3. FDM-Werkstatt | 31.03.-02.04.2025 | Universität Münster

The 3. FDM-Werkstatt was organized by fdm.nrw in cooperation with the University of Münster, the University and State Library of Münster and NFDI4BIOIMAGE.

Program

Monday, 31 March 2025, Multiscale Imaging Centre (MIC), University of Münster

- 3:00 PM: Arrival & Registration
- 3:45 PM: Official Welcome
- 4:00 PM: Tour of the Multiscale Imaging Centre (MIC)
- 5:30 PM: Manakish & Merriment

Tuesday, 01 April 2025, Torhaus, University of Münster

- 9 AM: Arrival & Registration
- 10 AM – 1 PM:
    - Session 1: Data Driven Visualization (with LaTeX)
    - Session 2: Getting up and running with InvenioRDM
    - Session 3: Metadata Magic: A Deep Dive into eLabFTW Customization
- 1 PM – 2 PM: Lunch
- 2 PM – 5 PM:
    - Session 1: OpenRefine for Beginners: Making Your Data Better
    - Session 2: Using the ELN Format
    - Session 3: Open up your research with the OSF and GitHub
- 7:00 PM: Self-Pay Dinner at Mocca d'or

Wednesday, 02 April 2025, Torhaus, University of Münster

- 9 AM: Arrival & Registration
- 10 AM – 1 PM:
    - Session 1: ARCify your research project
    - Session 2: Embedding of OMERO.iviewer in eLabFTW
    - Session 3: From Data to Visuals: A Hands-On Workshop with Wikidata, Gephi, and Inkscape
- 1 PM – 2 PM: Lunch
- 2 PM – 4 PM:
    - Discussion: What makes RDM service infrastructures "good" as a service?

Abstracts

Data Driven Visualization (with LaTeX)

Effective data visualization transforms complex information into clear, engaging graphics. Our workshop, "Data-Driven Visualization with LaTeX," is designed for students, researchers, and enthusiasts eager to create dynamic visualizations using LaTeX. Participants will learn how to automate the creation of visual elements by leveraging data from CSV files, ensuring that updates and modifications are seamlessly integrated through data changes.

The workshop will guide attendees through the process of importing and managing data from external sources, emphasizing best practices for maintaining data integrity and facilitating easy updates. We will explore how to tweak and toggle certain modification possibilities of a provided script adjusted the outcome to your desires and needs.

By the end of the workshop, participants will have developed their own set of dynamic, data-driven visualizations, fully integrated with their datasets. They will leave equipped with the knowledge and skills to utilize LaTeX as a powerful tool for producing high-quality, automated graphics that enhance their research and presentations.

Target Group

- Participants should have basic knowledge about LaTeX
- If you work through https://learnlatex.org you are well prepared

Requirements

- Properly working LaTeX distribution on laptop or access to Overleaf / ShareLaTeX

Person/s in charge

- Lukas C. Bossert (RWTH Aachen University/DKZ.2R)

Metadata Magic: A Deep Dive into eLabFTW Customization

In this session, participants will be introduced to advanced features of the eLabFTW electronic lab notebook. We will focus on customizing experiments and database elements using metadata. Participants will be able to work on their own templates during the course of the workshop.

We will cover the design and implementation of custom templates. Metadata in the form of so-called extra fields can be used to create purpose-specific templates with a well-structured layout that can be filled in intuitively, thus improving standardization across projects.

Building on this, we will explore ways to manually construct metadata in JSON format that is compatible with eLabFTW and can be attached to experiments as well as database elements.

Experienced participants will have the opportunity to extract and manipulate metadata via the eLabFTW API, allowing for automated metadata enrichment.

Target Group

- All workshop contents do not require any coding or programming knowledge
- A basic understanding of structured data formats (e.g. JSON) is helpful, but not required
- Knowledge of basic usage of eLabFTW (creating and editing experiments and database items) is required

Requirements

- Text editor capable of JSON syntax highlighting (Notepad++, Visual Studio Code or similar)

Person/s in charge

- Dr. Alexander Minges (SFB 1430/Zentrum f. Medizinische Biotechnologie, Universität Duisburg-Essen)

Getting up and running with InvenioRDM

In this hands-on workshop we will guide participant through the installation of the latest InvenioRDM version on their computers and start with some simple customization tasks (changing the Logo, colors or modifying the UI in general), and answer questions that come up.

