



User interface design principles for peer-to-peer distributed databases for ecological citizen science projects

Julien Jean Malard-Adam^{1,2,3,i}, Wietske Medema⁴, Nallusamy Anandaraja^{3,i}, Joel Harms⁵,
Johanna Dipple⁵, Sheeja^{5,i}, and Palanivelan Jaisridhar^{6,i}

¹G-Eau, Institut de recherche pour le développement – IRD, Université de Montpellier, Montpellier, France

²Institut français de Pondichéry – IFP, Puducherry, India

³Directorate of Extension Education, Tamil Nadu Agricultural University, Tamil Nadu, Coimbatore, India

⁴University of California, Berkeley, California, USA

⁵Department of Bioresource Engineering, McGill University, Montréal, Québec, Canada

⁶ICAR – Krishi Vigyan Kendra, Tamil Nadu Agricultural University,
the Nilgiris, Tamil Nadu, Doddabetta, India

ⁱThe original author and affiliation names can be found at the end of the paper.

Correspondence: Julien Jean Malard-Adam (julien.malard@mail.mcgill.ca)

Received: 25 March 2025 – Revised: 18 July 2025 – Accepted: 24 July 2025 – Published: 28 October 2025

Abstract. Citizen science, where participants from outside of academia contribute to data collection or analysis, is an important approach in ecological studies that can significantly improve both modelling outcomes and community participation. However, all ecological citizen science platforms developed to date rely on centralised server architecture for data storage and communication with citizen scientists, which can lead to unsustainable server maintenance costs for project managers as well as data sovereignty issues for the concerned communities, thereby endangering project resilience and sustainability after the end of a funded project. Distributed databases, which rely on peer-to-peer technology to store and share data, can address these concerns, but they are complex and conceptually different from centralised systems. As such, their use involves a very steep learning curve that hinders their adoption by citizen science practitioners in ecology, where neither project leaders nor end users are experts in peer-to-peer technologies. In this article, the authors use formal and open-ended feedback from workshops with academics to discuss how well-planned user interface design can be used to facilitate the adoption of peer-to-peer distributed databases in citizen science and provide generalisable key recommendations for the implementation of user interfaces in citizen science applications. In particular, we discuss several key conceptual differences between centralised and distributed applications, such as key-pair authentication and eventual consistency, that must be efficiently and visually communicated to end users. While there is extremely limited literature available on user interface design for distributed systems (and none so far in the ecological field), we find that lessons learned from other fields transfer well to the field of ecological citizen science, that well-designed user interfaces are key to the adoption of new technologies, and that simplicity and efficiency in interface design are more important than showing average users the details of how the underlying technology works. We propose these recommendations as a blueprint for future research and development of citizen science applications based on peer-to-peer distributed database technologies.

1 Introduction

Digital technologies offer significant potential to strengthen biodiversity and conservation efforts across the globe (Freitas and Gouveia, 2025), and citizen science, where individuals or communities not traditionally involved in academia or research institutions contribute to the scientific process (for instance, data collection, hypothesis generation, analyses, or policy design), has important potential to improve both the scope of ecological data collection and the involvement of partner communities in the research and in the application of its outputs (Cieslik et al., 2019; von Gönner et al., 2024; Khu et al., 2023). While many forms of citizen science participation exist, most methodologies involve, at least in part, the use of a community of citizen scientists to either collect or digitise data for use in a research project (Haklay et al., 2021; Lee et al., 2020), often (though far from always) through the use of a digital platform such as a website or mobile phone app (Teacher et al., 2013).

Applications of citizen science span many fields, most notably ecology (Fraisl et al., 2022) and hydrology and natural hazards (Lee et al., 2020). These fields have in common the frequent use of highly geographically and temporally dispersed data, which makes crowdsourcing an interesting solution to the cost of realising large-scale data collection (Avelaneda et al., 2020); these fields are also notable because of the extent to which research often aims to modify non-academic citizens' behaviours in response to evolving information (Hall et al., 2024). The field of hydrology has many examples of successful citizen science projects, including in flooding (See, 2019), streamflow and water level monitoring (Lowry et al., 2019; Seibert et al., 2017; Strobl et al., 2019), pollution monitoring (Moshi et al., 2022), and disaster management (Hicks et al., 2019). Important contributions to hydrological modelling and flood forecasting have also been obtained from citizen science and crowdsourced data (Avelaneda et al., 2020; Lumbrazo et al., 2022; Mazzoleni et al., 2017; Mengistie et al., 2024; Songchon et al., 2023). While the use of citizen science for live forecasting is less common in the ecological domain, ecological applications of citizen science tend to employ user-contributed observations for species detection and monitoring, including machine learning models for species identification from user photographs (Bonnet et al., 2024). These examples include BirdNET for avian populations (Wood et al., 2022), Pl@ntnet for botanical data (Joly et al., 2016; Lefort et al., 2025), and iNaturalist for entomology (Pernat et al., 2024).

These project setups work very well when the primary goal of the activity is to collect data from a large pool of volunteers for use in a particular scientific project, and the main concern is the potential for the motivation of citizen scientists to wane before that of academics. Significant research has accordingly been dedicated to developing user interfaces that improve user interest and data contributions to the project (Giuliana, 2017; Liu et al., 2021; Sprinks et al.,

2015). When the potential for citizen scientists to feel invested in the project and use the collected data and scientific results to inform their behaviour or decisions in an adaptive management framework is instead an important goal of the project, however, the opposite concern – that of academics' interest (or funding) waning before that of participating communities' – becomes an important consideration. As citizen science should extend beyond simple data collection to more empowering aspects of the scientific process such as methodology and policy design (Hall et al., 2024), data sovereignty and the sustainability of the software tools developed for data collection are key challenges to maintaining long-term community-led citizen science projects. Indeed, the financial sustainability of such projects is a well-known problem, since, once academics have moved on, the platforms used to receive, store, and transmit citizen science data require regular expenditure to maintain server and data costs, which is generally not feasible for communities to maintain (Liu et al., 2021). If data have been archived on a public research archive, it may remain technically feasible for communities to access them; even so, new data cannot be contributed and citizen science apps or websites built to access data from a dedicated project server will cease to function, raising issues of project sustainability and of data ownership and sovereignty for communities. Alternatively, the platform may continue in a commercial form (Liu et al., 2021), which, depending on the nature of the project and the type of community data involved, may involve conflicts of interest that jeopardise the initial goals of the research.

