#1 How can FAIR data management be integrated into the academic curriculum? Table host: Cord Wiljes, University of Bielefeld

In this group we collected existing formats for teaching FAIR principles to students and researchers. In addition we identified motivations for participating in the courses.

Course Formats

We identified a wide range of different course formats for teaching FAIR data management:

- Discipline-specific courses for working groups (2h length), also useful in preparation for individual support
- Workshop for PhD candidates in context of a graduate school (half-day/full day/two day length)
- Integration into existing full-semester lecture series in one session (1,5 h length)
- “Coffee Lecture”, no registration necessary (1 h length)
- Full-day seminar, registration necessary
- Topic-specific courses (e.g. “Metadata”) (length varying from 1,5 h to 2 days)
- Webinar, MOOC
- 4 days course at Fraunhofer (paid)
- Full semester course (in information science for Bachelor and Master)

Motivations for participants to learn FAIR data principles

For FAIR data management courses to be accepted, motivations of participants are identified:

- Credit points are a huge motivation
- Experience with FAIR data management and data data management technologies (e.g. electronic lab notebooks) are a plus if applying for a job in industry
- Interesting topic (intrinsic motivation)
- Increases (academic) status, e.g. “Open Science Ambassadors”)
- Requirements by funding agencies, especially DMP required for proposal
- Courses in the context of a working group are often mandated by the group head

Additional comments

- It is difficult to formally integrate FAIR data management into the medical curriculum, because the curriculum is already very tight. This holds true for many other disciplines, too.
#2 For what and how are you searching for FAIR training materials?  
Table host: Karsten Peters, German Climate Computing Center – DKRZ

Topic and Goals of Table #2

The general topic of Table #2 of the Knowledge Cafe Session at the GOTRAIN workshop was the current state of the FAIRness of research data management (RDM) training materials and how the current state can be improved towards improving FAIRness of such materials.

Several keynote talks of the morning workshop session mentioned the (potential) importance of collecting training materials in a centralized location and illustrated current developments in this direction. Development of ontologies and metadata schemas for training materials is needed such that published training materials become increasingly FAIR.

The goals of the discussions at Table #2 included a mapping of the current state of searching for and obtaining RDM training materials, identification of problems with the current situation and brainstorming on future ontologies and metadata entries for centrally collected and published RDM training materials.

A summary of the discussed topics is presented in the following – accompanied by photographs of the used flip chart sheets.

**Current state of RDM training material FAIRness (Fig. 1)**

Asked the question of

How and by which means do you search for and obtain RDM training materials for your purposes as RDM-Trainer?

the participants reported on a „mixed bag“ of current strategies:

- google the topic and browse through the results
- Personal network/word-of-mouth
- Browsing through already existing collections and link lists, e.g. on the German website forschungsdaten.org – a community based, wiki-like platform
- Discipline specific conferences (IT, RDM, library)
- Organisation of workshops and invitation of community experts
- Already existing eLearning platforms, e.g. LIBER, DataONE, MANTRA

It is fair to say that the result of this discussion topic exactly met the expectations in the sense that there currently exists a wide variety of possibilities to obtain RDM training materials and a centralised collection might indeed be useful.
Fig. 1: Participant input regarding the current state of searching for and obtaining RDM training materials

How to find RDM-Tri.

Input

- generic material: by chance while browsing for RDM-related
  - forschungsdaten.org
- decentralized communication
- community experts
- Google
- conferences
- personal network
- eLearning platforms
  "campunities"
Problems/shortcomings of the current state (Fig. 2)

As the question of:

Where do you currently see shortcomings or problems when searching for RDM training materials?

The discussion evolved around the following items:

- generic vs. discipline specific materials
- generic RDM training materials in one flavour or the other dominate the available resources
- discipline specific approaches and use cases are hard to find or not present – as every research discipline has different needs of RDM and features available methods/services facilitating efficient RDM, discipline specific knowledge is a staple of RDM training
- licences of published RDM training materials are often not clear OR it is not entirely clear under which kind of licence compiled training materials can be published for reuse. Institutional policies are oftentimes not clear in this regard.
- A large part of available materials is published on institutional webpages, thereby having no assigned PIDs. URL links often change and material becomes unfindable.
- Materials are often not up-to-date or updates are not communicated – keeping up with current developments in RDM and the increasing awareness of the FAIR data principles and their incorporation in funder, journal and institutional policies is essential.

Another point raised was the general lack of incentives for researchers to attend RDM trainings and to act as multiplicator (“data champion”) in their home institutions and/or community networks. Oftentimes, the same „usual suspects“ attend the trainings. Establishing reward systems could improve the situation.
Fig. 2: Participant input regarding problems/shortcomings of the current state of searching for and obtaining RDM training materials

Problems
- discipline specific search impossible
- generic material is omnipresent
- links disappear → PDA’s
- licences
- updates!
- motivation!
- no reward system
Increasing the FAIRness of training materials (Fig.3)

This brainstorming evolved around the thematic anchor. Given the existence of a centralized repository of FAIR RDM training materials, which keywords/search facets would you use to locate materials relevant for your purpose?

