

NTK

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

Pilotování “Data Stewardship” kurzu pro doktorandy

23. 11. 2023, EOSC-CZ Seminář, online

Eva Hnátková, Martin Schätz, Mili Losmanová

eva.hnatkova@techlib.cz



Harmonogram

- **09:30 - 10:30 Sekce 1**
- 10:30 - 10:45 Coffee Break
- **10:45 - 12:00 Sekce 2**
- 12:00 - 13:00 Obědová pauza
- **13:00 - 14:00 Sekce 3**

Technikálie

- mikrofon, video, dotazy, chat, nahrávání
- prezentace bude dostupná

Průzkum účastníků

- Podpora doktorského studia
- Doktorandi/vědci
- Data stewardi
- OS manažeři / koordinátoři / knihovníci

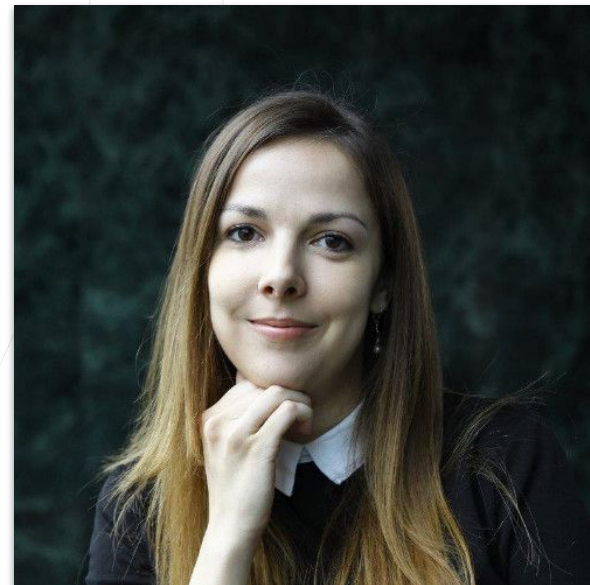
O nás



Eva Hnátková 
Open Science



Martin Schätz 
Research Data Management



Mili Losmanová 
PhD Support

Témata

- Představení DocEnhance
- Data stewardship kurz pro doktorandy
- Teoretická a praktická část kurzu
- Spolupráce s datastewardy
- Komunikace a organizace kurzu
- Spolupráce s průmyslem

Představení DocEnhance



DocEnhance



DocEnhance

How DocEnhance started ...

The role of professionals in DE

7) Involved in projects related to doctoral training

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)

Vienna, April 2018

Shaping Doctoral Education – the Role of Professionals at Universities

**PRIDE
Network**

ASSOCIATION FOR
PROFESSIONALS IN
DOCTORAL EDUCATION



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



DocEnhance



Partners of DocEnhance project during the kick-off meeting in Tromsø

DocEnhance

- EU project 2020 - 2022
- 19 partners



UiT The Arctic
University of Norway



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



Why DocEnhance?

- PhD holders more likely to follow a **career outside the university** as numbers of PhDs on the rise (OECD, Education at a Glance, 2019; Kehm, 2007)
- Reported mismatches between employers' needs and skills developed in doctoral programmes (DeGrande et al, 2014; Bebiroglu et al, 2022)
- ***"Offering training in transferable skills ... is central, and should be a priority for doctoral schools and programmes"*** (Salzbourg II recommendations, EUA, 2010)
- ***"Identifying missing elements of skills training remains an important task..."*** (Survey "Doctoral Education in Europe today", EUA-CDE, 2022)





DocEnhance

DocEnhance aims

- Collect **evidence on PhDs' employment** and **skills** usage
- Develop and integrate **transferable skills training** into **existing PhD programmes**
- Increase **interaction with non-academic sector**



**Broadening PhD expertise
through transferable skills**



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



DocEnhance

DocEnhance outcomes

- PhD career-tracking survey
 - Transferable skills courses
- } will be **freely available** at the DocEnhance platform

DocEnhance resources are designed to be easily adapted & integrated at your institution



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



DocEnhance

PhD Career-Tracking Survey

+ good practice guide
for implementation



DocEnhance
Deliverable 1.2

Authors: Julia Boman,
Harry Beeson, Mabel
Sanchez Barrioluengo,
Mihaela Rusitoru

Organisation responsible:
European Science
Foundation (ESF)

What comes after a PhD?

**Findings from the DocEnhance
survey of doctorate holders
on their employment
situation, skills match,
and the value of the doctorate**

6 December 2021



Report on career-tracking of PhD graduates



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



PhD Career-tracking survey

Questionnaire

- based on ESF career-tracking surveys + developed with DocEnhance partners

Target population

- PhD graduates between 2016 and 2020 from **9 universities** (NL, DE, NO, GR, ES, PT, SK, IT, CZ)

Survey

- **2,217** responses collected (response rate: **23%**)



DocEnhance

PhD Career-tracking survey



PhD employment by sector



All findings from DocEnhance Career-tracking survey at www.docenhance.eu

37%

in universities

22%

in private sector

12%

in research organisations

10%

in health care

8%

in the government

11%

in other sectors



DocEnhance

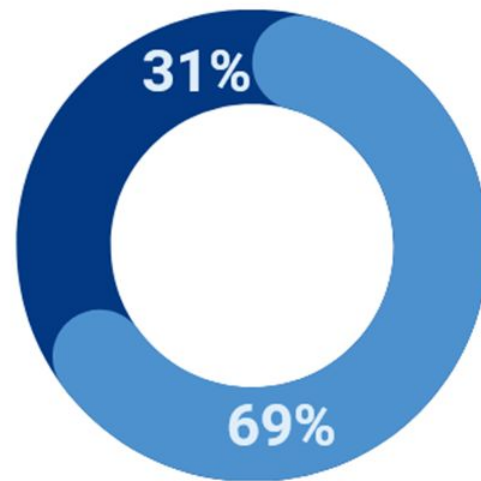
PhD Career-tracking survey



How many PhD holders are engaged in research in their current job?



All findings from DocEnhance Career-tracking survey
at www.docenhance.eu



-  Work in research
-  Do not work in research



DocEnhance

Satisfaction with doctoral training

(1 = very dissatisfied, 5 = very satisfied)



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



DocEnhance

Transferable skills courses



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



DocEnhance

Design of 3 pilot courses

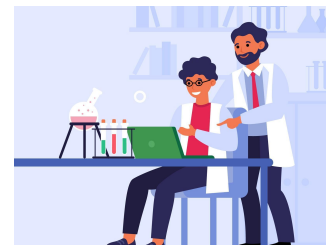
on transferable skills for doctoral candidates/supervisors



- Data Stewardship



- Career management
& entrepreneurship



- Supervision

All modules include exam tests and certification



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



Data stewardship

Aims of this course: to introduce PhD candidates the complexity of research data management from finding relevant data to publishing an own dataset

PhD candidates, learn:

- How to develop and use a Data Management Plan
- How to store and archive data
- How to use licenses, repositories, database platforms, data search engines, apps...



...and much more!



Career management & entrepreneurship

Aims of this course: to acquire PhD candidates the competencies needed to advance their career and to become more entrepreneurial.

PhD candidates, learn:

- How to build a personal branding
- How to build a career plan and define personal development goals
- How to communicate and coach others effectively...



...and much more!



PhD Supervision

Aims of this course: high-quality PhD supervision is the crucial tool that supports impactful doctoral education and prepares PhD candidates to contribute to research and society.

PhD Supervisors, learn:

- How to build a professional relationship with doctoral candidates
- How to define roles, functions and expectations through the Supervision process
- How to use Supervision models, tools, resources...