Distributed (peer-to-peer) databases have been recently proposed as a solution to these issues, since these databases do not require a central server and can function directly between users' end devices (phone or computer) (Malard-Adam, 2024, 2025; Medema et al., 2024). As of now, very few examples of distributed or peer-to-peer scientific database software exist, with the notable exception of qri, which ceased operations in 2022 (Parsons et al., 2019). A new software tool, Constellation, has been recently developed for distributed scientific databases (Malard-Adam, 2025), which differs from qri in that it aims to offer *mutable* database functionalities (i.e., databases that can be continuously updated) for rapidly evolving citizen science projects characteristic of ecological applications. Data mutability means that the data contained in a given database can change without requiring the database's unique identifier or address to also change, which allows databases to keep evolving over time as new data are collected (e.g., as in a Google Sheets document), contrarily to *immutable* databases, which require a new version and associated URL or other identifier to be regenerated if the data change (e.g., as with Zenodo and other data repositories).

As no centralised server is used to manage users and data, however, three main functionalities generally provided by servers, namely access control, data storage, and data discovery, must be reimplemented according to a distributed-

system philosophy. Although the approaches used to implement these functionalities in distributed systems avoid many limitations of centralised systems (Malard-Adam, 2024), they also require a significant change of mental model from the currently nearly ubiquitous centralised server–client paradigm. Among the most important differences are (1) the lack of a single source of truth (copies of data on different end-user devices are not ephemeral but rather equally valid sources of data for sharing with further users), (2) eventual consistency (there is no guarantee of a unique absolute state to the system at any given time, but if users remain connected to each other for long enough, they will eventually converge to a single shared state), and (3) content-based addressing (data are accessed through keys that correspond to the desired content, rather than according to their location, as is commonly done with domain names) (Perry, 2020). In many ways, these characteristics more closely mirror real-world human-to-human interactions (Malard-Adam, 2024); however, they also risk introducing unnecessary layers of complexity and cognitive load if not appropriately abstracted away from end users.

The adoption of new and alternative technologies, as such, is highly dependent on the availability of user-friendly interfaces, particularly when similar (even if less performant) alternative technologies exist, as exemplified by the experiences of tool developers in related fields such as statistics (Lunn et al., 2009; Salvatier et al., 2016). User interface design is also known to have an important impact on the quality of data collected (Torre et al., 2019). At the same time, there is very little literature on user interface design for citizen scientists, and there is none at all on design in the context of the newly emerging field of citizen science powered by peer-to-peer distributed databases, whether for the citizen scientists collecting the data or for the academics leading the project. This is a particularly critical area for ecological research, since distributed systems, despite their advantages for resilience and data sovereignty, require a paradigm shift from the dominant server-centric data worldview and risk alienating potential users if their interfaces are not designed to minimise user exposure to more complex aspects of peer-to-peer technologies that are not relevant to their immediate needs.

In this paper, the authors present their experiences in the participatory development of a graphical user interface for Constellation, a new software tool for distributed scientific databases that can be used both to share existing databases and to collect new data in a citizen science framework. In the case of the latter, the main goal of Constellation is to create a streamlined platform for managing citizen science project data in a peer-to-peer manner that avoids dependencies on centralised servers for data management and safeguarding. This involves two main aims: (1) academic-oriented interface functionalities for developing and managing citizen science projects and (2) citizen-scientist-oriented interfaces that guide their contributions to a particular citizen science project. The surveyed literature on user inter-

face design for citizen science in centralised systems concentrates on the latter, though the former is also critical to ensuring academics without advanced computer knowledge skills can effectively deploy a citizen science platform for their projects. While the Constellation platform can also be used to store and distribute scientific data that do not originate from a citizen science initiative, this article will concentrate on the citizen science use case.

This paper presents the authors' experiences in developing the current user interface for Constellation in collaboration with end users in the context of ecological citizen science, and it proposes a set of generalisable recommendations and "lessons learned" for user interface design for citizen science applications in the emerging field of distributed scientific databases, with the expectation that these guidelines will help ensure that peer-to-peer scientific database tools are rapidly adopted by the larger ecological scientific community instead of remaining a niche product for distributed-system enthusiasts.

2 Approach and goals

The results in this section stem from the authors' experience in developing Constellation's interface and user-testing it in various forums and workshops while receiving formal and informal feedback from participants during the training and interactive activities. In these activities, participants were presented with a brief introduction to the general concepts of distributed databases before engaging in an interactive session using Constellation to complete basic database tasks (account creation and data upload and sharing).

Since Constellation is a peer-to-peer distributed database system for citizen science, the needs of two main user types must be met – firstly, the academics setting up citizen science projects and, secondly, the citizen scientists participating in these projects. The needs of these groups are very different; academics need an interface to define variables to be collected and to build data tables, while citizen scientists require more directed interfaces to correctly enter the data that they collect (e.g., photos or numbers) with a maximum level of automation to remove repetitive tasks (e.g., geolocalisation of observations, automatic filling of observation dates and times). In addition, priorities for data access and visualisation will differ; academics are more likely to prioritise automated or programmatic access to data for data science and formal analyses, while citizen scientists will have their own priorities and interests in data, particularly in visual interfaces (maps, graphs).

Some (centralised) citizen science platforms have opted to combine both roles in one application; KoboToolBox, for example, allows academics to develop a data table structure and then publish this as an online form that end users can use to contribute data (Lakshminarasimhappa, 2022); similarly, Google Forms has also been used (Árvai et al., 2023). An al-

ternative would be to have academics develop their own user interface specific to their project needs; while potentially allowing for more project-appropriate user interfaces, this approach also requires academics to access significant coding expertise and resources and adds a very high entry barrier to deploying citizen science projects (Howard et al., 2022).

In the case of Constellation, we have decided to pursue both routes, allowing for both generic form generation and customised application-specific apps. Constellation itself is divided into an API (application programming interface) as well as a GUI (graphical user interface) tailored for academics; this interface will allow academics to automatically deploy project-specific data entry form stand-alone apps for their research projects whose code can also be adjusted, if desired, for a more project-tailored user experience. In the next sections, we will discuss general concepts and challenges applicable to both academic and citizen scientist user interfaces specific to the context of distributed data platforms. While, to the best of our knowledge, no research on this topic has been published yet in any field, including ecology, the literature from user surveys in the field of distributed systems in general has proposed some general guiding principles for user interface design when porting centralised services to distributed alternatives (Kosem and Dietrich, 2020). This research concluded that users mostly prefer for distributed software to resemble, visually and in functionality, comparable centralised software alternatives; the unique aspects of distributed software should therefore only become visible to the user when absolutely necessary or when offering a unique functionality unavailable in centralised systems.

2.1 General concepts

There are several main conceptual and structural differences between distributed and centralised systems, which can lead to differences in how user interfaces are designed and how information is then presented to users. In this paper, we discuss five key differences: the concept of (not having) a single source of truth for the system's state, eventual consistency, content-based addressing, key-pair-based authentication, and concentric networking.

