

EGU22-5147, updated on 12 Aug 2022

<https://doi.org/10.5194/egusphere-egu22-5147>

EGU General Assembly 2022

© Author(s) 2022. This work is distributed under the Creative Commons Attribution 4.0 License.



Distributed databases for citizen science

Julien Malard-Adam¹, Joel Harms², and Wietske Medema²

¹UMR G-EAU, IRD (Institut de recherche pour le développement), Université de Montpellier, Montpellier, France

(julien.malard@mail.mcgill.ca)

²Génie des bioressources, Université McGill

Citizen science is often heavily dependent on software tools that allow members of the general population to collect, view and submit environmental data to a common database. While several such software platforms exist, these often require expert knowledge to set up and maintain, and server and data hosting costs can become quite costly in the long term, especially if a project is successful in attracting many users and data submissions. In the context of time-limited project funding, these limitations can pose serious obstacles to the long-term sustainability of citizen science projects as well as their ownership by the community.

On the other hand, distributed database systems (such as Qri and Constellation) dispense with the need for a centralised server and instead rely on the devices (smartphone or computer) of the users themselves to store and transmit community-generated data. This new approach leads to the counterintuitive result that distributed systems, contrarily to centralised ones, become more robust and offer better availability and response times as the size of the user pool grows. In addition, since data is stored by users' own devices, distributed systems offer interesting potential for strengthening communities' ownership over their own environmental data (data sovereignty). This presentation will discuss the potential of distributed database systems to address the current technological limitations of centralised systems for open data and citizen science-led data collection efforts and will give examples of use cases with currently available distributed database software platforms.