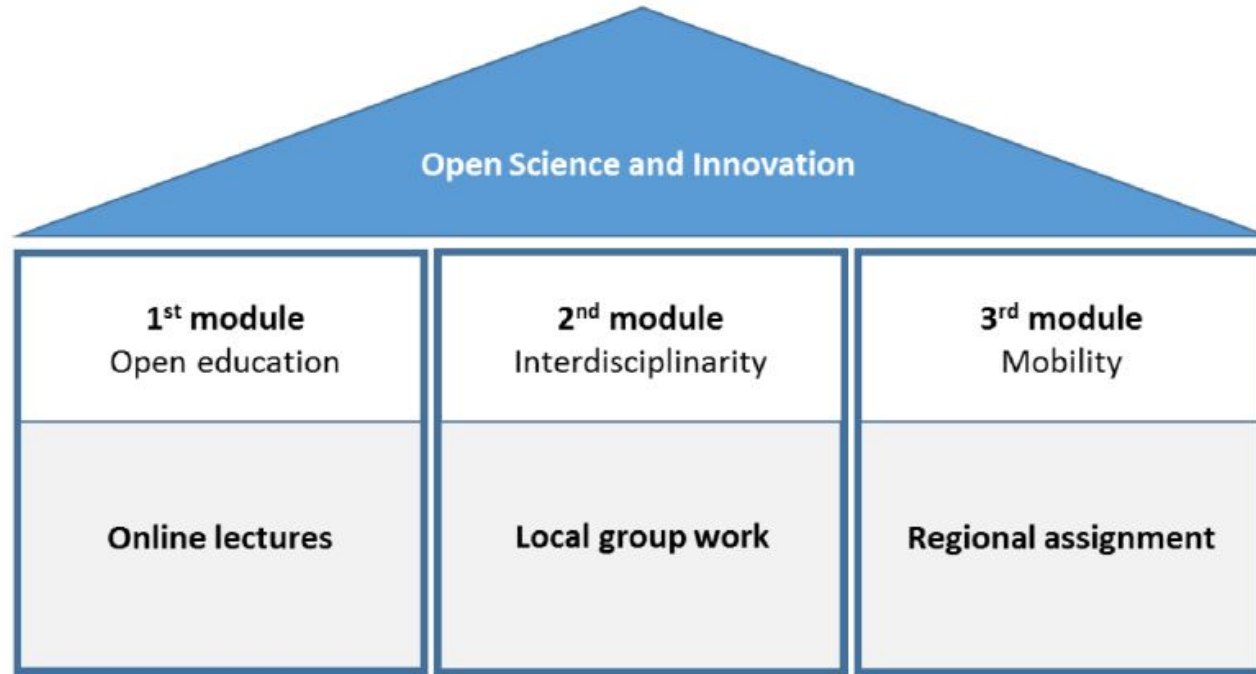


...and much more!



DocEnhance

A Novel Course Concept



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

Data stewardship kurz pro doktorandy



DocEnhance

Data stewardship kurz

Vyvinula skupina odborníků na Research Data Management (RDM) na UiT, kteří se podílejí na výuce RDM na této instituci



UiT The Arctic
University of Norway

Tříčlenná pracovní skupina
za vše odpovídá

- [Helene N. Andreassen](#)
- [Henning Hansen](#)
- [Leif Longva](#)



Pilotování Data stewardship kurzu

Vysoká škola chemicko-technologická v Praze



~ 25 doktorandů (doctoral candidates)

~ 8 lektorů

1. Pilot - finanční odměna za aktivní účast
a následnou zpětnou vazbu

Karlstad University, Švédsko

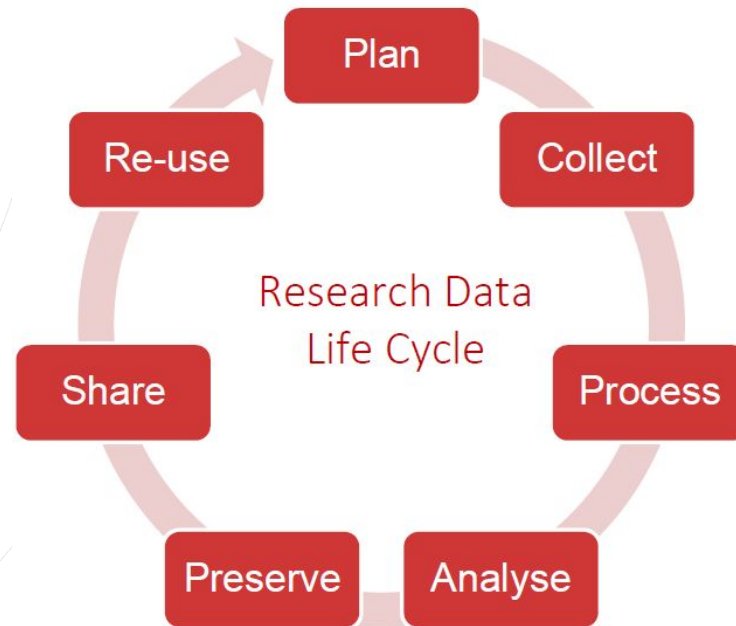


~ 25 doktorandů



Cíle Data Stewardship kurzu

- **seznámit** doktorandy s komplexností a významem správy dat
- **projít** všemi kroky správy výzkumných dat - od vyhledání relevantních dat až po zveřejnění vlastní datové sady
- **pomoci** doktorandům s vypracováním disertační práce a lépe je připravit na práci mimo akademickou sféru po ukončení doktorského studia
- **propojit** praktická cvičení s teoretickým obsahem



Data Stewardship course

MODUL 1	MODUL 2	MODUL 3
Teoretická část - online (Moodle)	Praktické cvičení a úkoly v pracovních týmech	Workshopy a regionální úkoly u zaměstnavatelů
• obecný úvod • základy zacházení s výzkumnými daty v doktorské práci	• učení se praxí • praktické úkoly • práce v malých týmech • diskuse • místní zdroje (DMP) • splnění místních požadavků #Interdisciplinarita #Týmová spolupráce	• praktické úkoly ve spolupráci s veřejným/soukromým sektorem • příprava na kariéru mimo akademickou sféru • uplatnění teorie a dovedností v praxi #Mezisektorová spolupráce

Modul 1: tematické sekce

Zaměřené na motivaci, teorii a koncepty (~ 16 hodin)

1. Introduction
2. About research data
3. How to search & cite research data
4. How to structure & document research data
5. How to store research data
6. Data cleaning, analyzation and visualisation
7. How to archive research data
8. Rights and licenses for research data
9. How to write a research DMP
10. Summary and exam quiz

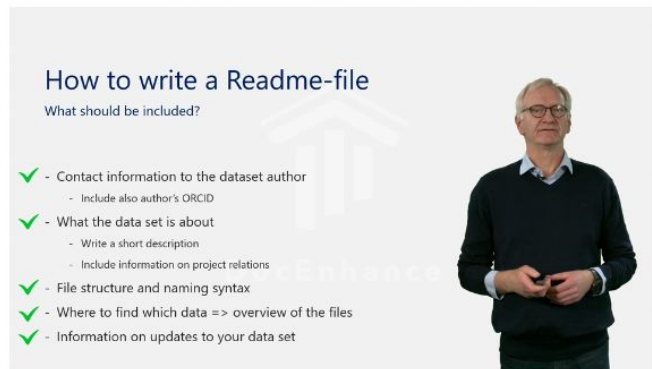
- **Video, doporučená zdroje**
- **Interakce prostřednictvím kvízů a otevřených "podnětů k zamyšlení"**

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.



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?

Modul 1: zkušební kvíz a certifikace

Závěrečný kvíz

pokud získáte 80 % nebo
24/30 správných odpovědí



Certifikát z Modulu 1
vygenerovaný v Moodle



Modul 1 dostupný na webu NTK

- Aktuálně máme nainstalovaný Modul 1
<https://moodle.techlib.cz/course/view.php?id=179>
- Je to otevřené pro kohokoliv
- Pokud uživatel chce udělat test a získat certifikát, musí se zapsat do kurzu.
- Pro zápis do kurzu - pokud je to registrovaný uživatel NTK, tak použije své údaje. Pokud ne, vyplní krátký registrační formulář. Více je na [How to enrol in a course](#) in NTK Moodle
- Administrator kurzu vidí záznamy o zapsaných účastnících (návštěvnost, známky za testy apod).



Modul 1 dostupný na webu NTK



0. Welcome

0.1 About the DocEnhanc...

0.2 About the Data Stewa...

0.3 How to use the online ...

0.4 Reuse conditions

0.5 About the teachers

0.6 Contact information

0.7 Acknowledgements

0.8 Announcements

1. Introduction

1.0 About: Introduction

1.1 The importance of res...

1.2 Transparency and ope...

0. Welcome

[Collapse all](#)

The online Data Stewardship course provides an introduction to research data management. It has been created within the framework of the international, EU funded project [DocEnhance](#) project by a working group of research data management specialists at UiT The Arctic University of Norway. It is provided under [CC BY 4.0](#) license.

The main target group for this course is PhD candidates at universities, research institutes and university colleges. It is primarily designed to function as a preparatory task for interactive seminars with peers, as well as collaborations with partners from the non-academic sector. It may equally well function as an independent teaching and learning resource, and the course is therefore open for anyone who wishes to improve their competences and skills in research data management.