Figure 1 shows the key structural differences between centralised, decentralised, and distributed systems. Centralised systems, which represent the near-entirety of the internet today (e.g., websites, email, and most messaging apps), route all communication between users through a central server which provides data storage, user authentication, and connectivity between users. Decentralised systems, the most well known of which are cryptocurrency and other blockchain projects, aim to replace centralised systems with a group of peers whose trustless collaboration ensures data replication and storage. As these systems, however, were primarily designed for financial transactions, the extremely high performance, financial, and environmental costs of the consensus algorithms used to prevent “double-spending” pre-

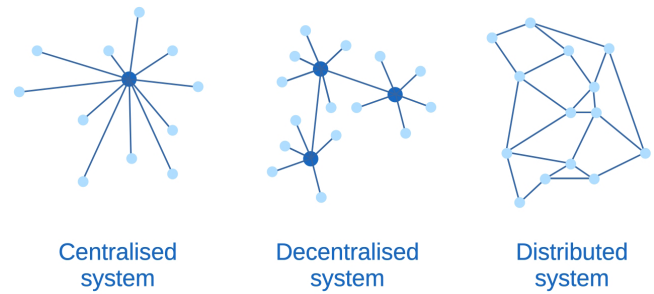


Figure 1. Structural differences between centralised, decentralised, and distributed systems. Reused with permission and minor modifications from Medema et al. (2024).

clude their use for practical data storage purposes (Weaver, 2021). Since these systems are computationally heavy, distributed networks generally become comprised of large peers (usually hosted by corporations) that network with each other and then offer services to end users with a traditional client–server interface. Distributed systems, on the other hand, are lightweight enough to fit on small end-user devices and as such allow all users to participate directly in the network (Medema et al., 2024).

2.1.1 No single source of truth

In centralised digital systems, there is usually a single “source of truth” regarding the current state of the system, generally in the form of a database on a server. This state is then updated by users with the appropriate permissions (as verified by the server) and then sent out to be viewed by any end user with read access to the data. As such, local copies of the data in users’ browsers are only meant for ephemeral access, and the central server is the single source of truth on the current state of the system. In contrast, in distributed systems, each copy of the data, whether located on a device used to create it or on one only used to access or visualise it, is equally valid as an official source of data and can be sent to other users who request it. This shifts the responsibility of data storage and safekeeping from the service provider (and its centralised server) to the users, who can locally “pin” data from other users if they wish to ensure the data’s guaranteed availability later on. At the same time, users (at least data producers) are likely to wish to see how many other users have pinned their data in order to assess the availability as well as usefulness to the community of the data they have contributed. Instead of a dichotomy of data sharing as in centralised systems (data are uploaded or not), we have instead a continuous gradient of availability (number and geographical location of nodes sharing one’s data). Constellation therefore uses icons to rapidly show the status of a database (pinned on this account, pinned on this device, not pinned) as well as more in-depth options in advanced menus (such as pinning of large media files and replication status across the network).

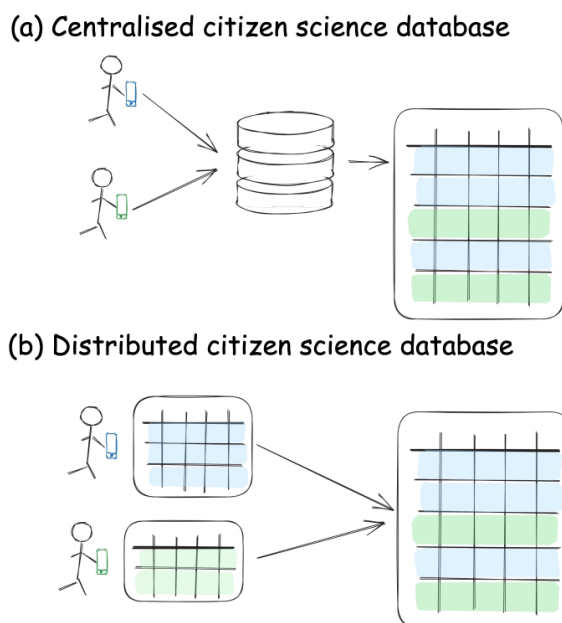


Figure 2. Structural differences between centralised and distributed citizen science databases. Adapted from the Constellation documentation (Constellation, 2025).

Figure 2 shows the structural differences between centralised and decentralised systems for citizen science databases. While centralised solutions save each user's data to a single database on a server, distributed solutions will create individual, identically formatted databases to which each user will add their own data. These databases are then visually combined to present a single dataset when visualised, but they technically remain, “under the hood”, individual and independent databases, where each user only has permission to write to their own database.

This system architecture works very well for living, evolving databases but does not lend itself well to permanent links and formal citations for scientific reproducibility. As such, the Constellation interface and API offer the option of exporting data to Excel, LibreOffice, or other formats, which can then be uploaded to standard data repositories or included as supplementary materials in a publication. (These data exports can also be configured to run either on a periodic schedule or else automatically upon changes to the data, allowing for configurable backups.) In this case, users are recommended to cite both the exported immutable (or “fossilised”) dataset for reproducibility purposes and the living dataset according to its Constellation unique identifier. This facilitates readers' access to the data as they were when the published analyses were conducted as well as to the most recent, up-to-date version of the dataset.

2.1.2 Eventual consistency

A second major difference between centralised and distributed systems is the concept of eventual consistency (Gomes et al., 2017), which means that data can be updated in parallel across different and potentially disconnected devices, with the software automatically merging changes once the devices are once again connected to each other. This is in contrast to centralised collaborative databases (such as Google Forms (Árvai et al., 2023)), where edits can only be made when devices are online so as to avoid potential editing conflicts.

2.1.3 Content-based addressing

A third major difference between centralised and distributed systems is content-based addressing. The majority of the internet works through location-based addressing (Trautwein et al., 2022), where users ask to retrieve content hosted at a specific address without consideration of what the actual content is (for example, accessing <https://appli.reseau-constellation.ca> (last access: 1 October 2025)) will request and serve the most up-to-date web page available on the server corresponding to this address). On the other hand, distributed systems can source the requested information from any user who has a local copy of those data (ideally a geographically close user to improve network response time). As such, data are requested not by a URL but rather by a unique identifier (hash) that uniquely matches the requested data (e.g., “zd-puAunUaEc6CRX8wAjNixNioB6hf5SYvwJytx4ynm5wQx1vH”) and can be used to verify the integrity of the returned data. Users of a distributed data interface for citizen science will likely need to access data contributed by others using these hashes. However, these hashes can be visually intimidating for non-experts, reflecting a need for interfaces to be designed, as much as possible, to hide these hashed links behind visual abstractions to increase the usability of these platforms, similarly to how URL addresses are often hidden behind visually aesthetic links and buttons on the centralised web. At the same time, these content-based addresses must also be available for more advanced users who wish to directly access or share a link to a specific database (for instance, as a reference in a scientific publication). Constellation addresses this by prominently displaying human-readable database names along with a small “link” icon that, when clicked, displays the database's unique identifier (Fig. 3).