Target Group

- Participants should have some experience in coding or programming (JavaScript/React and/or Python)

Requirements

- Linux/MacOS Computer with admin rights, git and docker installed

Person/s in charge

- Sarah Wiechers (University Münster)
- Werner Gresshoff (University Münster)
- Markus Klöpper (University Münster)
- Karl Krägelin (University Münster)

OpenRefine for Beginners: Making Your Data Better

In this practical 2-hour workshop we will introduce you to OpenRefine, a powerful yet easy-to-use tool that helps you make sense of messy data.

We'll start the workshop with an overview of what OpenRefine is and the different scenarios where it can be useful, from preparing research data to improving the quality of spreadsheets. You'll learn how to easily install OpenRefine on your computer and navigate the user interface, including the project panel and data table.

As we delve into the basics, you'll discover essential data cleansing techniques. We'll cover how to sort and filter data, and how to import datasets from CSV files and Google Sheets. You'll get hands-on experience with faceting, which allows you to effectively analyse and visualise your data. We'll also explore clustering methods to identify duplicates and standardise records, making it easier to ensure the integrity of your dataset.

In our data transformation session, you'll learn how to link your data to external databases and perform basic transformations using built-in functions - no programming skills required!

The workshop includes a guided hands-on exercise where you can practice your new skills on a sample dataset. We'll conclude with a Q&A session where you can ask questions and troubleshoot any challenges you may face.

Target Group

- Participants may have little or some experience in coding or programming

Requirements

- None

Person/s in charge

- Dennis Voltz (Service Center Digital Humanities, University of Münster)
- Maike Sommer (University and State Library of Münster, University of Münster)

Using the ELN Format

The ELN Consortium (https://github.com/TheELNConsortium) aims to increase interoperability of electronic lab notebooks (ELNs) by establishing common specifications that support the exchange of (meta)data between different ELN software.

In this session we will take a closer look at the ELN file format. What information is contained, how is it structured, and how can we use this in our workflows, for example, to exchange data between ELNs, transfer datasets to Coscine, or publish work in repositories? Depending on the skills of participants, we may also look into current issues in the style of a hackathon, solve them or suggest how they may be solved or discuss what workarounds there are.

Target Group

- Participants are encouraged to bring their own ideas and challenges as a basis for discussion and programming work
- (Python) programming skills are required to fully participate and code, while anyone is welcome to bring their ideas

Requirements

- Python programming and a Jupyter Instance or similar environment (We will try to set up a central Jupyter instance for all to use)

Person/s in charge

- Nicole A. Parks (RWTH Aachen University)
- Note: This is a general idea and hosts and contributors are welcome.

Open up your research with the OSF and GitHub

Researchers have been recognizing their responsibility in increasing the openness of research materials, data, and articles. At the same time, funders and institutions have been raising their requirements. In this session, we will present the following tools that enable researchers to open up and enrich their research without placing burdens:

- Open Science Framework (Lukas Röseler): Researchers can use this open-source tool to document materials and data from their entire research process, to structure complex projects with hierarchical structures and links between projects, make contributions transparent, (pre)register hypotheses, and publish and review research via preprint servers.
- GitHub (Ole Hätscher): A considerable portion of research involves programming and other types of software development. GitHub provides an infrastructure for teams of researchers to work on software together.

Target Group

- Participants may have little or some experience in coding or programming:
    - A minimal understanding of code is recommended for the Github part.
    - No understanding of code is necessary for the OSF part.

Requirements

- OSF Account
- Github Account

Person/s in charge

- Lukas Röseler (Münster Center for Open Science)
- Ole Hätscher (University of Münster)

ARCify your research project

This hands-on workshop introduces the concept of the Annotated Research Context (ARC) – DataPLANT's implementation of a Fair Digital Object (FDO).

Learn how to use the tool ARCitect to create ARCs (https://arc-rdm.org) for your datasets, annotate them with metadata and share them via the DataHUB (https://git.nfdi4plants.org).

