

ORGANISING HIGH QUALITY RESEARCH DATA MANAGEMENT SERVICES FOR OPEN SCIENCE



Organisers: *Peter Wittenburg* (RDA/ MPCDF), *Daan Broeder* (EUDAT/ Meertens Institute), *Ellen Leenarts* (EUDAT/ DANS), *Maggie Hellstrom* (ICOS/ ULUND), *Stefano Nativi* (RDA/ CNR), *Rebecca Grant* (RDA/ National Library of Ireland), *Simone Sacchi* (RDA/EUDAT/LIBER)

Moderator:

Simone Sacchi (RDA/EUDAT/LIBER)

Presenters:

- *Daan Broeder* (EUDAT/ Meertens Institute) | [Service provisioning for excellent sciences](#)
- *Stefano Nativi* (RDA/ CNR) | [Brokering services facilitating interoperability and data management provision](#)
- *Wolfram Horstmann* (LIBER/UGOE) | [Data services from the librarian's perspective\(s\)](#)

Duration: 1.5 hours

Workshop description:

Open science needs high quality data management where researchers can create, use and share data according to well defined standards and practices. this is one of the pillars of Open Science. In the data management landscape we find quite a few organisations that aim at achieving this, however to get it right, a collaboration is called for where all can play a suitable role and present this in a consistent way to the researcher.

Libraries and librarians are playing a pivotal role in enabling services and nurturing practices in support of a more open and transparent approach to research., one that is in line and contributes to the emerging European Open Science vision. In this context, libraries and librarians are increasingly called up on to help managing the ever increasing quantities of data produced by research groups and communities. Libraries are excellently positioned to advise e.g. university based research groups and individual users, but currently some libraries still lack long time experience and an existing e-infrastructure to help doing this.



Figure 1: Simone Sacchi introducing the workshop

At the same time we have the experiences of a number of initiatives and e-infrastructure projects that have a long time experience working on data management services for research. This is augmented by a number of 'accepted' best-practices development that are valid cross-domain such as the FAIR data principles. The Research Data Alliance (RDA) is currently the main platform for discussion and breeding ground for data management best-practice development. The EUDAT initiative is currently providing general data management services rooted in principles developed or recommended by the RDA. This workshop goal was to discuss how RDA, EUDAT and LIBER can collaborate to provide high-quality data management to researchers.