2.1.4 Key-pair-based authentication

A key difference between centralised and decentralised or distributed systems is the logic behind user authentication. In centralised systems, servers are responsible for verifying users' identity, generally through usernames and associated passwords or other authentication procedures (e.g.,

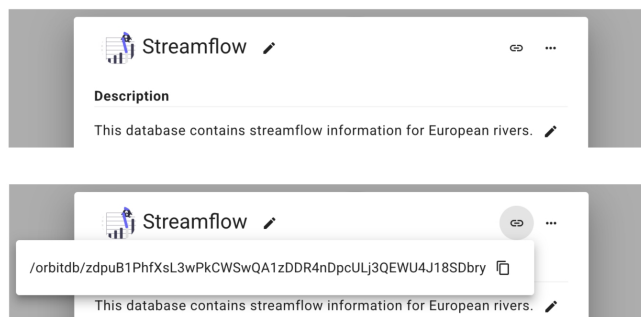


Figure 3. Example of a database information card header in the Constellation interface. The link icon, when clicked (lower), opens a pop-up with the unique database identifier that is normally kept hidden for user friendliness.

two-factor authentication and one-time passwords). In a non-centralised system, however, there is no central authority and users themselves must be able to verify each other's identity. This is generally done through a key-pair authentication system, where two mathematically linked numbers, called keys, are generated by each user. One of the two numbers, the public key, is shared with other users and is used as an account identifier, while the private key never leaves the user's device and is used to cryptographically sign messages, thereby providing a proof that the message indeed originates from the user in question.

A consequence of this system is that there is no longer a need for user registration and login; the simple act of opening the app on a device generates the public-private key pair that uniquely identifies the user. Our first version of the Constellation app therefore opened directly on the user's dashboard, ready to search and create datasets. Initial user feedback, however, indicated that this was confusing to many users habituated to having to first create an account by choosing a username before logging in to an online service. As such, we reworked the app to add a user registration login process, where, upon first login, the user is guided through the choice of their display name and profile photo before accessing the app.

Similarly, as user credentials and data are stored on the local device, there is no risk of loss of user data or account access from forgotten passwords as in centralised services; at the same time, the physical loss of a device or wiping of its memory, unlike in centralised systems, can lead to irrevocable loss of user accounts. To mitigate this risk, users should sign in to the distributed platform on more than one electronic device if possible or, alternatively, export their account credentials and potentially data to a backup drive. As this represents a change in user behaviour as compared to centralised systems, distributed platforms should take specific care to encourage users to link devices to their account. In Constellation, this is achieved by sending periodic notifications to users who have not yet linked a second device to

their account, especially when on the browser version of the app.

2.1.5 Concentric networking

A second consequence of the abovementioned distributed-system structure is that, just as user clients must verify other users' identities, individual users must also curate their network on the platform. The impacts of this are seen both in the context of data indexing and searching and in the case of content moderation (an important concern, particularly in distributed systems, as historical precedence shows that any reasonably popular online platform will sooner or later attract less well intentioned actors).

Regarding data indexing and searching, two main approaches would be possible: (1) a distributed hash table (DHT) created to create a shared data index between peers (the approach used by the HoloChain network (Parmar et al., 2023) for user data as well as by libp2p (Baumgart and Mies, 2007) for peer lookup) or (2) concentric searching, where a query is first sent to the requester's immediate network of contacts and collaborators and then sent on recursively to the contacts and collaborators of these contacts until sufficient data corresponding to the request are found (see Malard-Adam, 2025, for more details). Constellation adopts the latter approach, but the details of implementation are hidden to the user – searches merely return a list of ranked results, with the search expanding further through a user's network as the user scrolls down to request more results. At the same time, the user page will allow users to visualise their user network in graph form as well as add or remove contacts, allowing for more explicit visualisation of the peer-to-peer nature of the Constellation network; user profiles are also shown with an icon indicating the level of trust along a gradient (ranging from 1 for explicitly added contacts to 0 for perfect strangers, with co-authorship and interactions with another user's data incrementally increasing trust).

2.2 Project-specific apps

The abovementioned considerations have been mostly made in the context of interfaces for academics managing citizen science projects. For most citizen science projects, however, two levels of user interfaces will generally need to be developed – one for the managers of the project and one for end users (data collectors). The former will be responsible for defining the variables of interest to collect, managing user invitations, and running data analyses and can be expected to have a digital literacy level that includes at least the use of spreadsheet software (LibreOffice/Excel/Google Sheets), while the latter can potentially have a much lower level of digital literacy and may, in some cases, be illiterate. The academics leading the project will be responsible for defining the project and scope, including variables of interest, their units, and the data format (e.g., rainfall, pictures or videos of

waterbodies, and qualitative amount of streamflow, as well as metadata such as time, date, and location) and will generally work with tabular data structures (Excel/LibreOffice spreadsheet-style) to define project data structure and even visualise collected data. Citizen scientists, on the other hand, will simply use the user interface to input data, generally through a form, and will usually prefer to visualise data in graphics or maps instead of tabular form (Sormunen et al., 2023; Uelmen et al., 2023).

This is usually managed by having separate administrator and user sign-ins for administrators and users or else through the automatic generation, based on administrator-defined data structures, of a form for end users to enter data, as in KoboToolBox (Lakshminarasimhappa, 2022). Early in the process of developing Constellation, the decision was taken to divide the codebase into two main parts, the networking stack and the user interface (Malard-Adam, 2025). The former contains all code functions necessary to create user accounts, create and modify data, and network with other users; the latter provides a user interface tailored to academics that allows for account management and creation and search of databases on the network, mostly presenting data in tabular form.

The development of a generic interface tailored to citizen scientists, however, is much more difficult to attain, even within a single field of study. While generic actions such as user onboarding can be relatively standardised, different citizen science projects will have vastly different user needs regarding data entry and visualisation (Giuliana, 2017). For instance, a crowdsourced precipitation measurement project may decide to present interpolated precipitation overlaid on local area maps, while a river flow monitoring app may instead wish to present photos or videos of local rivers along with AI-generated flow rate estimates, and an insect tracking app may present different users' photos of insects along with AI-based species identification. This diversity in end-user needs produces a conundrum for the development of end-user interfaces for Constellation: on the one hand, the diversity of potential ecological projects makes the development of a single high-quality generic platform for end users unfeasible; on the other hand, development of application-specific user interfaces is highly time-consuming and requires non-trivial specialised programming skills that pose a high entry barrier for the vast majority of would-be citizen science project founders.

