

Supporting Text 1

Appendix 1

Literature searches

Research questions: 1) What challenges are CS projects facing in terms of data management? 2) Are the FAIR principles applied for data in CS projects? The research questions were searched simultaneously as shown below. Besides this, a Google search and a search in BASE (Bielefeld University Library, no date) was performed and focused on the terms “citizen science” and “data management” and was ended on 11 March 2020. Datasets and non-English literature were omitted from the search results.

Scopus	#1 TITLE-ABS-KEY(“citizen science”)
	#2 TITLE-ABS-KEY((data W/2 management) OR RDM OR DMP OR legal OR ethic* OR "intellectual property rights" OR IPR OR licens* OR (data W/2 sharing) OR publication OR publish* OR re*use OR (FAIR PRE/1 principles) OR metadata OR findable OR findability OR (data W/2 access*) OR interoperab* OR "code of conduct" OR repository OR (best W/0 practice*) OR guideline\$ OR (data W/2 storing) OR (data W/2 storage))
Performed 03.03.2020 excluded: data papers limited to language: English	#1 AND #2 = 1107

Web of Science	#1 TS=”citizen science”
	#2 TS=((data NEAR/2 management) OR RDM OR DMP OR legal OR ethic* OR "intellectual property rights" OR IPR OR licens* OR (data NEAR/2 sharing) OR publication OR publish* OR re-us* OR (FAIR NEAR/1 principles) OR metadata OR findable OR findability OR (data NEAR/2 access*) OR interoperab* OR "code of conduct" OR repository OR (best NEAR/0 practice*) OR guideline\$ OR (data NEAR/2 storing) OR (data NEAR/2 storage))
Performed 03.03.2020 excluded: data papers limited to language: English	#1 AND #2 = 630

References

Bielefeld University Library (no date) *What is BASE?*, <https://de.base-search.net/about/en/index.php>. Available at: <https://de.base-search.net/about/en/index.php> (Accessed: 14 September 2020).

Appendix 2

Screening of publications

All records were imported to Rayyan (Ouzzani *et al.*, 2016) and title and abstracts screened independently by the authors by allocating the authors to two “screening” groups. Discrepancies were determined by JSH. Records selected for full-text screening were read and selected by the authors. JSH extracted data from eligible articles.

Eligibility

Publications were included if they describe discussions, challenges, or investigations of or tools or guidelines aimed at RDM practices in CS projects. If adherence to the FAIR principles was described in relation to CS projects, the publication was also included. If the publications had a main focus on other aspects than RDM, but still contained relevant RDM challenges, they were included.

Publications describing development, handling or presentation of individual projects, project databases or results are generally excluded if they do not directly describe RDM challenges and/or their solutions. Documenting quality control and data validation may be regarded as part of the data management of CS projects. However, this aspect is omitted from the literature search and analyses, because it is more related to the precision and reliability of the data and is widely discussed in other papers (Kosmala *et al.*, 2016; Bowser *et al.*, 2020). Articles evaluating user accessibility, recruitment and engagement or uptake of technology by users are excluded as was also slide presentations (e.g. powerpoints) and short conference abstracts.

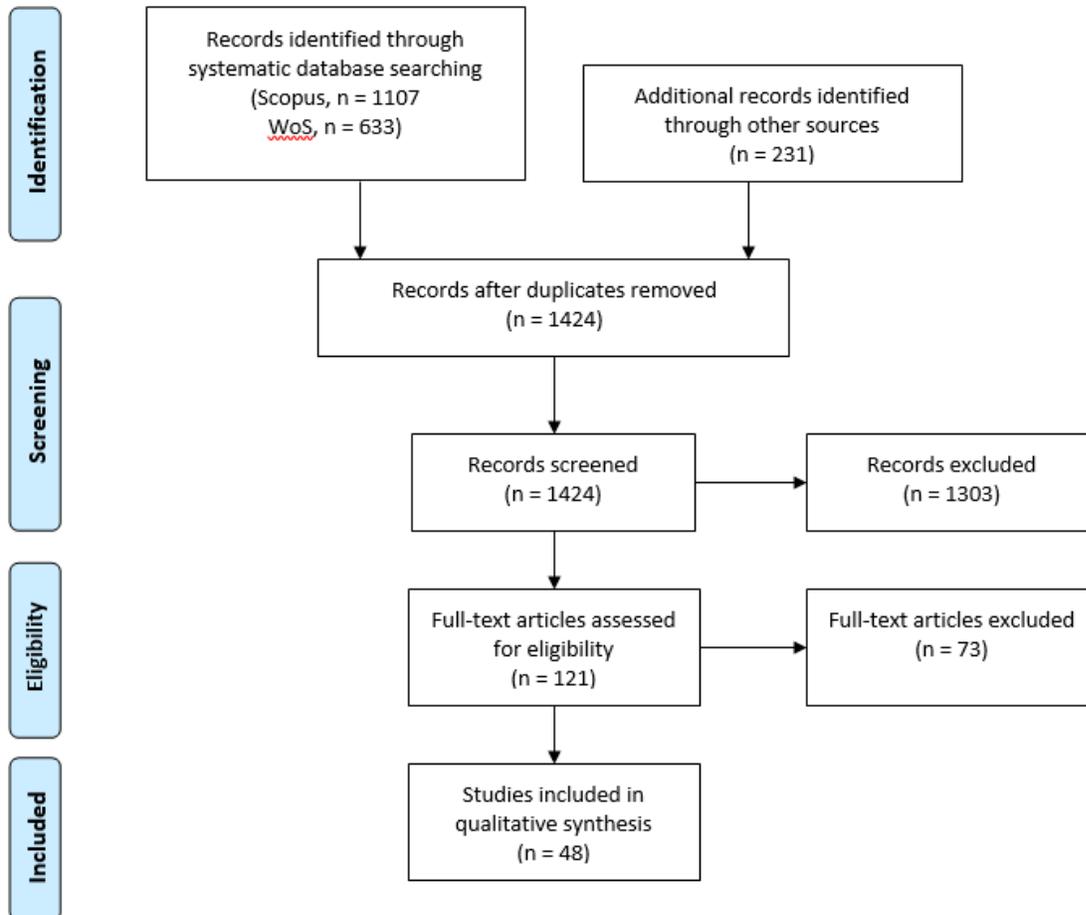


Figure 1 PRISMA diagram of systematic literature search and selection.

References

Bowser, A. *et al.* (2020) 'Still in Need of Norms: The State of the Data in Citizen Science', *Citizen Science: Theory and Practice*, 5(1), p. 18. doi: 10.5334/cstp.303.

Kosmala, M. *et al.* (2016) 'Assessing data quality in citizen science', *Frontiers in Ecology and the Environment*, 14(10), pp. 551–560. doi: 10.1002/fee.1436.

Ouzzani, M. *et al.* (2016) 'Rayyan—a web and mobile app for systematic reviews', *Systematic Reviews*, 5(1), p. 210. doi: 10.1186/s13643-016-0384-4.

Appendix 3

The semi-structured interviews were based on the following questions:

- Describe the project
- Describe the data flow of the project
- Are the data described by relevant metadata?
- How are the data documented?
- What data can the citizens/participants/the public access?
- If data are shared then describe how, and to whom?
- Does/did anything prevent data sharing?
- Who owns the data?
- How is data quality ensured?
- How are personal data handled?
- Did you seek and find help for data management at your institution?
- Did you create a data management plan?
- Are you aware of the FAIR principles?

Answers were noted by the interviewer in Danish. The interviews are therefore not available as transcripts. The information presented in Table 5 was extracted from the notes and approved by the interviewees.