The course can be taken at any point of time or place. To gain access to the final exam, which enables you to receive a certificate, please enrol in the course.

[How to enrol in a course in NTK Moodle](#)

Registered NTK patrons can register by first logging into Moodle using their NTK credentials (upper right corner - *Log in - Log in using your account on: NTK Shibboleth*) and then clicking the *"Enrol me in this course"* option from the course's main menu, confirming with the blue *"Enrol me"* button.

If you are not registered with NTK, first fill in this [form](#). After confirming your email address, log into Moodle (upper right corner - *Log in - enter Username and Password - Log in* button) and then click the *"Enrol me in this course"* option from the course's main menu, confirming with the blue *"Enrol me"* button.

If you have any questions about this course, please contact us at: reference@techlib.cz.

Modul 1 dostupný na webu NTK

How to enrol in a course in NTK Moodle

1. [For registered NTK patrons](#)
2. [For patrons who are not registered in NTK](#)

For registered NTK patrons

Registered NTK patrons can use their NTK credentials to enrol in a course.

1.

The screenshot shows the NTK Moodle interface. At the top, there is a header with the NTK logo, a search bar, a language selector set to 'English (en)', and a message 'You are currently using guest access'. A red box highlights the 'Log in' button, with a red arrow pointing to it from the right. Below the header, the course title 'Data Stewardship – Module 1: MOOC' is displayed. Underneath, there is a section titled '0. Welcome' with a 'Collapse all' link. The main content area contains a paragraph about the course: 'The online Data Stewardship course provides an introduction to research data management. It has been created within the framework of the international, EU funded project DocEnhance project by a working group of research data management specialists at UiT The Arctic University of Norway. It is provided under CC BY 4.0 license.'

Co plánujeme dál?

- Na Moodlu NTK bude stránka, kde budou umístěny pro stažení materiály (všechny 3 Moduly) + další DocEnhance kurzy
- Soubory MBZ (formát Moodle) je jednoduché nahrát do standardního rozhraní Moodle.
- Měl by to zvládnout každý administrator Moodle.
- Je potřeba dodatečně připojit [plugin "mod_customcert"](#), který pomáhá přizpůsobit certifikát.
- NTK dodalo svoje logo a certifikát se vygeneruje automaticky každému uživateli po absolvování závěrečného testu v Modulu 1.
- U Modulu 2 a 3 by mělo jít také nastavit generování Certifikátu po splnění některých povinností.

NTK

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

Děkuji za pozornost

Otevřená diskuse

Eva Hnátková

eva.hnatkova@techlib.cz



Harmonogram

- **09:30 - 10:30 Sekce 1**
- **10:30 - 10:45 Coffee Break**
- **10:45 - 12:00 Sekce 2**
- **12:00 - 13:00 Obědová pauza**
- **13:00 - 14:00 Sekce 3**

Praktická část kurzu a spolupráce s datastewardy

Martin Schätz

martin.schatz@techlib.cz

Agenda

- Přehled Modulů Data Stewardship kurzu
- Vstupní podmínky pro Modul 2
- Přehled Modul 2
- Přehled a požadavky jednotlivých sekcí modulu 2
- Data Stewardship kurz jako možnost školení Data Stewardů

Data stewardship - Modul 2

MODUL 1	MODUL 2	MODUL 3
online lekce	Workshopy a úkoly v pracovních týmech	Workshopy a regionální úkoly u zaměstnavatelů
• obecný úvod • základy zacházení s výzkumnými daty v doktorské práci.	• učení se praxí • praktické úkoly • práce v malých skupinách • diskuse • Data Steward. Wizard • místní zdroje • splnění místních požadavků	• praktické úkoly ve spolupráci s podnikatelským subjektem / subjektem veřejných služeb. • příprava na kariéru mimo akademickou sféru. • uplatnění teorie a dovedností v praxi

Kontrolní seznam

Kontrolní seznam pro lektory:

- Potvrzení o absolvování Modulu 1: Ověřte certifikáty získané v modulu 1.
- Přístup k testovacímu repozitáři.
- Pochopení požadavků na lektory.

Přípravné práce účastníků:

- Znovu si (případně) projít videa z Modulu 1.
- Dokončit předběžné úkoly: Pro některé sekce je potřeba se připravit dopředu.

Kontrolní seznam

Modul 1 Zkouška a certifikace

- Modul 1 je dobré nechat otevřený po celý rok.
- Ne všichni kdo získají certifikát pokračují do Modulu 2.



Data stewardship - Modul 2

Workshopy a zadání

- Materiály pro lektory na platformě Moodle.
 - Místní lektori mohou rozhodovat o tom, který obsah a aktivity použijí.
 - 6 seminářů (2-3 hodiny)
 - Účastníci pracují v malých skupinách.
- 
- Úvod (pro lektory)
 - Sekce 1: Význam RDM a otevřené archivace
 - Sekce 2: Plán správy dat
 - Sekce 3: Strukturování a dokumentace výzkumných dat
 - Sekce 4: Ukládání dat, formáty souborů a zásady FAIR pro data
 - Sekce 5: Archivace výzkumných dat
 - Sekce 6: Vyhledávání, citování a licencování výzkumných dat
 - Zadání zkoušky

Přehled Modulu 2

Každé sekce má následující strukturu:

- Přehled sekce s časovým rámcem.
- Pokyny pro lektora sekce.
- Texty úkolů a další materiály.

Formát

- Sekce mohou probíhat buď fyzicky, nebo prostřednictvím digitální platformy.
- Sekce jsou především o aktivním zapojování účastníků.
- Úkolem lektora je zprostředkovat aktivity a diskuse.
- Účastníci mají sdílet informace o svém vlastním výzkumu a výzvách souvisejících s daty.
- Pro účastníky je připraven průvodce modulem.

#1 Význam RDM a otevřené archivace

Cíle sekce:

- Seznámit se s různorodostí požadavků práce s daty v různých oborech.
- Prodiskutovat specifika otevřené archivace v různých oborech.
- Prodiskutovat tyto požadavky z hlediska interdisciplinarity.
- Prodiskutovat hlavní obavy týkající se archivace otevřených dat.

Požadavky na lektora:

- Připravit se na diskuzi obavy týkající se archivace otevřených dat (podkladový materiál).
- Udělat si přehled o vědeckých oborech na instituci.

#1 Význam RDM a otevřené archivace

- Účastníci představí sebe, své projekty, typ dat, na kterých pracovali, a případné problémy se správou dat.
- Skupinová diskuse o:
 - Transparentnosti, archivaci otevřených dat a FAIR dat v různých oborech.
 - Otázky mezioborové spolupráce.
 - Vyjasnění terminologie.
 - Obavy týkajících se archivace otevřených dat.
- Skupinová diskuze o společných obavách, zhodnocení obav a argumentů pro nebo proti.
- Výstup: poznámky z diskuse o všech klíčových obavách.

#2 Plán správy dat

Cíle sekce:

- Naučit se interpretovat a zhodnotit Plán Správy Dat.
- Naučit se tvořit Plán Správy Dat.

Požadavky na lektora:

- Mít praktickou zkušenost s tvorbou DMP.
- Znalost práce s alespoň jedním nástrojem tvorby DMP.
- Předpřipravit si knowledge model/template a přístup do nástroje.
- Mít přehled o možnostech a komunitních standardech pro vědecké obory instituce.
- Mít přehled o požadavcích grantů.

#2 Plán správy dat

- Praktická část v DMP nástroji.
- V našem případě vlastní instance nástroje DSW.



- Online výuka: zkrácený Horizon Europe Template
- Prezenční výuka: celý Horizon Europe Template

- Co je Data Stewardship Wizard?
 - Online kolaborativní nástroj pro vyplnění DMP.
 - Filtrování otázek podle typu templatu.
 - Systém uzavřených otázek.
 - Zvýrazňování otázek podle fáze projektu.
 - Propojení na ROR, FAIRsharing, a další ...
 - Různé formy výstupu.

#2 Plán správy dat

DS Wizard

Users

Knowledge Models

Projects

List

Importers

Documents

Settings

Albert Einstein
Admin

« Collapse sidebar

My Experiment

Questionnaire

Metrics

Preview

Documents

Settings

View

Import answers

Comments 4

TODOs 1

Version history

Current Phase

Before Submitting the Proposal

Chapters

I. Administrative information1

II. Re-using data4

III. Creating and collecting data7

IV. Processing data4

V. Interpreting data2

VI. Preserving data4

VII. Giving access to data3

IV. Processing data

In the processing phase, the data will be undergoing the mostly automated steps for processing, before the analysis and interpretation.