Constellation therefore chooses a middle route by allowing academics, once they have specified their project data structures in the Constellation interface, to automatically generate a working stand-alone app for data entry and visualisation by end users. This one-click app code generation will automatically configure deployable code for the specific citizen science project in question and allow users with minimal to no coding experience to rapidly deploy a minimal viable interface for data collection by their citizen scientists. At the same time, more advanced users can use this automatically

generated scaffolding to then customise the app according to their specific project needs.

The aforementioned points regarding best practices for distributed citizen science application development are particularly relevant in the case of end-user app development, especially regarding the recommendations on avoiding the complexity of peer-to-peer implementation details and the importance of making apps “feel” just like existing centralised solutions.

3 Lessons learned

We identify three main lessons learned from our experiences in developing user interfaces for distributed scientific databases for ecological sciences: firstly, that design principles are transferable across the peer-to-peer field, with insights into user behaviour from other (even non-scientific) applications largely applicable to scientific users; secondly, that simple user interfaces that users find intuitive to use are key to driving adoption of new technological advances; and thirdly, that user interfaces for peer-to-peer systems should aim to remain as simple as possible for the user and to mimic centralised systems as much as possible.

3.1 Transferability of lessons

While there is extremely limited literature on the design of user interfaces for distributed systems, we find that the conclusions of the limited studies available do provide relevant recommendations for the case of citizen science, even if these studies were conducted in different contexts. In particular, Kosem and Dietrich (2020) studied user preferences of three user groups (experts, early adopters, and potential users) in the context of a distributed mobile chat and file-sharing app and presented a list of general user interface design recommendations for distributed apps. These recommendations fall under the general philosophy that distributed app interfaces should aim to mimic, as much as possible, the user experience of centralised apps that users have become used to and only show peer-to-peer-specific features where these are unavoidable for app functionality (e.g., local pinning of data, identity recovery seed phrases) or else offer functionality unavailable in centralised apps (e.g., connecting to peers without internet access). Of the specific recommendations, many were also adopted in Constellation, including the discreet display of data hash links alongside more prominent, user-friendly names and links, the use of subtle icons (pin) to show whether a specific dataset is safely stored locally, and the avoidance of long text descriptions of how peer-to-peer systems work in favour of simply showing the user how the app functions through icons and workflows.

3.2 User interfaces as drivers of technological adoption

Well-designed user interfaces have previously been identified as crucial drivers of adoption of citizen platforms by citizen scientists (Giuliana, 2017), and previous experience strongly indicates that the adoption of cutting-edge digital technologies in the field of science is highly dependent on ease of use and quality of the interface.

We draw here on experience from a different field, that of the adoption of Monte Carlo Markov chain Bayesian inference in statistical methods, which shares important characteristics with the currently new field of distributed scientific data, namely theoretical potential for improved performance or results combined with steep learning curves and important practical difficulties of implementation. The development of the OpenBUGS software opened up the field of Monte Carlo Markov chain Bayesian inference, a notoriously complex algorithm to programme, to general researchers. While arguably more powerful than classical statistics and offering more easily interpretable analyses, Bayesian statistics (1) require a shift in frame of mind to change from classical statistics and (2) are non-trivial to implement and require specialist knowledge to code, with many potential pitfalls and a rapidly evolving set of current “best practices,” two characteristics shared by distributed scientific databases. While OpenBUGS made Bayesian inference accessible to a wider scientific audience, the user interface was, however, notoriously difficult to navigate, and the software was eventually superseded by the much more ergonomic PyMC (Salvatier et al., 2016) and RStan (Stan Development Team, 2025) software.

We argue that the use of peer-to-peer distributed databases in citizen science initiatives today is in much the same situation as Bayesian inference was before the invention of OpenBUGS. Both are at a point in development where the required technology has been developed as scholarly proofs of concept but remains technologically challenging to implement correctly and as such has not been adopted by the vast majority of use cases for which it could be appropriate.

However, and contrarily to the state of Bayesian inference at that time, there is large availability of alternative software for real-time data sharing (Google Sheets, KoboToolBox, and others), all with very user-friendly interfaces. Although these are centralised solutions, many of them offer offline capacities and use paid tiers to offer free levels of data storage for their users. Alternatives must therefore be similarly intuitive to use if they are to compete with established platforms for users’ time and effort (Lemmens et al., 2021; Liu et al., 2021). While poor data management practices have been identified in general as significant barriers to open data reuse in the sciences (Lansana et al., 2020) and building citizen science platforms on open distributed database systems brings long-term financial sustainability advantages for larger projects as well as better data access for the scientific community in general, the vast majority of potential users will be concerned less immediately by these issues and more

by the ease of onboarding and the possibility of rapidly setting up a functioning project.

3.3 The importance of simplicity

Previous research has highlighted the importance of simplicity in citizen science user interface design to improve data collection (Liebenberg et al., 2017), particularly when working with less literate user groups. Citizen science platforms based on distributed databases should attempt, as much as possible, to hide the more technical aspects of peer-to-peer communication behind interfaces that mimic the user experience of centralised systems with which users are well familiar while only mentioning aspects specific to the peer-to-peer architecture when these are absolutely necessary (e.g., showing whether data are pinned and thus safely stored on the local device) or else offer an advantage over centralised systems (e.g., showing the number of peers pinning a user’s data). The concerns of ecological practitioners involved in citizen science projects principally regard the usefulness of the project for data collection and education of citizen scientists (Minkman et al., 2017), and convincing academics leading such projects to switch to a distributed database software system will require user interfaces to leverage the potential of distributed systems to further these goals while ideally decreasing system maintenance costs. In the end, users care less about the distributed or centralised nature of the system and more about its ease of use and the concrete advantages that distributed systems can bring, such as lower costs, avoiding server setup, and local-first data access (Kosem and Dietrich, 2020). Distributed database software solutions for scientific data, as such, need to meet a high standard of user interface simplicity and ease of use if they are to attract potential users away from alternative centralised database systems, whether commercial or open-source.

4 Perspectives

As uptake of distributed databases increases and projects built on their infrastructure increase in size, questions of scalability will likely come to the forefront of scientific research in this field in the near future. Software solutions will need to address scalability in terms of both data volume (storage on limited end-user devices) and peer connectivity (scaling peer and data discovery across a large number of peers, where connecting directly to each one is infeasible). Centralised database systems have also evolved towards data *sharding* (i.e., sharing the burden of hosting data) between multiple servers for reasons of data volume and geographical proximity (e.g., Cassandra (Carpenter and Hewitt, 2022)); distributed databases solve the same problem but with the added complexity of running not on servers but on end-user devices with limited processing power and storage space and uncertain connectivity.

