

Finnish Research Environment and Infrastructure at the Service of Citizen Science

Konsta Happonen

Open Knowledge Finland - Open Science Working Group

@Koalha

Research Environment and Infrastructure

What digital services that support open citizen science exist or are being developed?

How could they be better?

How do the guidelines for research conduct take into account citizens as researchers?

How do copyright laws affect crowd-sourcing research?

DMPTuuli for data management plans

Specialized services for crowd-sourced data collection

Finnish data repository IDA not open for citizen scientists

European EUDAT and ZENODO are.

International NGO-operated Dryad supports versioning of datasets.

Development needs

DMPTuuli recommends CC-BY licences for data - unnecessarily restrictive.

Adequate European services for archiving complete datasets exist.

No services for reliable collection of general observational data currently exist.

Current guidelines, citizens and researchers

Existing guides and guidelines are mostly as relevant for citizen scientists as they are for full-time researchers.

Responsible conduct of research (TENK): principles regarding authorship, data storage and data usage rights should be negotiated in the research group beforehand.

Agreements of participating citizen scientists on said policies should be gotten in some traceable form, for example by email or by making data collectors agree to terms of use of a data collection interface.

Copyright law and crowd-sourced science

Current Public Administration Recommendations (JHS) state that data should be licenced under a CC-BY 4.0 licence. JHS recommendations are a widely influential and probably the reason why DMPTuuli recommends it.

CC-BY imposes unneeded restrictions on the use of data. The legal text of CC-BY is complicated and imposes restrictions for the derivative use of data.

The moral rights of data authors, such as right to attribution, are more easily protected by the internal control mechanisms of science than by complicated licences. Public data providers should care more about the usability of their data products than if their names are mentioned in the derivative works.

Copyright law and crowd-sourced science

One of the promises of decentralized research is that huge numbers of small data flows can combine to make large-scale observational science easier.

A key requirement for this is that data can be easily combined from several sources with minimum hassle with licences. CC-BY licences do not fulfill this condition.

This is why data overall and data from citizen science projects especially should be released to the public domain with, for example, a CC0 dedication.

Infrastructure: top-down or bottom-up?

National infrastructure for well-defined, broadly interesting observations. (Top-down)

Training in the use of moldable, less standardized, more lightweight solutions such as wikis (on all levels).

Important to remember: standardized national infrastructure can't mean there's no functionality for the user. Motivation, motivation motivation!