In this chapter, many questions are focusing on the compute environment that is used to process the data and make it available for interpretation by project partners. Some of those questions (e.g. on workflow systems and data provenance) are also relevant for the work in the interpretation phase.

1 Will you be using a shared working space to work with your data?

Horizon 2020 DMP

Horizon Europe DMP

Science Europe DMP

Will you be using a working space containing data and software specific to the project that is shared between all the people working on the data in the project? Sometimes such a system is called a *Virtual Research Environment*.

☒ Desirable: Before Submitting the Proposal

☐ a. No, participants in the project each have different collections of data and tools

☒ b. Yes

Clear answer

Answered 5 minutes ago by Albert Einstein.

1.b.1 Will this work space be run by dedicated specialists?

TODO

If your work space is run and maintained by specialists, e.g. the ICT department of one of the institutes involved in the projects, this means that backup and restore as well as access management is properly addressed.

☒ Desirable: Before Submitting the DMP

Named versions only

October 2022

▼ 4, 10.

9:38

Current

1.0.0

Will this work space be run by dedicated specialists?

☐ Yes

Albert Einstein

9:38

Will you be using a shared working space to work with your data?

☐ Yes

Albert Einstein

9:37

Data format/type

XML Query Language

Albert Einstein

9:34

Cleared reply of Data set:

#2 Plán správy dat

FAIR metriky

Metric	Measure	
Findability	0.70	
Accessibility	0.22	
Reusability	1.00	
Good DMP Practice	0.67	
Openness	0.69	

FAIR - zpětná vazba

9 Will you need consent for any newly collected personal data? + !

☒ a. No, We do not collect any new personal data
Reusability

☐ b. No, our work on personal data can be done using another legal base ☰
Reusability

☐ c. Yes, we will collect consent for our specific use of the data
Reusability

☐ d. Yes, we will collect consent for our use as well as reuse of the data
Reusability

#3 Strukturování a dokumentace dat

Cíle sekce:

- Zhodnocení datasetu z hlediska strukturování, dokumentace a verzování datasetu.
- Hodnocení důležitosti metadat a správné dokumentace.
- Naučit se správnému strukturování a dokumentace datasetů.

Požadavky na lektora:

- Mít zkušenosti nebo přehled o práci s různými druhy dat (data z průzkumů, data ze sčítání lidu, agregovaná data, klinická data, ...).
- Mít praktickou zkušenost se znovupoužitím datasetů.
- Připravit si publikvaný dataset ke zhodnocení.

#3 Strukturování a dokumentace dat

- Předpříprava: zamyšlení se nad konceptem metadat, strukturování a dokumentování výzkumných dat ve vlastním projektu.
- Skupinové diskuze:
 - Metadata ve vlastních projektech.
 - Zhodnocení vybraného datasetu z hlediska metadat, strukturování a dokumentace.
- Plenární diskuze: Důležitost “readme” souborů, jejich obsahu vzhledem k strukturování a dokumentace publikovaných dat.

#4 Ukládání dat, formáty souborů a FAIR

Cíle sekce:

- Pochopit nároky na úložiště z hlediska aktivního používání dat.
- Naučit se plánovat využití jednotlivých typů úložišť.
- Naučit se hodnotit datové formáty z hlediska dlouhodobého využití.
- Naučit se používat FAIR principy k hodnocení práce s daty.

Požadavky na lektora:

- Mít zkušenost s prací s daty z vědeckého projektu (aktivní používání, záloha, archivace).
- Mít zkušenost s prací s jak otevřenými tak proprietárními datovými formáty.
- Mít přehled o doporučených datových formátech pro vědecké obory.

#4 Ukládání dat, formáty souborů a FAIR

- Předpříprava: Zhodnocení vlastního projektu z hlediska citlivosti dat, nároků úložiště, řízení přístupu k datům, záloh dat, datových formátů. Krátká esej na téma proč používat FAIR principy.
- Skupinové diskuze:
 - Představení a zhodnocení ukládání dat ve vlastních projektech.
 - Představení a zhodnocení používaných datových formátů ve vlastních projektech.
 - FAIR principy pro data.
- Skupinová nebo plenum diskuze:
 - Zhodnocení datových formátů ve vybraných datasetech.
 - Oborové zvyklosti a standardy z hlediska datových formátů.

#5 Archivace výzkumných dat

Cíle sekce:

- Naučit se jak připravit soubor dat k publikaci.
- Naučit se jak najít vhodný repozitář pro daný soubor dat.
- Při praktickém cvičení budou účastníci pracovat na souborech dat, které jsou buď vytvořeny, nebo upraveny pro vzdělávací účely.
- Naučit se publikovat data v “trénovacím” repozitáři.
- Naučit se vyhledávat vhodné repozitáře.

Požadavky na lektora:

- Připravit si a naučit se/mít zkušenosti např. s Zenodo sandboxem.
- Mít zkušenost s vyhledáváním a hodnocením nejen oborových repozitářů.
- Připravit tréninkový dataset.

#5 Archivace výzkumných dat

- Předpříprava: Podle instrukcí lektora se zaregistrovat do tréninkové verze vybraného repozitáře. Nastudovat si návod pro používání vybraného repozitáře. Nastudovat si požadavky na přípravu dat (pojmenovávací konvence, datový formát, ...)
- Skupinové úkoly:
 - Předpřípravit dataset k publikaci (přejmenování, konvertování dat, ...)
 - Vyplnění metadat v repozitáři.
 - Nahrání dat do repozitáře.
- Diskuze:
 - Vzniklé problémy a otázky při publikování testovacího datasetu.
 - Výběr vhodného repozitáře pro předpřípravené datasety.

#5 Archivace výzkumných dat



Why use Zenodo?

- **Safe** — your research is stored safely for the future in CERN's Data Centre for as long as CERN exists.
- **Trusted** — built and operated by CERN and OpenAIRE to ensure that everyone can join in Open Science.
- **Citeable** — every upload is assigned a Digital Object Identifier (DOI), to make them citable and trackable.
- **No waiting time** — Uploads are made available online as soon as you hit publish, and your DOI is registered within seconds.
- **Open or closed** — Share e.g. anonymized clinical trial data with only medical professionals via our restricted access mode.
- **Versioning** — Easily update your dataset with our versioning feature.
- **GitHub integration** — Easily preserve your GitHub repository in Zenodo.
- **Usage statistics** — All uploads display standards compliant usage statistics

About
About
Policies
Infrastructure
Principles
Projects
Roadmap
Contact

Blog
Blog

Help
FAQ
Docs
Guides
Support

Developer
s
REST API
OAI-PMH

Contribute
GitHub
Donate



Funded by

#5 Archivace výzkumných dat

**UNIVERSITY OF
CHEMISTRY AND TECHNOLOGY
PRAGUE**

UCT Prague Zenodo Community
<https://www.vscht.cz/> [University of Chemistry and Technology ROR](#)



[Records](#) [Members](#) [About](#)

11 results found

Sort by Newest

Versions

☐ View all versions

Access status

☐ Open

☐ Restricted

Resource types

☐ Publication

☐ Dataset

☐ Software

☐ Presentation

September 23, 2023 (1.0.0) Presentation Open

Research Data Managemet and how not to get overwhelmed with data

Martin Schätz 

Research data management and how not to get overwhelmed with data presentation is an overview of bioimage analysis with a focus on the basics for data management planning, FAIR principles, and how to practically organize folders and prepares naming convention. The presentation includes an overview of metadata, Creative Common licenses, and a sum...

10 Uploaded on September 23, 2023

1 more versions exist for this record

346 115

June 21, 2023 (v1) Dataset Restricted

REPARES interlaboratory study (ring test) on qPCR quantification of antibiotic resistance genes and other related genes in DNA isolated from sewage sludge and wastewater

Bartacek Jan ; Purkrtova Sabina ; Hanakova Katerina

This dataset collects data on interlaboratory study (Ring test), performed within project REPARES in two phases (Ring test phase 1, Ring test phase 2). Sampling and analyses for phase 1 started on 25.2.2020 and for phase 2 on 27.9.2021. The sampling site was a wastewater treatment plant (> 100 000 inhabitants) in the Czech Republic. REPARES: R...