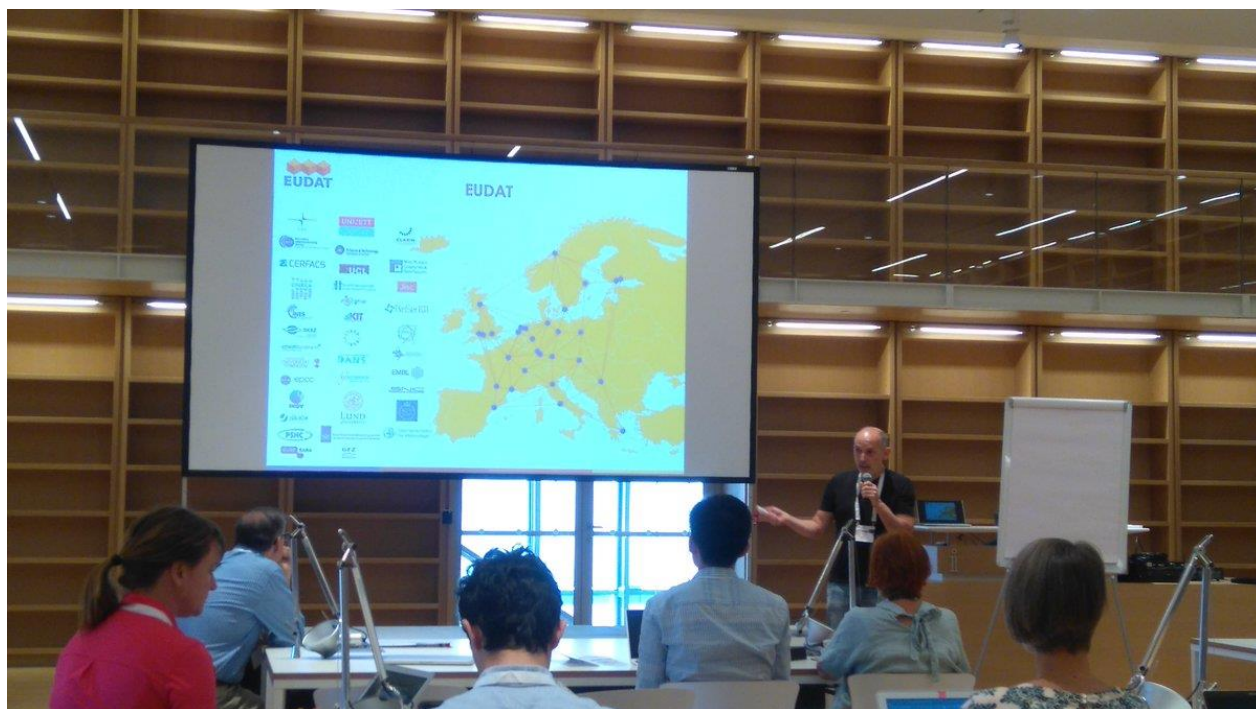


Figure 2: Daan Broeder presenting

Workshop outcomes:

The workshop was very well attended with 66 participants plus organizers and presenters. The workshop started with an introduction to the theme of the workshop by Simone Sacchi, followed by three presentations by Stefano Nativi, then Daan Broeder, and then Wolfram Horstmann. The three presentations reflected the different organizations represented, and their different roles in fostering the adoption and successful implementation of high quality research data management services in support to Open Science. Stefano Nativi for focused on a concept developed at RDA to brokering services to facilitate data management interoperability, Daan Broeder on RDM service provisioning in EUDAT and the broad community engagement around them, and Wolfram Horstmann on the role of libraries in RDM.

The presentations was followed by an interactive session with the presenters and the audience engaged in a vivid discussion. While quite different, these presentations reveled “hooks” that enabled the identification, and potential leverage, of synergies and collaborations, with clear themes emerging: we need to harmonize activities, clearly distinguish roles in the process to avoid unnecessary operational overlap, while strengthen collaborations according to the respective strengths. Specifications, implementations, and communities: we need to go from the abstract, to the development and operations, to the training and outreach. Each play a significant role, and (some) research libraries are active throughout the spectrum and the entire research lifecycle. It was also reiterated that diversity is there for a reason: the landscape is complex and we need balance between generalizable services and specific disciplinary community needs. We need to work together to leverage our networks and point researchers to the best solution to fit their needs



Figure 3: Stefano Nativi presenting

Stress was given to the modes of collaboration, with request fro clarification on how standardization bodies like RDA and e-infrastructures like EUDAT can affect and influence each other (i.e. by adopting RDA specifications on the one end, to adapting to user feedback on the other end).

The LIBER “Ten recommendations for libraries to get started with research data management”(2012) was mentioned by a participant as a mean to start promoting alignment and synergies, with examples form Portugal and the collaborations between research libraries, their institutions and the government. This comment was also followed by similar propositions from Sweden, where also international collaboration was put at the front line of success, mentioning again the active participation of research libraries in RDA.

Discussion on policy and governance to support the uptake of effective RDM services concluded that the institution level is the most effective one, and the one where the majority of policies and implemented, and the role of research libraries, as members of their own academic institutions, was deemed important to move forward the adoption of RDM policies that comply with open science principles.



Figure 4: Wolfram Horstman presenting

Final comments were made with respect to the different approach that different disciplines bring to the subject of effective RDM services, and how local and national disciplinary communities are independently moving forward. A clear conclusion is that the situation is very complex and as broad as the range of stakeholders involved. The landscape is extremely heterogenous, that's why there are multidisciplinary and more generic initiatives.

The discussion at the end was lead by a specific question from the audience: "Where do you want to go if you have data?" This question nicely summarizes the main concern when it comes ot RDM services: the complex landscape in terms of options, actors, and possibilities, creates unnecessary confusion and lack of clarity within the different research communities, especially those in the long tail of science. Suggestion was to focus on the boundaries of the institutional level and try to organize it at national level, with the caveat that from a financial perspective a generic initiative makes more sense. We will see in the next few years where the landscape is going, but it is important to keep steering the discussion towards a more effective use of resources, adoption of shared (e.g. FAIR) principles, and interoperable services.