.eurodoc.net/

- non-profit organisation
- based in Brussels (2002)
- 28 national associations
- funded by member fees
- supported by volunteers
- develop & inform policy
- run workgroups & events

**eurodoc**

The European Council of Doctoral
Candidates and Junior Researchers

# Issues for early-career researchers

Career Development

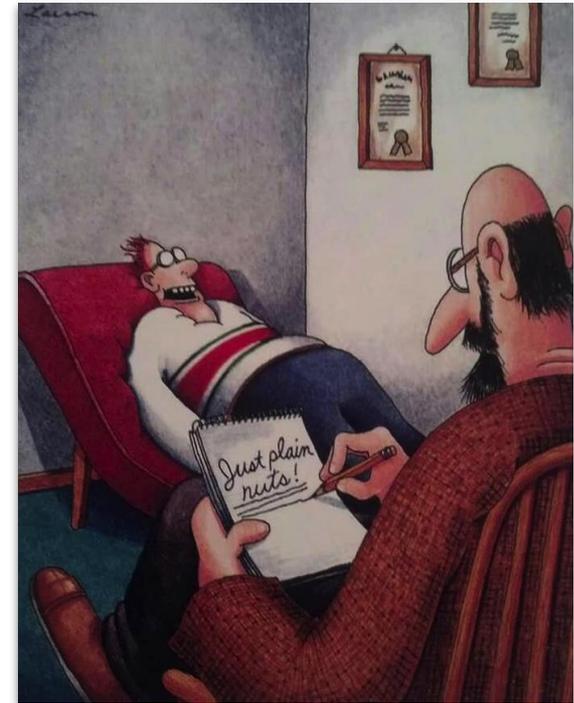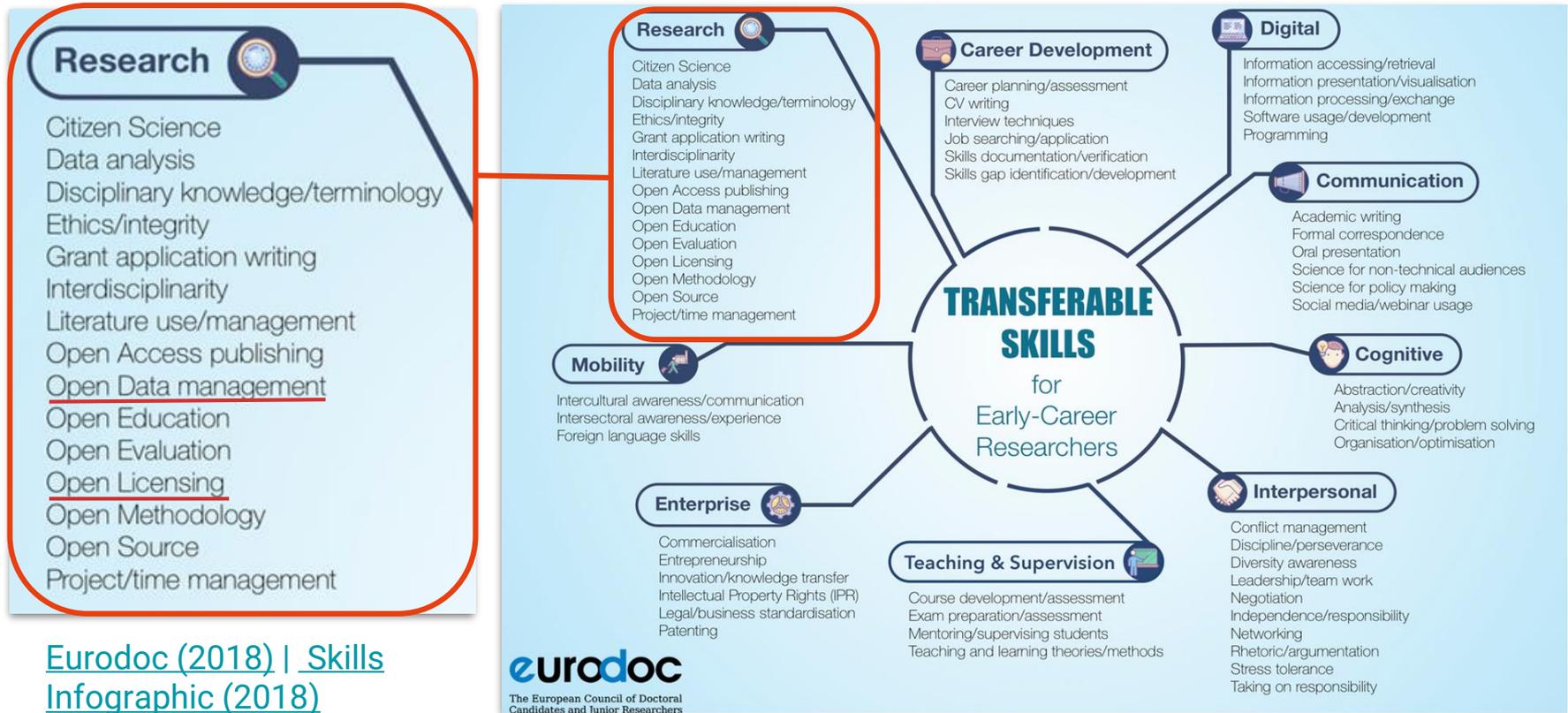Open Science