Constellation itself is built on OrbitDB (Faria and Pereira, 2025), a low-level database that uses the InterPlanetary File System (IPFS) (Trautwein et al., 2022) for data storage and transmission and uses libp2p (What is libp2p, 2024) for peer-to-peer connectivity. The software currently mitigates the data volume footprint on end-user devices by saving media files to IPFS, a peer-to-peer protocol that allows users to retrieve content from other peers in the network. As only the file's address is saved to the OrbitDB database, individual databases remain relatively small, even if voluminous data such as pictures or video are included. This allows users to decide whether they wish to “pin” (locally copy and replicate) the associated media files to their local device or not, which will keep most databases to a manageable size. OrbitDB also uses the GossipSub protocol (Vyzovitis et al., 2020) to efficiently propagate database updates throughout a network of peers which are not all directly connected to each other, which allows the system to scale to a large number of users.

In addition, since citizen science projects are created not as a single database containing every user's data but rather as a swarm of identically formatted databases belonging to each user, users can search incrementally and concentrically through their network for data contributors and potentially stop searching once a certain number of data points are found. Since user contributions are contained in individual databases, malicious users and data can also be easily blocked without polluting the entire project's history, as would be the case with a single central database. This also sidesteps the problem of data consistency between different users, since each user is only effectuating changes to their own individual database. In the case where multiple users (or multiple devices from a single user) do collaborate on the same database, OrbitDB offers *eventual consistency*, which means that, although the state of a database on different devices may differ at any given moment, they will eventually converge to the same state if connected to each other for long enough.

An alternative approach has been adopted by PeerBit (PeerBit development team, 2025), a distributed database system originally based on OrbitDB. In this case, a single database is created to hold all users' data; this database is then sharded amongst users so that each user only holds a local copy of part of the data. This allows for easier automated load sharing between users; on the other hand, the use of a single database with an append-only log (i.e., previous entries cannot be erased, only hidden from display) makes the system more vulnerable to chain pollution attacks (e.g., adding voluminous, objectionable, or illegal content) from malicious or compromised peers (Matzutt et al., 2018).

As the use of distributed databases in citizen science research increases, we expect many opportunities to arise for future research aiming to quantify and optimise network latency and data availability, as well as in user interface design

improvements to effectively communicate connectivity and data availability information to users.

5 Limitations

The results presented in this paper were obtained with a limited number of participants (between one and three dozen) connected to the app at the same time. Larger-scale use of the app, with more users concurrently connected, will likely lead to further refinement of its interface design and associated recommendations for best practices, especially in communicating network structure and connectivity status. In the future, formal user type analysis, such as the Hexad user type scales (Krath et al., 2023; Tondello et al., 2016), could offer more quantitative and actionable directives for user interface design of distributed peer-to-peer database apps for different target user groups.

6 Conclusions

Citizen science is an important tool for ecological research, contributing to both data collection and improved community participation and leadership in research initiatives. All citizen science data collection and storage systems to date, however, are based on centralised database systems that come with important infrastructure costs and data sovereignty issues for contributing communities. Distributed peer-to-peer data systems have recently been developed and proposed as a solution to these issues but involve several key conceptual differences from the centralised systems to which internet and mobile application users have become accustomed to, which hinders the potential for more general adoption in ecological citizen science projects. This paper discusses the potential for user interface design of ecological citizen science apps to lower the learning curve for the use of distributed databases in this field and to encourage adoption of this new technology in the ecological community. In a context where relatively little literature on the design of user interfaces for citizen science applications (and none regarding distributed database systems) in ecology is available, we draw on lessons learned from similar fields as well as on formal and informal feedback from early-adopter users in workshops to provide general recommendations for user interface design for citizen science projects using peer-to-peer distributed databases.

Code and data availability. No code or data were directly used to produce the results presented in this paper. Code for the Constellation interface can be found at <https://doi.org/10.5281/zenodo.17252582> (Malard-Adam and Sheeja, 2025), while the code for the entire Constellation suite of software is available under the following GitHub organisation: <https://github.com/reseau-constellation> (last access: 1 October 2025). All of these codes are available under open-source licences.

Original author names.

Nallusamy Anandaraja நல்லுசாமி ஆனந்தராஜா

Sheeja சூ ஷீஜா

Palanivelan Jaisridhar பழனிவேலன் ஜெய்ஸ்ரீதர்

Original affiliation names.

3 விரிவாக்கக் கல்வி இயக்ககம்,

தமிழ்நாடு வேளாண்மைப் பல்கலைக்கழகம்

6 भारतीय कृषि अनुसंधान परिषद – कृषि विज्ञान केंद्र

தமிழ்நாடு வேளாண்மைப் பல்கலைக்கழகம்

Author contributions. JJMA – conceptualisation, methodology, software development, writing. WM – conceptualisation, methodology, writing. NA – conceptualisation, writing. JH – conceptualisation, writing. JD – conceptualisation, methodology, writing. KS – software development. PJ – conceptualisation, writing.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. The authors would like to express their gratitude to Constellation workshop participants and early adopters for their valuable feedback, criticism, and recommendations which have informed this article, as well as for their kind patience with our alpha-stage software.

Review statement. This paper was edited by Sara Gamboa and reviewed by Francisco Sanz García and Sara Gamboa.

References

Árvai, M., Takács, T., Kovács, Z. A., Takács, K., Balog, K., László, P., Imréne Takács, T., Mészáros, J., and Pásztor, L.: Az ‘Alsóban az élet’ című hazai talajállapotot célzó közösségi tudomány program első tapasztalatai és eredményei – The First experiences and results of the Hungarian citizen science program (“Life in Undies”) aimed at soil properties, *Agrokem. Talajtan*, 72, 25–43, <https://doi.org/10.1556/0088.2022.00136>, 2023.

Avellaneda, P. M., Ficklin, D. L., Lowry, C. S., Knouft, J. H., and Hall, D. M.: Improving Hydrological Models With the Assimilation of Crowdsourced Data, *Water Resour. Res.*, 56, <https://doi.org/10.1029/2019WR026325>, 2020.

Baumgart, I. and Mies, S.: “S/Kademlia: A practicable approach towards secure key-based routing”, 2007 International Conference on Parallel and Distributed Systems, Hsinchu, Taiwan, 1–8 pp., <https://doi.org/10.1109/ICPADS.2007.4447808>, 2007.

Bonnet, P., Joly, A., and Munoz, F.: Complementarity of Big Data and Citizen Participation in Monitoring Plant Biodiversity, in: *Linking with Nature in the Digital Age*, 1, 151–159, <https://doi.org/10.1002/9781394297580.ch8>, 2024.

Carpenter, J. and Hewitt, E.: CASSANDRA: The Definitive Guide, (revised) Third Edition, O’Reilly Media, Inc., 433 pp., ISBN 9781098115159, 2022.