4 3 3 1

Uploaded on June 21, 2023

45 0

#6 Vyhledávání, citování a licencování dat

Cíle sekce:

- Naučit se vyhledávat datasety relevantní pro vlastní projekty.
- Diskutovat nad propojením a citováním dat ve vědeckých publikacích.
- Naučit se vybírat a hodnotit licence přidělené datovým souborům.
- Porozumět důsledkům které licence mají pro opakované použití datových souborů.

Požadavky na lektora:

- Mít zkušenost s publikováním dat jak samostatně tak v rámci článků.
- Rozumět různým druhům licencím pro data a software.
- Mít zkušenost s vyhledáváním a znovu použitím publikovaných dat.

#6 Vyhledávání, citování a licencování dat

- Skupinové úkoly:
 - Vyhledávání výzkumných dat, která jsou relevantní projekty.
 - Diskuze nad různými způsoby citování a odkazování na soubory dat.
 - Zhodnocení vyhledávacích prostředí.
 - Vyhodnocení jednotlivých licencí přidělené datasetům a důsledků, které mohou mít pro opakované použití datasetů.
- Skupinové úkoly nebo plénum, Etika výzkumu:
 - “The Dilemma game” - diskuze nad rozhodnutím u představených dilema na témata: sdílení přístupu k nepublikovaným datům, práce s chybami v datech, rozhodování o spoluautorství při použití mnou získaných dat, ...

Ověření znalostí Modulu 2

- Téma: Výběr vhodného úložiště výzkumných dat.
 - Rozsah: 1000-1500 slov (včetně odkazů)
- Úkolem je vyhledat a vybrat datové úložiště pro data, která (budou) použita v reálném projektu. Práce musí obsahovat následující tři součásti:
 - 1. Popis souboru(ů) dat, např.
 - oblast studia
 - typ(y) dat
 - velikost souboru dat
 - citlivost dat
 - 2. Popis procesu vyhledávání, tj.
 - místa vyhledávání
 - způsob vyhledávání
 - 3. Popis procesu hodnocení, tj.
 - hlavní kritéria použitá při hodnocení
 - argument pro konečné rozhodnutí

Shrnutí Modulu 2

- Modul 2 cílí na to prakticky procvičit teorii z Modulu 1.
- Většina sekcí je závislá na dobré individuální přípravě.
- Sekce nejsou specificky zaměřené na určité vědecké obory.
- Předpřipravené návody pro lektory ulehčují udržení stejného stylu výuky.
- Modul 2 se velmi jednoduše dá podle potřeby zaměřit na jakékoliv vědecké obory, je potřeba ale mít zkušené lektory.
- Naše zkušenost je, že část absolventů celého kurzu se v budoucnu buď zapojí jako lektori, nebo jsou alespoň ochotni sdílet své praktické zkušenosti.

Zajímavosti z praxe

- Rozmanitý mix vědních oborů vedl k mnohem podnětnějším diskuzím.
- V jednotlivých oborech je obrovský rozdíl mezi mírou definice požadavků na Open Science. Většina oborů má minimum - ale to neznamená že se nedá převzít dobrá praxe z pokročilejšího oboru.
- Pokud to jde, je dobré mezi sekcemi dát dostatečný časový odstup. Dobrá příprava je důležitá.
- Výstupy diskuzí a feedback může velmi ulehčit práci Data Stewardům.

Data Stewardship kurz

Jako možnost školení Data Stewardů

Obdobné jako pilotování kurzu

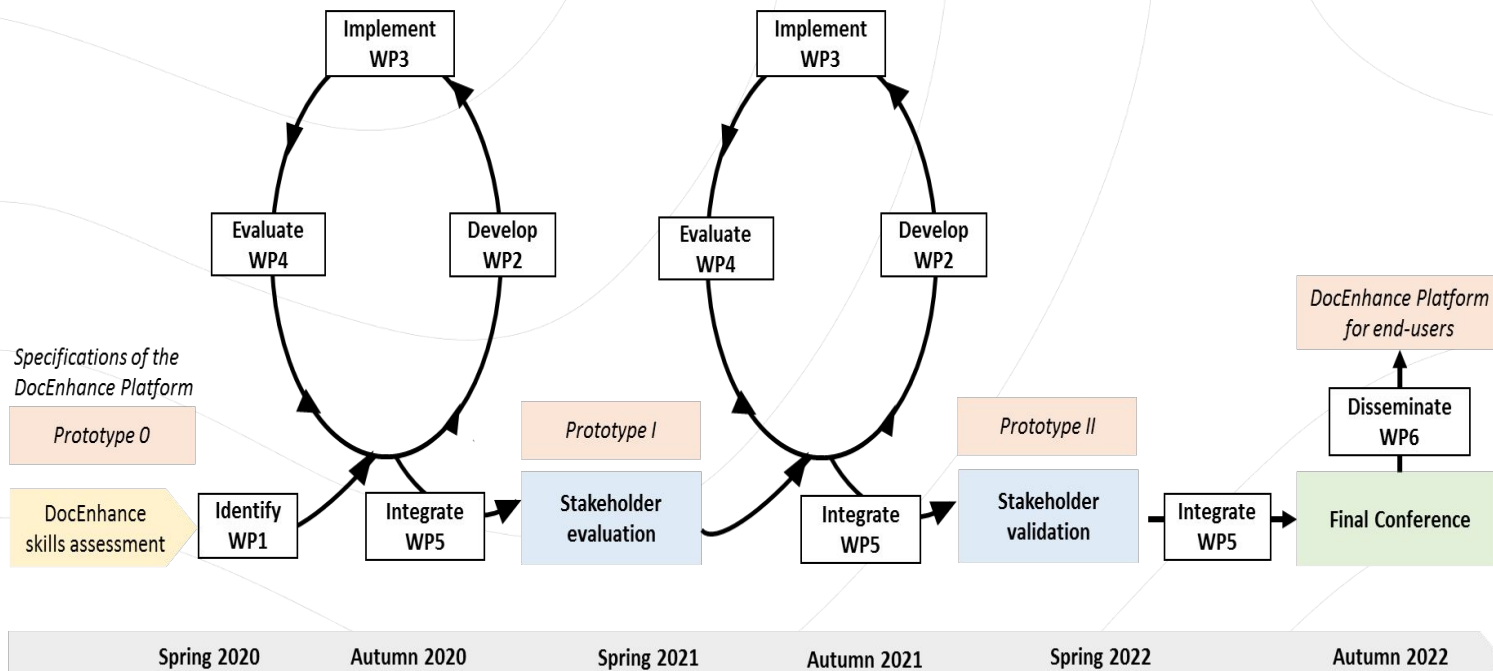
Co-creation approach



Skills intelligence



Skills integration



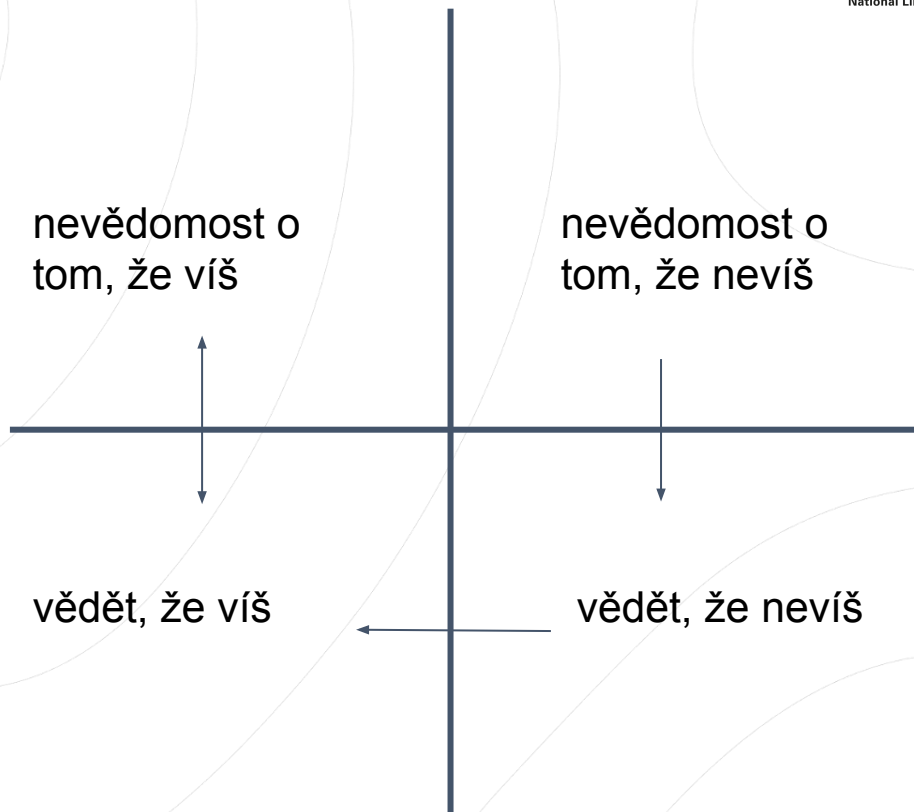
Získávání znalostí

Od účastníka kurzu, k lektorovi

- Sebeuvědomění
- Znalosti

Od lektora k Data Stewardovi

- Praktické znalosti.
- Aktivní zapojování se.



