

Developing metadata management training: Experiences and future plans

Becky Oldroyd, Hayley Mills, & Jon Johnson

Outline

- What is CLOSER?
- Existing metadata management training
- Current CLOSER training
- Previous CLOSER training
- A new training offer
- Possible opportunities
- Future plans



What is CLOSER?

- The UK partnership of leading social and biomedical longitudinal population studies (LPS), the UK Data Service (UKDS) and The British Library
- Mission: Increase the visibility, use, and impact of longitudinal population studies, data, and research
- Areas of work:
 - Data discoverability
 - Policy and dialogue
 - Training and capacity building
 - Data innovations



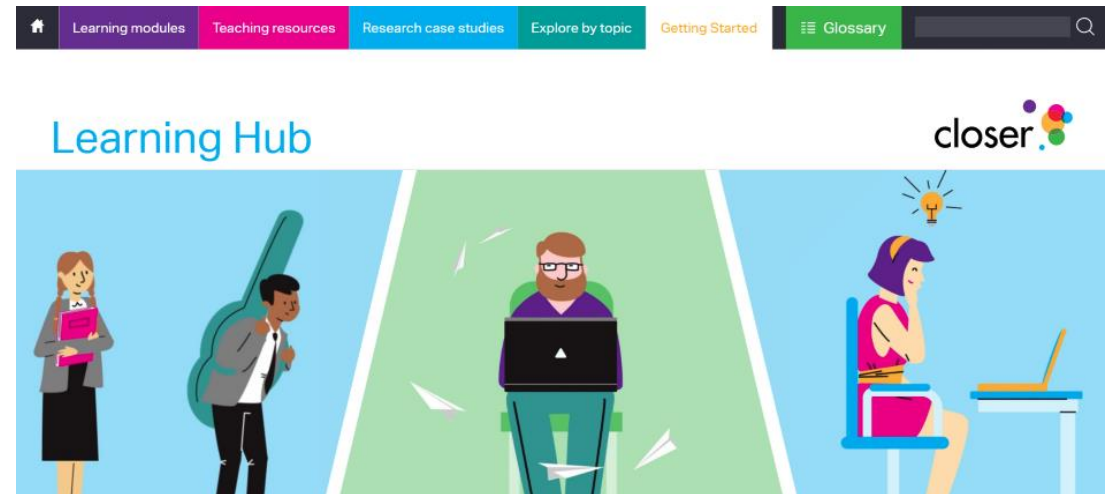
Existing metadata management training

- Limited existing training which focuses on the importance and application of metadata management for sharing purposes
- Mostly ad-hoc, focusing on a specific topic or standard, with the majority done through on-the-job training
- Existing training available for:
 - [Research Data Management \(RDM\)](#)
 - [FAIR Principles](#)
 - [Data Documentation Initiative \(DDI\)](#)



Current training - Learning Hub

- Information and resources to help explore LPS and get started using the data
- Aimed at those who are new to LPS
- Modules:
 - Introduction to longitudinal studies
 - Study design
 - Data harmonisation
 - Understanding metadata
 - Analysing longitudinal data
 - Research communication



<https://learning.closer.ac.uk/>



Learning Hub - Understanding metadata

- Advanced topics but at a high level
- Helps undergraduates understand research tools (e.g., CLOSER Discovery and the UKDS data catalogue)
- Content:
 - What is metadata and why is it important?
 - Structured metadata
 - Documentation
 - Using metadata to discover relevant variables
 - Test your knowledge

<https://learning.closer.ac.uk/learning-modules/understanding-metadata/>



Previous training – London (2019)

- **Understanding metadata management in LPS**
 - Two-day workshop
 - 28 attendees across the two days
 - Data managers/people involved in LPS
 - Overview of structured metadata, use cases, and data managers sharing experiences of working with metadata
 - Feedback:
 - Networking and group discussions
 - More examples/case studies
 - More practical activities

Previous training – Malawi (2021)

- **Health data research and metadata management**
 - Two-day workshop
 - 19 attendees
 - Primarily data managers
 - Knowledge transfer and capacity building
 - Feedback:
 - Enjoyed demonstrations of tools/talks on practical usage
 - Appetite for a longer course
 - More practical elements

A new training offer

- Identify and fill the gaps in metadata training and capacity
- Provide professionals and researchers with the foundational knowledge and understanding to:
 - *effectively manage quantitative data*
 - *in the social, humanities, and population health sciences*
 - *for data sharing and discovery.*

How do we do this?