Target Group

- Participants may have little to no experience in coding or programming
- While the workshop is designed around demo data from plant sciences, the tools are likewise adoptable in other (natural) sciences

Requirements

- Requires a laptop with ARCitect installed
- Please sign up for a DataPLANT user account in advance of the workshop
- An online tech-check meeting is offered on March 25th, 14:00 - 14:30

Person/s in charge

- Dominik Brilhaus (Heinrich-Heine-University Düsseldorf)

Embedding of OMERO.iviewer in eLabFTW

To improve the user exprience for the bioimaging community it makes sense to implement a link to the widely used Image Data Management Platform OMERO. As the user details their experiments in eLabFTW an intuitive addition would be an embedded view of the microscope image in question. For other applications (Project cLovid from Münster) the OMERO.iviewer has been adapted to be embedded into e-Learning platforms. In a similar fashion we will try to implement the simple embedding of an OMERO image into an experiment in eLabFTW. This is supposed to lay the groundwork for an official implementation of this feature by the eLabFTW core team, as external/custom plugins are not supported.

Target Group

- Participants should be fluent in programming

Requirements

- DIE of your choice ready to go

Person/s in charge

- Jens Wendt (University Münster, Imaging Network, NFDI4Bioimage)

From Data to Visuals: A Hands-On Workshop with Wikidata, Gephi, and Inkscape

Join our workshop, "From Data to Visuals: A Hands-On Workshop with Wikidata, Gephi, and Inkscape" where you'll learn to transform data into compelling visualizations using powerful tools. We'll start by querying Wikidata to gather geo-coordinates from NFDI members and exporting this information as a .csv-file. Next, you'll import the CSV into Gephi, a network visualization software, to create and analyze dynamic network graphs. Discover how to uncover patterns and relationships within the data through hands-on exercises.

After analyzing your networks in Gephi, we'll move to Inkscape, a versatile graphic editor, to polish your visualizations. Here, you'll add final touches, customize designs, and create professional-quality graphics that effectively communicate your findings. This workshop provides a complete workflow from data extraction to stunning visual presentations, perfect for students, researchers, and data enthusiasts looking to enhance their data visualization skills.

By the end of the session, you'll have the knowledge and experience to create insightful and attractive visual stories from complex data sets.

Target Group

- Participants may have little to no experience in coding or programming

Requirements

- Please have installed Gephi (https://gephi.org/) and inkscape (https://inkscape.org/)

Person/s in charge

- Jonathan Hartman (RWTH Aachen University/DKZ.2R)
- Lukas C. Bossert (RWTH Aachen University/DKZ.2R)

What makes RDM service infrastructures “good” as a service?

Offering research data management (RDM) services at an institution for higher education comes with the need to legitimize the portfolio of provided services. Thus, supporting stakeholders may want proof that the resources invested in RDM infrastructures actually support the researchers of their institution. The attempt to measure this inevitably leads to the "McNamara fallacy": only those observations that can be measured are being considered and all others ignored to determine the value of a service. Yet, to these stakeholders, a lack of numbers is reason enough to jeopardize the continuation of services.

The question is: How can we prove services to be important without providing quantitative observations to illustrate their use. What do we do if the significance of a service cannot be illustrated in numbers? This service may be a key to a hidden workflow and its discontinuation would cause delays, errors, setbacks for researchers, labs, projects and entire institutions. Bryn Nelson (Nelson, 2009) made the point in research data management that setting up a repository might not suffice for it to become important for researchers.

This session will provide the opportunity to discuss and define key measures enabling services to be perceived as relevant as they are. Building on the experience of setting up NFDI4Plant infrastructure within our organization, we will combine all kinds of expertise to communicate the relevance of RDM infrastructures. Is the availability of a service already enough? Which additional factors, like curation, helpdesks, user SOPs or additional high-level services, can make a difference? Is it possible to make

the RDM services appear as essential so that quantitative proof of their relevance becomes inconsequential?

With perspectives from different organizations on this question, we would like to write a white paper on: “What makes a good RDM service infrastructures “good” as a whole service?”

Target Group

- Little to no experience in coding or programming
- Some experience: Programming is not at the hard of this workshop, but the knowledge of how software and useer support and data curration work together. It can also be from interest to the result of this workshop if participants have experienced data submissions or data extraction at data repositories at first hand

Requirements

- None

Person/s in charge

- Dirk Fleischer (Heinrich-Heine University Düsseldorf)
- Dominik Brilhaus (Heinrich-Heine University Düsseldorf, CEPLAS)