NTK

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

Děkuji za pozornost

Otevřená diskuse

Martin Schätz

martin.schatz@techlib.cz



Harmonogram

- **09:30 - 10:30 Sekce 1**
- **10:30 - 10:45 Coffee Break**
- **10:45 - 12:00 Sekce 2**
- **12:00 - 13:00 Obědová pauza**
- **13:00 - 14:00 Sekce 3**

Komunikace a organizace kurzu, spolupráce s průmyslem

Mili Losmanová

mili.losmanova@techlib.cz

Agenda

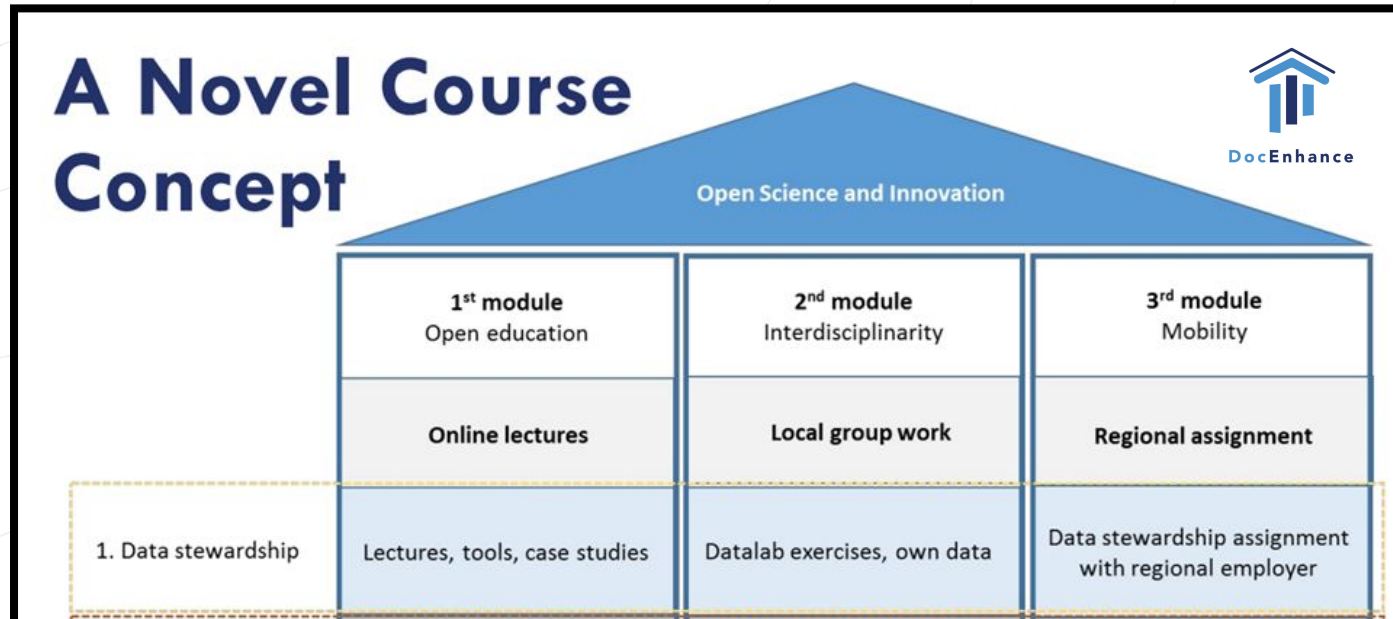
- Organizace kurzu
- Komunikace kurzu
- Spolupráce s průmyslem (Modul 3)
- Feedback doktorandů a lektorů
- Naše doporučení a tipy na pilotáž

Organizace kurzu

Modul 1: online videa + kvízy (otevřený cca měsíc)

Modul 2: práce na univerzitě ve skupinách (6x3h v rámci týdne)

Modul 3: spolupráce s privátním sektorem (2x2h + celodenní výjezd)



Organizace na VŠCHT

Aktuálně již 3x pilotováno

- 2021: kompletně online M1-M3
- 2022: M1 online, M2 - M3 hybridně
- 2023: M1 online, M2 - M3 prezenčně

Harmonogram kurzu:

- M1 - březen
- M2 - duben
- M3 - duben/květen

Rok konání	Počet absolventů
2021	14
2022	25
2023	17

Organizace kurzu: timing

Modul 1: online

délka videí 5h a 12 min

164 stran čtení

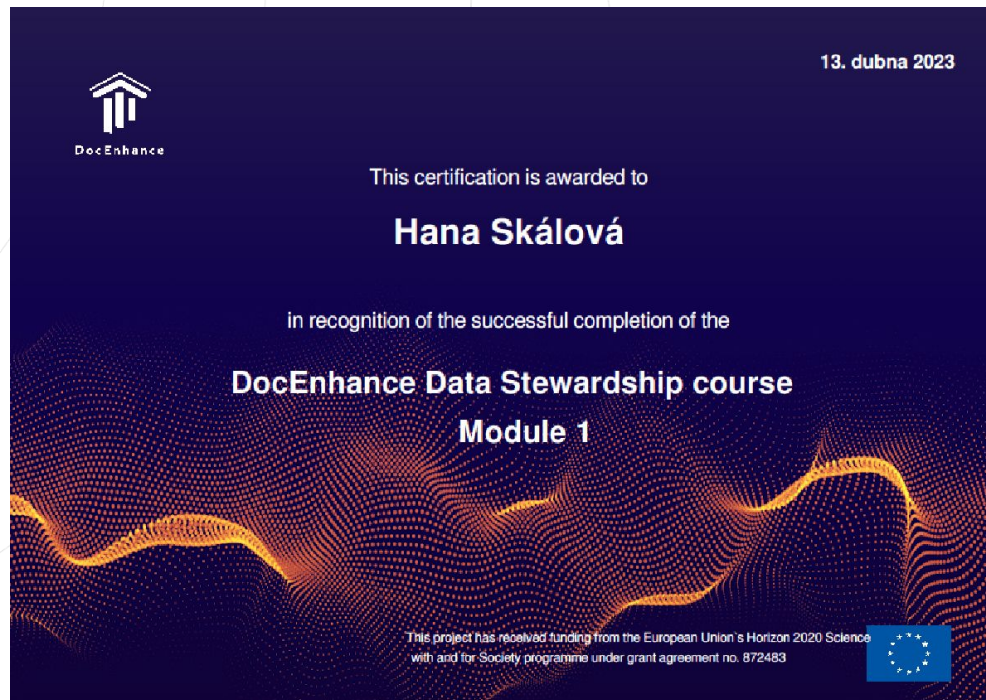
6 kvízů multiple choice (8 otázek) + závěrečný kvíz (30 otázek)

Modul 2: 18h výuky a práce ve skupině

Modul 3: 4h výuky prezenčně + 1-2 výjezdy do firem (1 + ½ dne)

Kurz v Moodle

- Vložení celého kurzu do Moodlu (po modulech)
- Modul 1 viditelný pro všechny
- Nastavení certifikátu dle šablony
- Generování certifikátu
- Modul 2 a 3 zejména pro lektory (event. lze nastavit část pro účastníky)



Organizace Modulu 1 v moodlu

Moodle

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](#)

≡ E-learning VŠCHT

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 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?

Organizace Modulu 2

- Domluva lektorů
- Nutné, aby lektori prošli modulem 1 + načtli materiály k modulu 2
- Každá hodina má specifickou strukturu a zadání
- Příprava úkolů & pre-assignments

Session 1, teacher guide: The importance of RDM and open archiving

Objective of this session

Think of this session as a useful and thought-provoking warm-up, where PhD candidates meet, get acquainted, and together brush up their knowledge on open science and transparency in research. They also get to map theory from Module 1 to their own research and discuss collaborative work across disciplines.