A comprehensive – and probably by no means exhaustive – list of possible keywords/metadata entries/search facets was compiled:

- **Scope, goals and documentation of the materials:**
  - What is to be achieved by the training?
  - Which capacities will the participants have attained?
  - How is the material to be applied?

- **Target Group**
  - Is the material and the scope presented suited for students, researchers, IT-service providers of librarians?
  - What are the prerequisites for attending, i.e. prior knowledge and experience

- **Scientific Discipline**
  - This comes back to the problem raised while discussing the current state (see above)
  - Generic RDM material of discipline specific?

- **Data type covered**
  - Here, data type refers to things like geospatial data, spectrometer data, simulation results, etc
  - This is important as the same type of data is used in different disciplines, thus possibly requiring similar approaches of RDM

- **Date published and information on updates**
  - See above for the discussion on keeping materials up-to-date

- **Ideal group size of participants**

- **Duration**
  - E.g. 15 Minutes, 90 Minute lecture, 2-day block workshop

- **Format**
  - E.g. workshop, lecture, hands-on session, eResource, etc.

- **Language of the material**

- **User/participant experience**
  - Interactive comments section on the web
  - „rating“
  - Publication of course surveys
Fig. 3: Participant input from the brainstorming of useful metadata entries/search facets of a centralized RDM training material repository


#3 Data Stewards - role or function? Table host: Jasmin Böhmer, UMC Utrecht

- with regards to Germany in comparison to the Netherlands
  - The strong requirements and support from the Dutch funders with regards to FAIR data, open data, DMP's and data stewardship does not find equal measures in Germany
  - the urgency of the reproducibly crisis exposed by nature in 2016 (?) perhaps has not yet arrived in the German collective mind

- role vs. function
  - depends on the discipline the data steward is applied
  - depends on the available funding (national, regional, or institutional)
  - nevertheless it seems to have data stewards on certain levels to function as beacons for this type of work and the importance of good RDM
  - depends on the line of work:
    - a. 0.5 working on policy development in addition to an already existing role makes sense
    - b. 1.0 FTE for someone actually managing research data from start to finish of a project is also reasonable

- data steward
  - title is not used that commonly in Germany; somewhat more popular in other European countries
  - overlapping function areas are of course existent with data curator, embedded librarians, subject librarians, data officers etc.

- data stewardship and quality control
  - with regard to the current application areas of data stewards and what data stewardship is in NL, the KC participants can envision that a DS should work closely with the institutional data protection officer
  - and generally cover data security, protection, and data quality control

- right level to employ data stewards within a institution
  - from an organisation centric view: when DS are employed on group level (e.g. partial role of a PhD student) they are taking the knowledge and skills with them when they leave -> no consistent competency development as a whole on the university level
  - from a people centric view: the new generations of PhD students and subsequent future professors and group leaders will incorporate these new ways of data stewardship and eventually reform the organisations from bottom up

- career paths?
  - currently there seems to be a lack of career and possibilities to "move upwards" on a career path as data steward
  - a mix of roles with data scientist, manager or similar in addition to data stewardship seems possible
  - with regards to possible career paths for former scientists:
    - a. it seems that data management will play bigger roles in being able to receive tenure positions as well;
    - b. being able to provide data management skills on top of domain knowledge could be a door opener for a lot of scientific staff;
c. There will be an increasing need to teach good (and domain specific) RDM, FAIR data and Open Science, therefore a lot more education and teaching path will open in the future (as showcased by the excellent RDM course of University Bielefeld since 2013)

- goFAIR or be square
  - it appears that goFAIR could be the missing link to enable collaborations and knowledge exchange for all the interested parties in Germany (and other countries)
  - goFAIR Germany could synergise with all available initiatives and also link to funders and policy makers on national level

#4 For what and how would you like to find collaborators nationally, internationally, (inter-)disciplinary? Table host: Silvia Wissel (GO FAIR Hamburg)

Participants shared which networks regionally, nationally or internationally they are already involved in relation to FAIR RDM in and what else might be needed. The participants stated current involvement in the following

- RDA
- RfII
- Liber
- TDWG
- National Open Science Initiatives
- Regional networks e.g. Thüringen

Some people in the group saw needs

- To be able to speak with one voice
- Centralized platform
- FAIR advice services (even if paid-for)
- Connecting the dots: the many resources which are available should be connected
- Funding for network events
- Exchanging national ways of implementing Open Science