DATA CITATION ROADMAP FOR FINLAND
Recommendations for National Stakeholders

www.fcrd.fi/data-citation

About the project

Finnish Ministry of Education and Culture Open Science and Research initiative tasked the Finnish Committee for Research Data (national committee of CODATA) with producing recommendations on introducing data citation practices to the Finnish research system. The project has used the FORCE11 Data Citation Synthesis Group: Joint Declaration of Data Citation Principles as a key point of reference and a framework for evaluating national state of the art. This poster summarizes the first published draft of the recommendations. They will be co-developed with stakeholders from the national research community during the fall of 2017. The recommendations and stakeholders are presented in no particular order.

Research institutions

- Include principles of data as evidence and data transparency into enforceable institutional data policies.
- Include principles and examples of data as evidence and data transparency to national research ethics MOOC and open science web course.
- Support and explore the development of data metrics in research evaluation. When implementing new metrics, pay special attention to the transparency of data and methods.
- Create and enforce institutional policies on licensing data, recommended licenses (i.e. CC-BY), and templates for data ownership agreements.
- Include addressing data authorship and ownership relevant questions to data management planning requirements.
- Include introduction to persistent identifiers, both as a concept and a practice, into basic researcher training, preferably starting already in the methods courses for undergraduate students.

National scholarly publishers

- Present all authors with a publication specific data reference model based on the recommendations made by this project (see website), and require its use.
- Include a hyperlink, preferably the PID, to underlying data description for all original research publications.
- Create discussions about possible national applications of the FORCE11 Data Citation Roadmap for Scientific Publishers, and Transparency and Openness Promotion (TOP) guidelines.

Data repositories (data centers, archives, libraries)

- All datasets intended for citation must have a globally unique persistent identifier that can be expressed as unambiguous HTTP URL.
- Finnish data repositories should use either DOI or URN as their PID of choice, since they are the best managed and most reliable PIDs in the Finnish environment.
- This persistent identifier must resolve to a landing page that supports access to the actual data set.
- Assigning PIDs and creating landing pages is the responsibility of the data repository.
- Landing page should facilitate access to metadata, either by holding metadata or a link to metadata.
- The landing page should include reference model for citation, and ideally also metadata helping with discovery, in human-readable and machine-readable format.
- National data centers, libraries and archives should agree on the required metadata content of a data landing page.
- Data that no longer exists should have a persistent landing page, which may direct the user to a current version of the old data set.
- License all metadata with a CC0 license or equivalent.
- Make metadata freely harvestable through open APIs.
- The persistent identifier must be embedded in the landing page in machine-readable format.
- Pilot the RDA data citation model for dynamic data in one or several national data centers.
- Release all data citation related content intended for broad audiences, such as guidelines and standards, in open format, i.e. CC-BY, or equivalent.

Researchers, learned societies

- Include principles of data as evidence and data transparency in next version of Finnish responsible conduct of research guideline.
- Recognize data authorship as a distinct issue and discussion in the upcoming national authorship guideline.
- Create multi-institutional, multi-disciplinary action to define principles for defining data authorship, or assign national representation to international action with the same goal.
- Organize multidisciplinary discussion on data management and citation, with the aim of creating interoperable practices.
- Promote the use of data reference model also when referring to authors own primary source data.
- Define field specific level of granularity for data citation.

Funders, policy makers

- Make data management planning required by all research funders, either in the application stage or after funding is granted.
- Explore mechanisms for evaluating the quality of published data sets for the purpose of assessing the impact of research institutions.
- When allocating research funding, take all research outputs into considerations instead of just publications, i.e. in the vein of US National Science Foundation (NSF), that asks a principal investigator applying for funding to list his/her research “products” rather than “publications”.
- Where relevant, adopt a (good quality, well described) data publication as the sole output from a research project.
- Give consistent, long-term support to data infrastructure necessary for data citation and access.

Contact: Project coordinator Heidi Laine, heidi.laine@helsinki.fi / @heidiklaine / +358405139593