Learning outcomes of this session

- Understand the background and rationale of good research data management (RDM), and its connection to transparency of science.
- Understand why open archiving is put forth as the preferred option, and why in some cases it is not possible.
- Understand potential advantages of and barriers to data sharing.

Required previous knowledge (Module 1)

- Section 1.2 – Transparency and openness: Requirements and expectations

Pre-session checklist

- Get familiar with the structure of the session (content, timeline, technical requirements)
- Inform the participants about required previous knowledge for this session
- Distribute the pre-assignment texts to the participants, well in advance

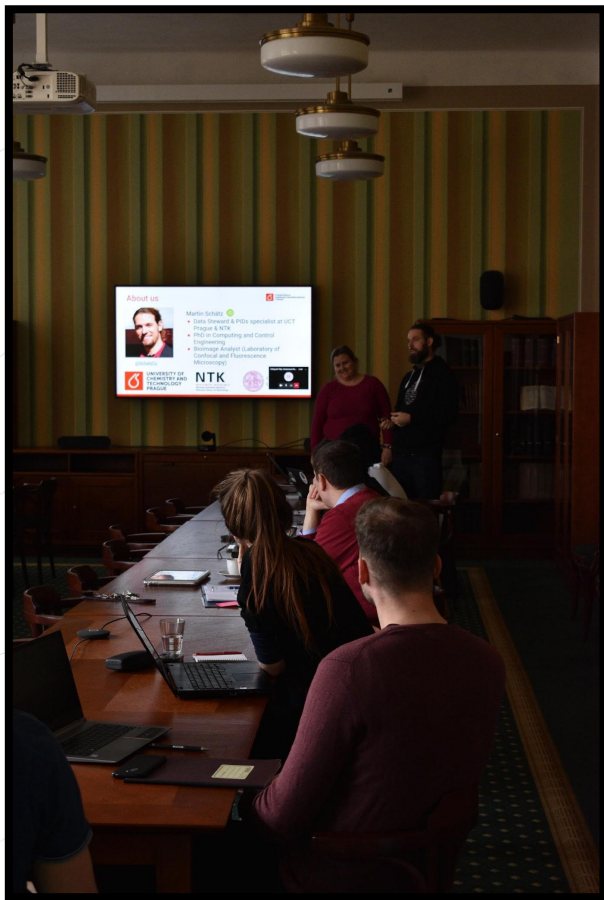
Organizace Modulu 2 v roce 2022

Datum	5.4.2022	6.4.2022	8.4.2022
	Zasedačka rektora VŠCHT	Zasedačka rektora VŠCHT	Zasedačka rektora VŠCHT
	10:00-12:00	10:00-12:00	10:00-12:00
Téma	Session 1: "The importance of RDM and open archiving" Session 3: "Structuring, documenting, and metadata"		Session 5: "Sharing and licensing research data"
<i>Hlavní lektor</i>	<i>Eva Hnatkova (NTK/VŠCHT)</i>	<i>Ivan Čmelo (VŠCHT)</i>	<i>Jiří Jiráť (VŠCHT/NTK)</i>
<i>Lektor 2</i>	<i>Martin Schätz (NTK/VŠCHT)</i>	<i>Jan Vrba (VŠCHT)</i>	
	14:30-16:30	14:00-16:00	13:00-14:55
Téma	Session 2: "The data management plan"	Session 4: "Data storage, file formats, and the FAIR principles" Session 6: "Visualising, reusing/searching, and citing research data"	
<i>Hlavní lektor</i>	<i>Eva Hnatkova (NTK/VŠCHT)</i>	<i>Marcela Dendisová (VŠCHT)</i>	<i>Radek Cibulka (VŠCHT)</i>
<i>Lektor 2</i>	<i>Robert Pergl</i>	<i>Eva Hnatkova (NTK/VŠCHT)</i>	<i>Marcela Dendisová (VŠCHT)</i>
<i>Lektor 3</i>	<i>Martin Schätz (NTK/VŠCHT)</i>		<i>Eva Hnatková (NTK/VŠCHT)</i>
			15:05 - 15:35
			Module 2 Wrap Up
			<i>Pavel Matějka (VŠCHT)</i>

Organizace Modulu 2 po finalizaci kurzu v roce 2023

Datum	14.4.2023	17.4.2023	21.4.2023
	9:00-12:00	9:00-12:00	13:00-16:00
Téma	Session 1: "The importance of RDM and open archiving"		Session 3: Structuring and documenting research data"
Hlavní lektor	Eva Hnatkova (NTK/VŠCHT)		Jan Vališ (VŠCHT)
Lektor 2	Martin Schätz (NTK/ VŠCHT)		Adam Sklenař (VŠCHT)
	13:00-16:00	21.4. 2023 9:00-12:00	24.4.2023 9:00-12:00
Téma	Session 2: "The data management plan"		Session 4: "Data storage, file formats, and the FAIR principles"
Hlavní lektor	Martin Schätz (NTK/VŠCHT)		Kristýna Sekyrová (NTK)
Lektor 2	Eva Hnatková (NTK/ VŠCHT)		Jan Vališ (VŠCHT)
Lektor 3			

Organizace Modulu 2



M3: Spolupráce s privátním sektorem

- Firmy blízké VŠCHT, kde naši doktorandi mohou najít uplatnění
- Firmy, kde působí školitelé specialisti
- Firmy pracující ve větší míře s daty

ORLEN Petrol

Zentiva

Organizace Modulu 3

Modul 3 program	
European and national view	2x 2h prezenčně na VŠCHT
Exkurze Zentiva	Půl dne návštěva firmy
Exkurze Orlen Petrol	Celodenní výlet

Spolupráce s privátním sektorem

ZENTIVA

- Exkurze přímo v sídle Zentivy v Praze
- Hodinová prezentace
- Následná ukázka laboratoří
- Prostor pro diskuzi s účastníky kurzu

Spolupráce s privátním sektorem

ORLEN PETROL

- Exkurze přímo v sídle Orlen Petrol v Litvínově
- Celodenní program
- Rozdělení účastníků do 3 skupin
- Účastníci si mohli vyzkoušet práci s daty přímo na místě
- Návštěva sídla petrochemie (velín)
- Prohlídky zařízení
- Prostor pro diskuzi s účastníky kurzu

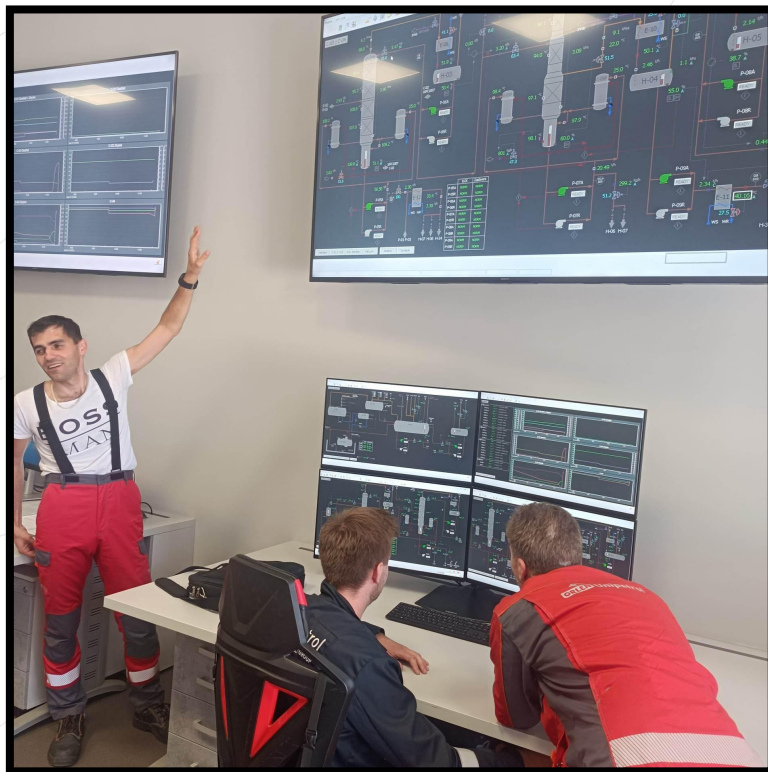
Spolupráce s privátním sektorem

ORLEN PETROL



Spolupráce s privátním sektorem

ORLEN PETROL



Výhody pro pilotáž

- Transkripce všech videí v Modulu 1
- Pokyny pro lektory k Modulu 2
- Předpřipravené úkoly po každé hodině
- Návod („cookbook“) jak na Modul 3 a privátní sektor

Kurz cílí na doktorandy ale mohli se ho účastnit i postdoci či studenti magisterského stud. programů.