Mental Health



O'Neill & Storti (2018) | O'Neill & Hnátková (2018) | O'Neill & Schroijen (2018)

# Development of transferable skills

**Research**

Citizen Science
Data analysis
Disciplinary knowledge/terminology
Ethics/integrity
Grant application writing
Interdisciplinarity
Literature use/management
Open Access publishing
Open Data management
Open Education
Open Evaluation
Open Licensing
Open Methodology
Open Source
Project/time management

Eurodoc (2018) | Skills Infographic (2018)

**Research**

Citizen Science
Data analysis
Disciplinary knowledge/terminology
Ethics/integrity
Grant application writing
Interdisciplinarity
Literature use/management
Open Access publishing
Open Data management
Open Education
Open Evaluation
Open Licensing
Open Methodology
Open Source
Project/time management

**Career Development**

Career planning/assessment
CV writing
Interview techniques
Job searching/application
Skills documentation/verification
Skills gap identification/development

**Digital**

Information accessing/retrieval
Information presentation/visualisation
Information processing/exchange
Software usage/development
Programming

**Communication**

Academic writing
Formal correspondence
Oral presentation
Science for non-technical audiences
Science for policy making
Social media/webinar usage

**TRANSFERABLE SKILLS**
for
Early-Career
Researchers

**Cognitive**

Abstraction/creativity
Analysis/synthesis
Critical thinking/problem solving
Organisation/optimisation

**Mobility**

Intercultural awareness/communication
Intersectoral awareness/experience
Foreign language skills

**Interpersonal**

Conflict management
Discipline/perseverance
Diversity awareness
Leadership/team work
Negotiation
Independence/responsibility
Networking
Rhetoric/argumentation
Stress tolerance
Taking on responsibility

**Enterprise**

Commercialisation
Entrepreneurship
Innovation/knowledge transfer
Intellectual Property Rights (IPR)
Legal/business standardisation
Patenting

**Teaching & Supervision**

Course development/assessment
Exam preparation/assessment
Mentoring/supervising students
Teaching and learning theories/methods

**eurodoc**
The European Council of Doctoral
Candidates and Junior Researchers

# DocEnhance

# DocEnhance - EU project

**Topic:** **SwafS-08-2019: Research innovation needs & skills training in PhD programmes**    **Forthcoming**

**Publication date:**    27 October 2017

**Types of action:**    CSA Coordination and support action
**DeadlineModel:**    single-stage                                    **Deadline:**    02 April 2019 17:00:00
**Opening date:**    11 December 2018

Time Zone : (Brussels time)

