Data papers – an introduction

Joachim Schöpfel
University of Lille, France

Workshop "Data Papers", NTK Prague, October 23, 2018

Research data

"Factual records (numerical scores, textual records, images and sounds) used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings" (OECD)

Archived data (485)
Audiovisual data (415)
Configuration data (59)
Databases (418)
Images (1093)
Networkbased data (126)
Plain text (1020)
Raw data (969)
Scientific and statistical data formats (1373)
Software applications (388)
Source code (120)
Standard office documents (1305)
Structured graphics (791)
Structured text (675)
other (775)

re3data.org

Research data management

biblio.uottawa.ca

Hot and cold data

hot
create, process, analyze

cold
preservation, publishing

metadata (standards: DataCite etc.), identifiers (DOI etc.), link to paper

Why publish data?

- research policy (institution, country, EU)
- funding agencies (including EC/H2020 program)
- academic journals

Data papers

"A data paper is a searchable metadata document, describing a particular dataset or a group of datasets, published in the form of a peer-reviewed article in a scholarly journal"

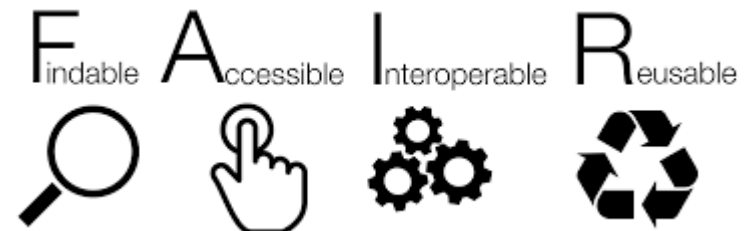(Global Biodiversity Information Facility)

Where and why

Where

- Repository
- Academic journal
- **Data journal**

Why

- To improve findability of data
  - cf. FAIR principles to increase machine readability of published datasets
- To increase impact

Format

Content

- Standard structure

- Description of dataset

- Link to dataset

- Short paper

- But a large diversity of formats

Link with data repository

- Metadata (export to data paper template)

- Enrichment
  - including elements from data management plan

- Importance of standards
  - http://www.dcc.ac.uk/resources/metadata-standards/list

Mandatory	Recommended	Optional
Identifier	Subject	Language
Creator	Contributor	Alternate ID
Title	Date	Size
Publisher	Resource Type	Format
Publication year	Related identifier	Version
	Description	Rights
	GeoLocation	

Data journals

Including data papers

- BMC Research Notes

- F1000 Research

- GigaScience (OUP)

- PLoS One

- Frontiers in Sociology

Exclusively data papers

- Data in Brief (Elsevier)

- Scientific Data (Nature)

- Journal of Open Archeology Data (Ubiquity)

- Genomics (Elsevier)

https://www.fosteropenscience.eu/foster-taxonomy/open-data-journals

Thank you !

Corresponding author:
joachim.schopfel@univ-lille.fr