Komunikace kurzu

- E-mailová rozesílka
- Spolupráce s proděkany pro VaV
- Obrazovky VŠCHT
- Webová stránka + kalendář VŠCHT

Data Stewardship course

News - DocEnhance *Data Stewardship* course in 2023

- The course starts with Module 1 (video lectures from March 1 to March 31, 2023) accessible for the enrolled participants through UCT Moodle.
- Modules 2 & 3 will take place in person at UCT Prague. See the description of Module 2 and Module 3 at the end of this page.
- The whole course will be in English.

Application

- **Application deadline: February 28, 2023**
- Enrolment confirmation will be sent out by the end of February, 2023
- **Online video lectures of Module 1 for enrolled participants will be available from March 3 to March 31, 2023.**

Course offer at UCT

The next round of **Data Stewardship Course** will start from March to May, 2023.

Apply here

Follow & Find more

Komunikace kurzu

- Komunikace probíhala v angličtině
- Cílení na doktorandy mimo jiné přes jejich školitele
- Snaha vysvětlit důležitost tohoto kurzu (např. DMP)
- Komunikace o exkurzích a firemních partnerech
- První dva běhy (2021 & 2022) byly proplacené i účastníkům v rámci projektu (5000,-Kč/ za účast a následný feedback)

Komunikace kurzu

Dear doctoral candidates,

We would like to inform you about the possibility of enrolling in a **Data Stewardship course**.
The whole course will be in English.

As research data becomes more openly available, data stewardship is a skill in high demand, both within and beyond academia. Starting March 1st the *Data Stewardship* course will examine and unfold the study of data management with a focus on how to put theoretical and practical skills into practice.

This course is divided into three parts.

Module 1 – online lectures (**whole March**)

Module 2 – workshops and assignments in local working teams at UCT, Prague (**April 14-21**)

Module 3 - workshops and regional assignments with employers (**April 28 + 2 excursions TBA**)

The course starts with Module 1 (video lectures 1, 3. – 31. 3. 2023) accessible for the enrolled participants in the UCT Moodle platform

See [more information about the course](#).

The deadline for application is on **March 13** - [Apply for the participation in the course here](#).

Feedback účastníků

- „Musím říct, že jsem se v kurzu naučil nejen nová témata, ale také jsem si ho užil a během skupinových setkání jsem si našel nové přátele (pravděpodobně na celý život). **Byla to pro mě životní zkušenost.**„
- „Páteční sekce v Unipetrolu byla za mě velmi povedená - tím, že se účastníci rozdělili na relativně malé skupiny, resp. podskupiny, **bylo poměrně dost prostoru si různé věci „ošahat“ a případně se i zeptat.**“

Feedback účastníků k Modulu 3

- „Den strávený v Unipetrolu byla asi nejzajímavější částí kurzu, z části protože je to vlastně součást mého oboru, ale **hlavně jsme se dostali do reálného provozu s okamžitou správou dat**, krásně hovořili o tom, jak se v dané oblasti vyvíjeli a tak.“
- „Prezentace pana Herinka byla vynikající, **představila nám procesy v rafinérii, vysvětlila nám role operátorů a zdůraznila roli dat pro současné řízení rafinérie a každodenní provoz**. Bylo také zajímavé se dozvědět o historickém vývoji sběru dat v rafinérii.“
- "Poslední exkurze do Unipetrolu byla mou oblíbenou ze všech modulů. **Byla to má první zkušenost v průmyslu**. Takže jsem si každý okamžik hodně užil.“

Feedback lektorů

- *„Myslím si, že je skvělé, že vzniká takovýto program, který může studentům usnadnit práci s výzkumnými daty. Jsem ráda, že jsem se toho mohla zúčastnit. Pevně věřím, že zájem o něj bude dlouhodobý.“* Petra Černošávková, NTK
- *„Osobně jsem opravdu rád, že se tento cíl podařilo naplnit. Z pozice školitele jsem velmi ocenil zápal doktorandů do tématu a množství dotazů během konání kurzu.“* Michal Zbuzek, Orlen Petrol

Naše doporučení

- Kurz by měl být začleněn přímo na začátku doktorského studia.
- Kurz by měl probíhat po celý semestr.
- Modul 2 a 3: Účastníci by měli mít dostatek času na práci s daty.
- Hodiny by měly probíhat v počítačové učebně nebo doporučujeme, aby si účastníci přinesli své notebooky.
- Modul 3: Ideálně by měl být pořádán mimo univerzitu, aby doktorandi zažili jiné prostředí např. privátní společnosti.

Tipy na pilotáž

- Modul 1 může být otevřen celý rok
- Modul 2 a 3 stačí pilotovat 1x ročně
- Komunikace s účastníky v průběhu celého kurzu
- Vytvoření skupiny v MST, kde bude kanál s přístupem pro účastníky i lektory (dostanou tam sdílené prezentace, tabulky, odkazy atd)
- Příprava lektorů (nutnost načíst materiály, projít Modul 1)
- Vysvětlit využití kurzu do budoucna ve vědecké kariéře ale i např. v práci mimo akademii
- Utužení vazeb s privátním sektorem

Co dělalo problém

- časování kurzu - lepší M2 workshopy rozdělit do více týdnů
- dodržet časový rozsah jednotlivých sessions
- myslet na reminder rozesílky ohledně přihlášení se do kurzu
- lektori v Modulu 2 nejprve učili svou látku, nenačítali materiály
- chyběli specialisté na určité sekce a data stewardi
- udržení účastníků v kurzu (z 56 → finálních 17 v roce 2023)
- UK měla problém najít firmy do M3

Administrativní role PhD Supportu

- **komunikační a organizační role**
- uploadování Modulu 1-3 do univerzitního moodlu (dohlédnutí na hladký průběh generování certifikátu) ve spolupráci s IT
- zajištění lektorů pro Modul 2
- propagace po univerzitě (příprava textů)
- registrace účastníků (moodle) + kontrola, že prošli M1
- zajištění místností/techniky na univerzitě + fixní termíny
- informování účastníků před startem kurzu, kde a kdy se hodiny konají a kdy jsou exkurze

Administrativní role PhD Supportu

- komunikace s účastníky, co je třeba mít s sebou na praktické workshopy, (zasílání pre-assignments a následných assignments po hodině v rámci Modulu 2)
- prezenční listiny
- domluva s privátním sektorem a zajištění exkurzí pro Modul 3
- zajištění autobusové dopravy do Orlen Petrol
- následný sběr feedbacků
- DPP pro lektory

Závěrečné video s účastníky



NTK

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

Děkuji za pozornost

Otevřená diskuse

Mili Losmanová

mili.losmanova@techlib.cz



Návod pro přenos kurzu do moodlu

- Co se týká podpory s naší strany a jednotlivých modulů k nahrání, kontaktovat Akademické služby NTK reference@techlib.cz
- Každý modul obsahuje již návody a texty pro lektory, úkoly, kvízy...
- Certifikát není automaticky nahraný, je potřeba ho nechat vytvořit na základě podkladů z VŠCHT
- Certifikát musí obsahovat loga EU a DocEnhance
- Nutnost vše před spuštěním nechat otestovat

Co plánujeme dál?

- Na Moodlu NTK bude stránka, kde budou umístěny pro stažení materiály (všechny 3 Moduly) + další DocEnhance kurzy
- Soubory MBZ (formát Moodle) je jednoduché nahrát do standardního rozhraní Moodle.
- Měl by to zvládnout každý administrator Moodle.
- Je potřeba dodatečně připojit [plugin "mod_customcert"](#), který pomáhá přizpůsobit certifikát.
- NTK dodalo svoje logo a certifikát se vygeneruje automaticky každému uživateli po absolvování závěrečného testu v Modulu 1.
- U Modulu 2 a 3 by mohlo jít také nastavit generování Certifikátu po splnění některých povinností.

Implementace kurzu

Jako součást kurikula doktorských programů

- Každá instituce/VŠ rozhoduje o tom, zda doktorandi získají za absolvování kurzu body/kredit.
- Všechny tři moduly jsou spojeny s úkoly, které mohou dohromady fungovat jako zkouška.

NTK

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

Děkujeme za účast