**Broadening PhD expertise**
**through transferable skills**

https://docenhance.eu/

# DocEnhance

- 3 years (2020 - 2022)
- funding from EU H2020
- 19 partners
- Project website

https://docenhance.eu/



UiT The Arctic University of Norway

Photo: Jonatan Ottesen, UiT

# Key deliverable - 3 pilot courses
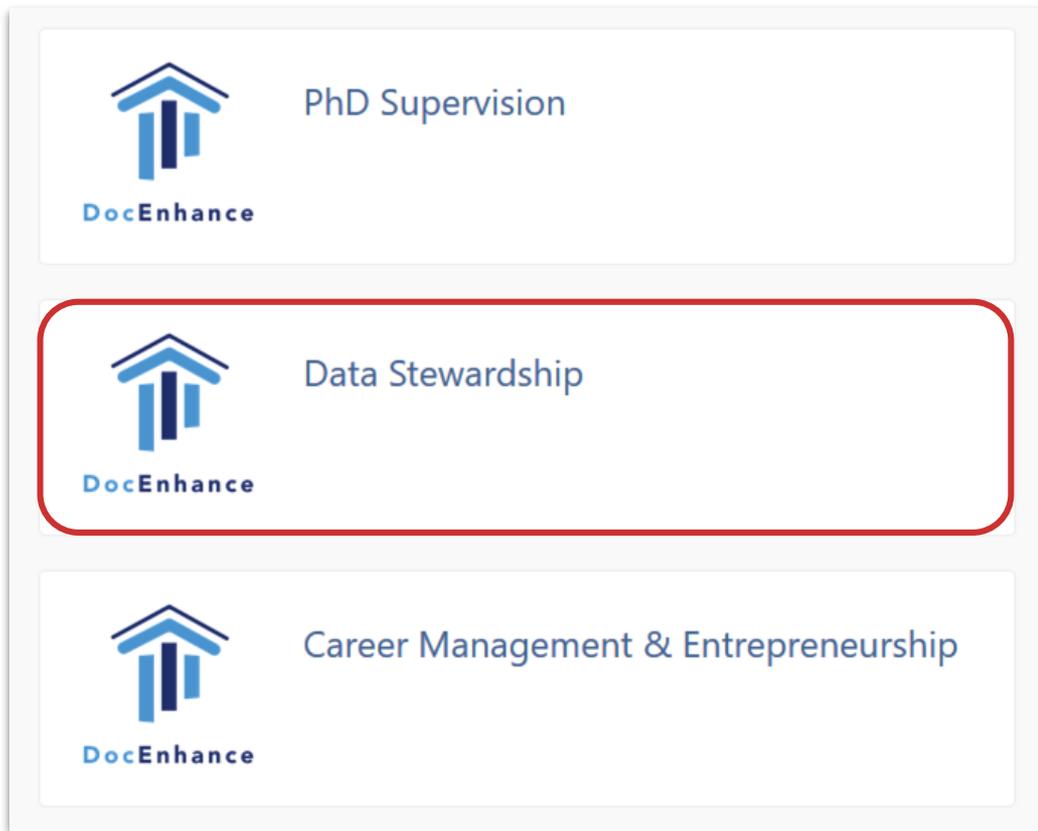


PhD Supervision

**DocEnhance**

Data Stewardship

**DocEnhance**

Career Management & Entrepreneurship

**DocEnhance**

Open for reuse on the DocEnhance Platform

- Under licence CC-BY



Planned release:
- End of 2022

# Data Stewardship Course

# 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

# Aims of the course

✓ to introduce to the complexity and significance of 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

