



# Data Curation & FAIR Data

**Rob Grim**

**June 23rd, 2020**

**Erasmus University Rotterdam (EUR)**



@LIBEREurope

# Data Curation Task Group

## **WG Members**

Rob Grim, Mari-Elisa Kuusniemi, Eliane Blumer, Philipp Conzett

## **Work plan focus**

1. Identify main aspects of data curation
2. Identify different roles of library staff in data curation, including how library staff can help researchers to create and publish FAIR datasets
3. Formulate recommendations on how to integrate data curation into library trainings



# Data Curation & Research Libraries

## Main aspects

1. Generic vs. specialized data curation services
2. Data curation skills and profiles: *basic, experienced, advanced*
3. Data curation vs. data publishing
4. Engaged communities

## Practical aspects

1. LIBER libraries don't have to start from scratch! E.g. CURATE network, DCC
2. Basic data curation activities resemble the work of metadata specialists
3. Many tools and platforms for data curation are already available e.g. Dataverse, Figshare, etc.
4. Many challenges ahead!



# Three Professional Data Curation Profiles



# 1. BASIC Data Curator (Msc)

The data professional has a basic understanding of common data formats used in science practice, has a basic understanding of metadata vocabularies.

Partly comparable to a metadata specialist and a cataloger

Manual curation tasks: dataset metadata entry, simple checks on datasets and supplementary materials .

Works with default metadata templates provided by data curation and annotation tools.



## 2. Experienced Curator (MSc, PhD)

**Synonyms: Experienced/Specialist Data Curator**

Has a good understanding of disciplinary practices or an advanced generic understanding of specialized disciplinary tools and statistical software and tools

This data professional can run code and can work e.g. on replication studies, specialized curation projects



### 3. Advanced Specialist (Msc, PhD)

Minimum of 3 yrs. experience

Data professional is able to code/program

Can *interoperate* with software engineers, research computation, and machine curation

Write programs e.g data annotation tools, comfortable with cloud computing infrastructures and platforms, familiar with statistical tools for data anonymization, data labeling and clustering





# Data Curation: Recommendations

- Articulate what data curation entails in the library and make the focus areas clear, e.g. archiving, preservation, support for publishing, special collections, topical collections, promoting reuse, etc.
- Establish a community of practice among LIBER libraries for the development of both generic and specialized curation services, comparable to e.g. CURATE network in the US
- Map the curation profiles to FAIR tasks and challenges
- Advocate the use of the curation profiles - basic, experienced and advanced - in research libraries
- Create and participate in projects where these curation profiles can be applied





# FAIR Data Task Group

## **WG Members**

Dr. Dragan Ivanovic, Eliane Blumer, Philipp Conzett, Liisi Lembinen, Dr. Žibutė Petrauskienė, Mari-Elisa Kuusniemi, Rob Grim & (Dr. Mijke Jetten)

## **Why update fact sheet on FAIR data?**

### **Planned deliverables**

1. Updating the LIBER 2017 factsheet on FAIR
2. A template for implementing and fostering FAIR data solutions
3. Explore FAIR opportunities for research libraries to support ongoing and after-the-fact research?



# Update fact sheet FAIR data

**F: Invest in developing search and discovery expertise in research libraries**

**A: Articulate the position and role of research libraries in supporting access to research data, *integration* with data platforms and its role in supporting digital data *conservancy***

**I: Create links to research information systems (CRIS, RAIDS), e.g. link DMPs to CRIS and data curation practice**

**R: Set up a community of practice within LIBER to build expertise around the *reuse of data, replication studies and data curation.***





# Thank you!



@LIBEREurope