



RDA basierte Module für eine Winter School zur Ausbildung von Daten-Managern und -Wissenschaftlern

Peter Wittenburg, Ari Asmi, Jacco Konijn
MPCDF, U Helsinki, UvA

research data sharing without barriers
rd-alliance.org

- Berufsbilder Definitionen vom RfII und anderen
- Berufsbilder noch nicht etabliert
- vor allem in der Datenpraxis verschiedene Individuen mit Wissen entlang eines Kontinuums

Data Manager

Data Librarian

Data Scientist

**Data Org
Expert**

**Analytics
Expert**

Lernziele

Die Teilnehmer sollen lernen

- wie man Daten FAIR-gerecht organisiert (Findable, Accessible, Re-Usable)
- wie man Standards einsetzt, um die Interoperabilität zu erhöhen
- wie man ein Repository sauber aufsetzt und wie seine Qualität geprüft werden kann
- wie moderne Konzepte wie Data Type Registry und Jupyter Notebooks eingesetzt werden können

Setup

	Module Name
8.30	Theoretical Session
11.00	Practical Session
12.30	Lunch break
13.30	Evaluation & explanations
14.30	continuation of Practical Session
18.00	end

- DFT: basic data model and terminology, DFT/FAIR compliant implementation
- PIT: concept of PIT API (profiles) and information types
- DTR: data typing and use of type registry
- DDC: recommendations about proper data citation
- MDC: use of metadata catalogue

Week Setup with Modules

Day	Module	Theoretical	Practical
1	Proper Repository	Repositories principles & functions, proper Data Organisation, Fedora setup, Certification Methods, FAIR	building a repository, create MD, store data, request PIDs, add PID to MD, make data accessible
2	Registering ENVRI Data	ENVRI metadata schema, PID types, data types, DTR intro, DCAT	get exp data, get checksum, create MD, DCAT, register PIDs, enter PID types in DTR
3	Metadata & Collections	collection forming & registering, citing data, OAI PMH, ResourceSync, indexes, search portals	expose metadata to harvester, build collections and register collections
4	Data Typing	data typing, design scientific DTR entries for visualisation, conversion & interoperability	create scientific DTR entries, carry out visualisations, conversion etc. incl. registration of new data
5	Using Data	Jupyter notepad, introducing data analysis using data and metadata, scientific value of stored data and potential	learn Jupyter, select data, write Jupyter scripts to extract metadata and data, carry out analysis, create prov metadata, register new data,

- es werden 2 verschiedene Datensätze von ENVRI verwendet und auch 2 Hauptgruppen gebildet
- darunter werden jeweils 3/4er Gruppen gebildet, in denen mindestens eine(r) versteht, Scripts zu schreiben
- es wird ein Scripting Fachmann die ganze Zeit anwesend sein, den RDA EU bezahlt
- dazu werden jeweils Spezialisten anwesend sein
- alles muss vorher ausgetested werden



Alles klar?