Plan

Collect

Re-use

Research Data Life Cycle

Process

Share

Preserve

Analyse
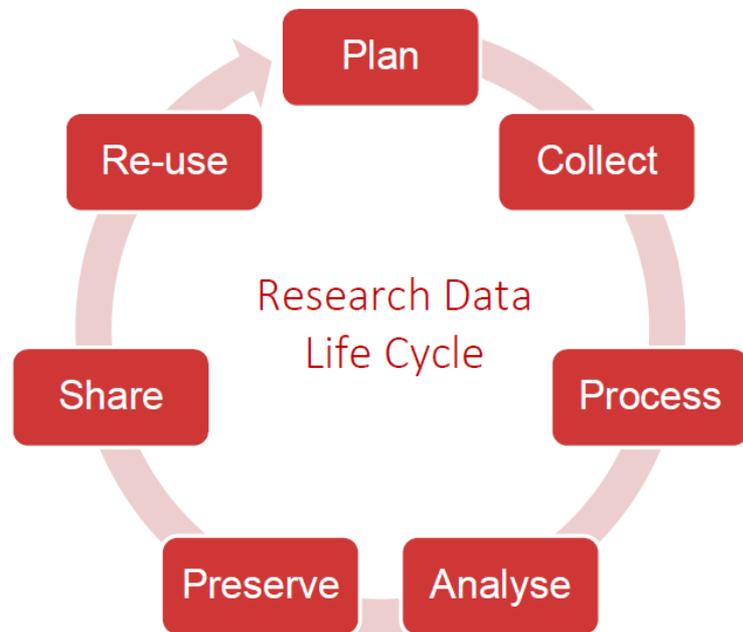
# Pilot courses

## 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

# Module 1, 2, 3

# Three modules

**Open Science and Innovation**

| 1st module | 2nd module | 3rd module |
|---|---|---|
| Open education | Interdisciplinarity | Mobility |
| **Online lectures** | **Local group work** | **Regional assignment** |
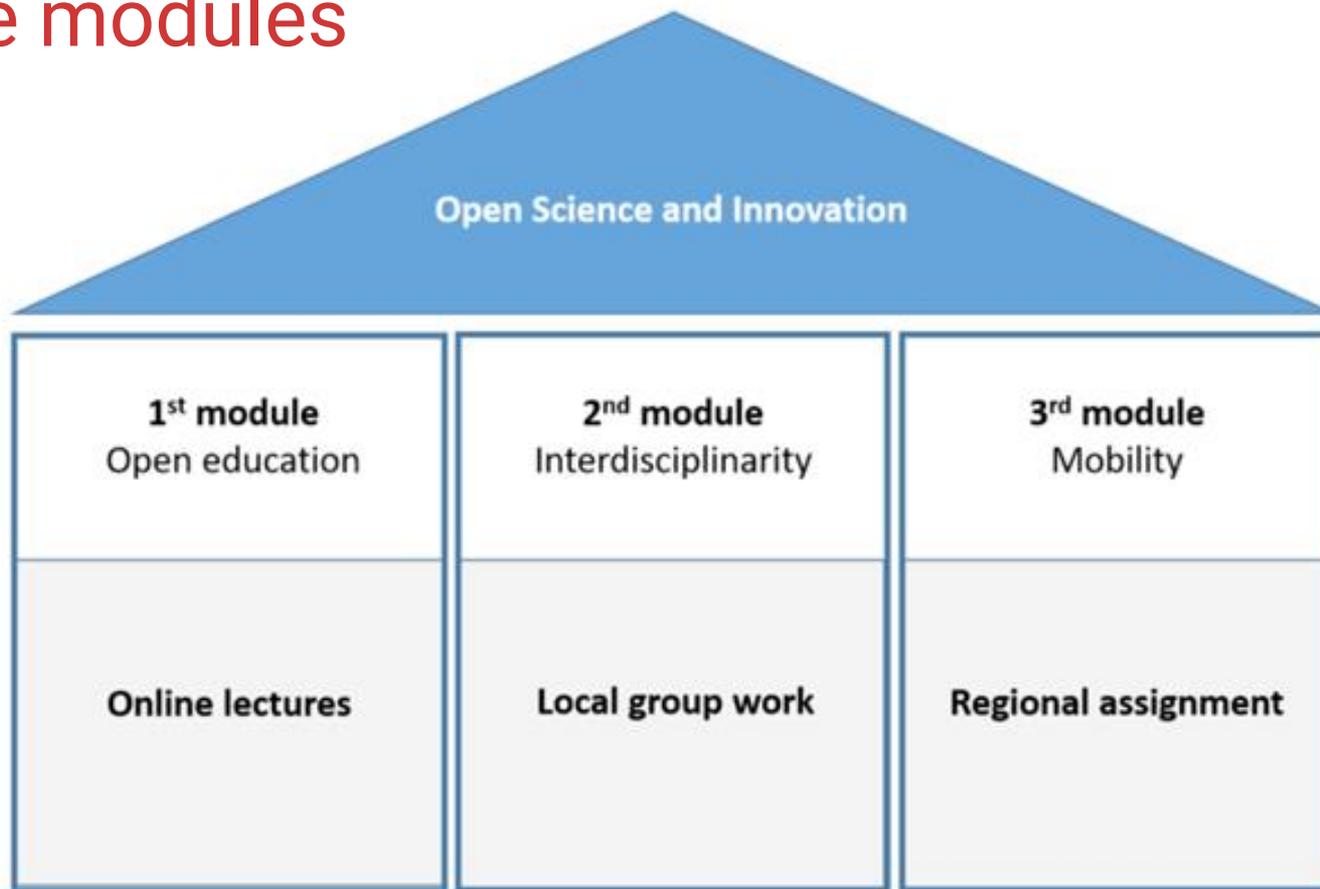
# Three modules

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

# 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 »

# Module 1

## DocEnhance Data Stewardship course: Module 1

Home / My courses / DocEnhance Data Stewardship course: Module 1 / 5. How to structure and document research data / 5.3 The Readme-file

## 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.

### E-learning VŠCHT

- DocEnhance Data Stewardship course: Module 1
- Participants
- Badges
- Competencies
- Grades
- Home
- Dashboard

### How to write a Readme-file
What should be included?

✔ - Contact information to the dataset author
  - Include also author's ORCID
✔ - What the data set is about
  - Write a short description
  - Include information on project relations
✔ - 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 are 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