Cieslik, K., Shakya, P., Upreti, M., Dewulf, A., Russell, C., Clark, J., Dhital, M. R., and Dhakal, A.: Building Resilience to Chronic Landslide Hazard Through Citizen Science, *Front. Earth Sci.*, 7, 278, <https://doi.org/10.3389/feart.2019.00278>, 2019.

Constellation: Constellation, <https://docu.reseau-constellation.ca/>, last access: 8 July 2025.

Faria, N. and Pereira, J.: CRDV: Conflict-free replicated data views, *Proceedings of the ACM on Management of Data*, Vol. 3, 25, 1–27 pp., <https://doi.org/10.1145/3709675>, 2025.

Fraisl, D., Hager, G., Bedessem, B., Gold, M., Hsing, P.-Y., Danielsen, F., Hitchcock, C. B., Hulbert, J. M., Piera, J., Spiers, H., Thiel, M., and Haklay, M.: Citizen science in environmental and ecological sciences, *Nat. Rev. Methods Primers*, 2, 1–20, <https://doi.org/10.1038/s43586-022-00144-4>, 2022.

Freitas, H. and Gouveia, A. C.: Biodiversity futures: digital approaches to knowledge and conservation of biological diversity, *Web Ecol.*, 25, 29–37, <https://doi.org/10.5194/we-25-29-2025>, 2025.

Giuliana, D.: Designing an interface for citizen science platforms ensuring a good user experience, Munich: Ludwig-Maximilians-Universität München, Institut für Informatik, <https://api.semanticscholar.org/CorpusID:32496628> (last access: 1 October 2025), 2017.

Gomes, V. B. F., Kleppmann, M., Mulligan, D. P., and Beresford, A. R.: Verifying strong eventual consistency in distributed systems, *Proc. ACM Program. Lang.*, 1, 109:1–109:28, <https://doi.org/10.1145/3133933>, 2017.

Haklay, M., Dörler, D., Heigl, F., Manzoni, M., Hecker, S., and Vohland, K.: What Is Citizen Science? The Challenges of Definition, in: *The Science of Citizen Science*, edited by: Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R., and Wagenknecht, K., Springer International Publishing, Cham, 13–33, https://doi.org/10.1007/978-3-030-58278-4_2, 2021.

Hall, D. M., Avellaneda-Lopez, P. M., Ficklin, D. L., Knouft, J. H., and Lowry, C.: How to close the loop with citizen scientists to advance meaningful science, *Sustain. Sci.*, 19, 1527–1542, <https://doi.org/10.1007/s11625-024-01532-3>, 2024.

Hicks, A., Barclay, J., Chilvers, J., Armijos, M. T., Oven, K., Simmons, P., and Haklay, M.: Global Mapping of Citizen Science Projects for Disaster Risk Reduction, *Front. Earth Sci.*, 7, 226, <https://doi.org/10.3389/feart.2019.00226>, 2019.

Howard, L., Van Rees, C. B., Dahlquist, Z., Luikart, G., and Hand, B. K.: A review of invasive species reporting apps for citizen