- CLOSER's experience
- Training reviews:
 - Assessing the impact of NCRM's Training and Capacity Building Activities 2014 - 2016 (2018)
 - Careers Research & Advisory Centre (CRAC) report National Data Skills learning and pilots for teaching data skills to noncognate undergraduates (2022)
 - A review of quantitative analytical training needs for users of longitudinal studies (Maddock, O'Neill, Rainsberry, & Hardy, 2019)

Possible opportunities

- UCL undergraduate training:
 - Data management dissertation sessions
 - Stand-alone extra-curricular sessions
- Degree-level module
- DDI training for archives and others setting up repositories
- New online courses
- Training Hub
- Collaborations using existing training (e.g., UKDS RDM training)



Future plans

Basic (face-to-face, later online)

Foundational (face-to-face, later online)

Advanced (face-to-face)



Overview of new training offer

Basic (face-to-face, later online)

- One day
- Postgraduates, new data managers, programmers
- Learning objectives:
 - Define structured metadata/other key terms
 - Explain why (meta)data management is important
 - Illustrate why using standards is important/describe how you would decide which standard to use
 - Apply the general principles of metadata to a dataset

Overview of new training offer

Foundational (face-to-face, later online)

- One day
- Data managers (small projects and studies)
- Overall learning objective:
 - Implement structured metadata using DDI codebook to enhance data sharing

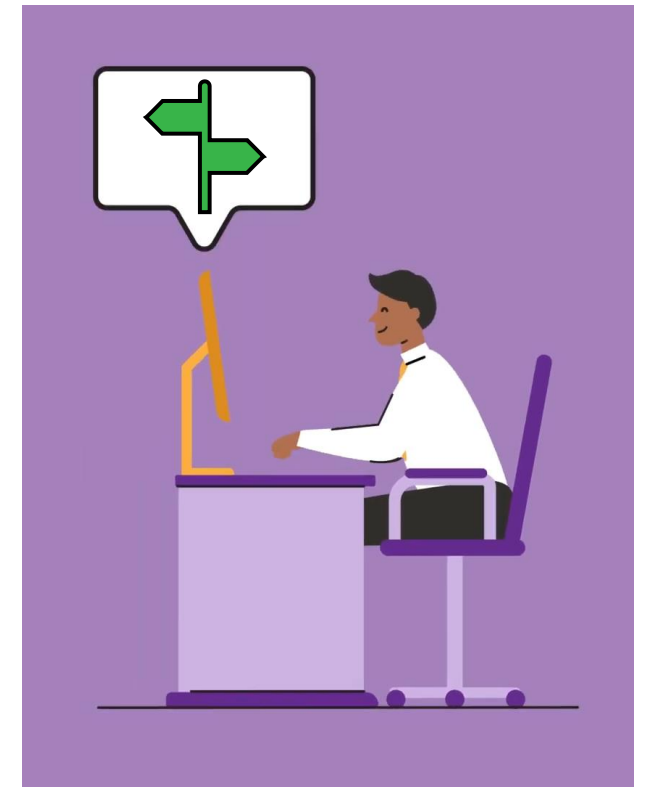
Overview of new training offer

Advanced (face-to-face)

- Two days
- Data managers (longitudinal, large projects/studies)
- Overall learning objective:
 - Implement structured metadata in the data lifecycle to enhance data sharing

Training Hub

- Overview of best practices and signposting to resources
- Aimed at those who are familiar with LPS data (e.g., postgraduates, early career researchers, and policy makers)
- Data management module
 - Overview of research data management, with focus on secondary & longitudinal data
 - Collaborating with the UKDS



Collaborations

- Current collaboration with the UKDS
- Annual virtual training course
 - *Accessing and using 'real-world' study data – a Stata-based introduction for newcomers to longitudinal research (2023)*
- Future collaboration with the UKDS
 - Ensuring no overlaps
 - Linking to each others' resources
- Ongoing discussions about working together in the future
- DDI Alliance Training Group

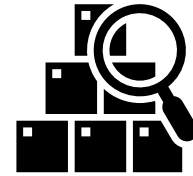
Summary

- There are gaps in metadata management training at all levels
- We have had success in running in-person metadata management training
- More training is needed
- We are using feedback from training events
- Future training will:
 - Have a clear offer/path
 - Be for a specific audience and objective
 - Be re-useable (train the trainers)
- Next step: finalise the basic online course

Get in touch!



closer@ucl.ac.uk



discovery.closer.ac.uk



@CLOSER_UK



Sign up to our
newsletters to get the
latest from CLOSER

Discussion points

- How to understand the attendees' baseline knowledge and cater to that?
- When to use neutral vs. contextualised examples?
- What are we missing?