UNIVERSITY OF
CHEMISTRY AND TECHNOLOGY
PRAGUE

29/04/21

This certification is awarded to

**Anna Chmelová**

in recognition of the successful completion of the

**DocEnhance Data Stewardship course Module 1**

DocEnhance has received funding from the European Union's Horizon 2020 Science with and for Society programme under grant agreement no. 872483

# 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 (via Teams) + extra event + wrap up
- Break out rooms - working in small groups

| Description | 24 Monday | 25 Tuesday | 26 Wednesday |
|---|---|---|---|
| | | 8:00-09:45 | 8:00-09:45 |
| Session | | #3: Structuring, documenting & metadata | #4: Data storage, file formats & FAIR principles |
| Main lecturer | The Module 2 starts at 10:00 on Monday | Ivan Čmelo | Marcela Dendisová |
| Lecturer 2 | | Jan Vrba | Eva Hnatkova |
| | 10:00-12:00 | 10:00-12:00 | 10:00-12:00 |
| Session | #1: The importance of RDM & open archiving | Czech Open Science Day @ National Technical Library | #6: Visualising, reusing/searching & citing research data |
| Main lecturer | Eva Hnatkova | P. Doleček | Radek Cibulka |
| Lecturer 2 | | Please use the following link to connect to this event: | Eva Svobodová |
| | | https://www.techlib.cz/en/84427-czech-open-science-day | |
| | 13:00-15:00 | 13:00-15:00 | 13:00-15:00 |
| Session | #2: The data management plan | #5: Sharing & licensing research data | #7: Wrap Up |
| Main lecturer | Eva Hnatkova | Jiří Jirát | Prof. Pavel Matějka |
| Lecturer 2 | Robert Pergl | | |

# Example DMP

## DSW online tool

# Example: Data storage & FAIR principle - Zenodo

**Upload type** <span style="float:right">required ▾</span>

| 📄 Publication ○ | ▦ Poster ○ | 👥 Presentation ○ | ▦ Dataset ⦿ | 📊 Image ○ | 🎬 Video/Audio ○ | </> Software ○ | 🎓 Lesson ○ | 📦 Physical object ○ | ✳ Other ○ |

**Basic information** <span style="float:right">required ▾</span>

▥ **Digital Object Identifier**
> e.g. 10.1234/foo.bar

Optional. Did your publisher already assign a DOI to your upload? If not, leave the field empty and we will register a new DOI for you. A DOI allows others to easily and unambiguously cite your upload. Please note that it is NOT possible to edit a Zenodo DOI once it has been registered by us, while it is always possible to edit a custom DOI.

▥ Reserve DOI

📅 **Publication date** *
> 2021-06-24

Required. Format: YYYY-MM-DD. In case your upload was already published elsewhere, please use the date of first publication.

▤ **Title** *
> 

Required.

👤 **Authors** *
> Family name, given names   |   Affiliation   |   ⓘD ORCID (e.g.: 0000-0002-1825-0097)   ↕ ×

Optional.

➕ Add another author

✏ **Description** * | 🗐 🗐 **B** *I* S̶ x₂ x² 🔗 ⦸ | 📶 ⠿ → ¶ 🗟 | ↶ ↷ I̶ₓ | Σ Ω | 📄 Zdroj | ⛶

## License

**Access right** * 
- ⦿ 🔓 Open Access
- ○ ⊘ Embargoed Access
- ○ 🔑 Restricted Access
- ○ 🔒 Closed Access

Required. Open access uploads h

✴ **License** *
> Creative Commons Attributi

# 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 Křižan** (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 Tkadlecová**

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

**Jaroslav Havlíček**
**Tomáš Pekárek**

# 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

# 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



https://phdhub.eu/

The aim is supporting innovative career oriented PhD training

**DocEnhance**

# Data Stewardship

**to be officially released by the end of 2022**

DocEnhance

# Horizon Europe - funding for OS

**European Research Executive Agency** ✓
@REA_research

🔵6🔵1 million Euros are available to fund projects that will reform 🇪🇺 #research & #innovation systems.

Check the 🆕#HorizonEU calls:
🌐#Openscience
🗣#Citizenscience
♂♀#Gender equality
⚖#RRI

Submit your proposal by 23 September!
#EUResearchArea

Horizon Europe funding for

**Reforming & enhancing the European R&I system**

Apply by 23 September 2021

#HorizonEU   #EUResearchArea

https://rea.ec.europa.eu/news/first-horizon-europe-calls-proposals-are-now-open-2021-06-22_en