- science and opportunities for innovation, *NeoBiota*, 71, 165–188, <https://doi.org/10.3897/neobiota.71.79597>, 2022.
- Joly, A., Bonnet, P., Goëau, H., Barbe, J., Selmi, S., Champ, J., Dufour-Kowalski, S., Affouard, A., Carré, J., Molino, J.-F., Boujemaa, N., and Barthélémy, D.: A look inside the PI@ntNet experience, *Multimedia Systems*, 22, 751–766, <https://doi.org/10.1007/s00530-015-0462-9>, 2016.
- Khu, S. T., Wang, J., and Wang, M.: Application Status and Future of Citizen Science in Hydrology, *Advanced Engineering Sciences*, 55, 141–148, <https://doi.org/10.1596/1/j.jsuese.202200369>, 2023.
- Kosem, J. and Dietrich, A.: IPFS Mobile Guidelines, <https://blog.ipfs.tech/2020-06-25-ipfs-mobile-design-guidelines/> (last access: 1 October 2025), 2020.
- Krath, J., Altmeyer, M., Tondello, G. F., and Nacke, L. E.: Hexad-12: Developing and Validating a Short Version of the Gamification User Types Hexad Scale, in: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 1–18, <https://doi.org/10.1145/3544548.3580968>, 2023.
- Lakshminarasimhappa, M. C.: Web-Based and Smart Mobile App for Data Collection: Kobo Toolbox / Kobo Collect, *Journal of Indian Library Association*, 57, 72–79, 2022.
- Lansana, A. S., Migisha, C., Minjire, D., Juma, L., Adan, S., and Alemu, W.: Drivers of Data for Development, <https://www.developlocal.org/wp-content/uploads/2020/07/D4D-report-2020.pdf> (last access: 1 October 2025), 2020.
- Lee, K. A., Lee, J. R., and Bell, P.: A review of Citizen Science within the Earth Sciences: potential benefits and obstacles, *P. Geologist. Assoc.*, 131, 605–617, <https://doi.org/10.1016/j.pgeola.2020.07.010>, 2020.
- Lefort, T., Affouard, A., Charlier, B., Lombardo, J.-C., Chouet, M., Goëau, H., Salmon, J., Bonnet, P., and Joly, A.: Cooperative learning of PI@ntNet's Artificial Intelligence algorithm: How does it work and how can we improve it?, *Methods Ecol. Evol.*, <https://doi.org/10.1111/2041-210X.14486>, 2025.
- Lemmens, R., Antoniou, V., Hummer, P., and Potsiou, C.: Citizen Science in the Digital World of Apps, in: *The Science of Citizen Science*, edited by: Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R., and Wagenknecht, K., Springer International Publishing, Cham, 461–474, https://doi.org/10.1007/978-3-030-58278-4_23, 2021.
- Liebenberg, L., Steventon, J., Brahman, N., Benadie, K., Minye, J., Langwane, H. K., and Xhukwe, Q. U.: Smartphone Icon User Interface design for non-literate trackers and its implications for an inclusive citizen science, *Biol. Conserv.*, 208, 155–162, <https://doi.org/10.1016/j.biocon.2016.04.033>, 2017.
- Liu, H.-Y., Dörler, D., Heigl, F., and Grossberndt, S.: Citizen Science Platforms, in: *The Science of Citizen Science*, edited by: Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R., and Wagenknecht, K., Springer International Publishing, Cham, 439–459, https://doi.org/10.1007/978-3-030-58278-4_22, 2021.
- Lowry, C. S., Fienen, M. N., Hall, D. M., and Stepenuck, K. F.: Growing Pains of Crowdsourced Stream Stage Monitoring Using Mobile Phones: The Development of CrowdHydrology, *Front. Earth Sci.*, 7, 128, <https://doi.org/10.3389/feart.2019.00128>, 2019.
- Lumbrazo, C., Bennett, A., Currier, W. R., Nijssen, B., and Lundquist, J.: Evaluating Multiple Canopy-Snow Unloading Parameterizations in SUMMA With Time-Lapse Photography Characterized by Citizen Scientists, *Water Resour. Res.*, 58, <https://doi.org/10.1029/2021WR030852>, 2022.
- Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N.: The BUGS project: Evolution, critique and future directions, *Stat. Med.*, 28, 3049–3067, <https://doi.org/10.1002/sim.3680>, 2009.
- Malard-Adam, J.: Données scientifiques distribuées et science citoyenne – Le réseau Constellation, in review, 2025.
- Malard-Adam J. and Sheeja, K.: IUG Constellation, v1.0.3, Zenodo [code], <https://doi.org/10.5281/zenodo.17252582>, 2025.
- Malard-Adam, J., (Kumar), Medema, W., (Anandaraja), Harms, J., and Dipple, J.: Distributed databases to improve data sovereignty in citizen science, *EGU General Assembly 2024*, Vienna, Austria, 14–19 Apr 2024, EGU24-12726, <https://doi.org/10.5194/egusphere-egu24-12726>, 2024.
- Matzutt, R., Hiller, J., Henze, M., Ziegeldorf, J. H., Müllmann, D., Hohlfeld, O., and Wehrle, K.: A Quantitative Analysis of the Impact of Arbitrary Blockchain Content on Bitcoin, in: *Financial Cryptography and Data Security*, Berlin, Heidelberg, 420–438, https://doi.org/10.1007/978-3-662-58387-6_23, 2018.
- Mazzoleni, M., Verlaan, M., Alfonso, L., Monego, M., Norbiato, D., Ferri, M., and Solomatine, D. P.: Can assimilation of crowdsourced data in hydrological modelling improve flood prediction?, *Hydrol. Earth Syst. Sci.*, 21, 839–861, <https://doi.org/10.5194/hess-21-839-2017>, 2017.
- Medema, W., Dipple, J., and Malard-Adam, J.: The role of communities in integrated water resource management, in: *Handbook on the Governance and Politics of Water Resources*, edited by: Benson, D. and Fristch, O., Edward Elgar, 83–101, ISBN 9781800887893, ISBN 9781800887909, <https://doi.org/10.4337/9781800887909.00015>, 2024.
- Mengistie, G. K., Haile, A. T., O'Donnell, G., Negash, E. D., Bekele, T. W., and Tedla, H. Z.: Citizen science data to improve rainfall-runoff model performance in urbanizing Akaki catchment, Awash Basin, Ethiopia, *J. Hydrol.*, 53, <https://doi.org/10.1016/j.ejrh.2024.101822>, 2024.
- Minkman, E., van der Sanden, M., and Rutten, M.: Practitioners' viewpoints on citizen science in water management: a case study in Dutch regional water resource management, *Hydrol. Earth Syst. Sci.*, 21, 153–167, <https://doi.org/10.5194/hess-21-153-2017>, 2017.
- Moshi, H. A., Shilla, D. A., Kimirei, I. A., O'Reilly, C., Clymans, W., Bishop, I., and Loisel, S. A.: Community monitoring of coliform pollution in Lake Tanganyika, *PLoS ONE*, 17, <https://doi.org/10.1371/journal.pone.0262881>, 2022.
- Parmar, J. K. and Vaghani, K. G.: A conceptual study on Holochain and blockchain technology, in: *Artificial Intelligence and Communication Technologies*, edited by: Hiranwal, S. and Mathur, G., 331–341. *Computing & Intelligent Systems, SCRS, India*, <https://doi.org/10.52458/978-81-955020-5-9-33>, 2023.
- Parsons, M. A., Duerr, R. E., and Jones, M. B.: The History and Future of Data Citation in Practice, *Data Science Journal*, 18, 1–10, <https://doi.org/10.5334/dsj-2019-052>, 2019.
- PeerBit development team: dao-xyz/peerbit, GitHub [code], <https://github.com/dao-xyz/peerbit/> (last access: 1 October 2025), 2025.
- Pernat, N., Memedemin, D., August, T., Preda, C., Reyserhove, L., Schirmel, J., and Groom, Q.: Extracting secondary

- data from citizen science images reveals host flower preferences of the Mexican grass-carrying wasp *Isodontia mexicana* in its native and introduced ranges, *Ecol. Evol.*, 14, <https://doi.org/10.1002/ece3.11537>, 2024.
- Perry, M. L.: The Art of Immutable Architecture: Theory and Practice of Data Management in Distributed Systems, <https://doi.org/10.1007/978-1-4842-5955-9>, 2020.
- Salvatier, J., Wiecki, T. V., and Fonnesbeck, C.: Probabilistic programming in Python using PyMC3, *PeerJ Comput. Sci.*, 2, e55, <https://doi.org/10.7717/peerj-cs.55>, 2016.
- See, L.: A Review of Citizen Science and Crowdsourcing in Applications of Pluvial Flooding, *Front. Earth Sci.*, 7, 44, <https://doi.org/10.3389/feart.2019.00044>, 2019.
- Seibert, J., Strobl, B., Etter, S., Vis, M., and van Meerveld, H. J.: CrowdWater: a new smartphone app for crowd-based data collection in hydrology, 11647, 19th EGU General Assembly, EGU2017, proceedings from the conference held 23–28 April, 2017 in Vienna, Austria, p. 11647, <https://ui.adsabs.harvard.edu/abs/2017EGUGA..1911647S> (last access: 1 October 2025), 2017.
- Songchon, C., Wright, G., and Beevers, L.: The use of crowdsourced social media data to improve flood forecasting, *J. Hydrol.*, 622, 129703, <https://doi.org/10.1016/j.jhydrol.2023.129703>, 2023.
- Sormunen, J. J., Kulha, N., Alale, T. Y., Klemola, T., Sääksjärvi, I. E., and Vesterinen, E. J.: For the people by the people: Citizen science web interface for real-time monitoring of tick risk areas in Finland, *Ecological Solutions and Evidence*, 4, <https://doi.org/10.1002/2688-8319.12294>, 2023.
- Sprinks, J., Houghton, R., Bamford, S., Morley, J., and Wardlaw, J.: Is that a crater? Designing citizen science platforms for the volunteer and to improve results, in: European Planetary Science Congress, 2015, EPSC Abstracts, 10, EPSC2015-694, 2015.
- Stan Development Team: rstan: R Interface to Stan, <https://mc-stan.org/rstan> (last access: 1 October 2025), 2025.
- Strobl, B., Etter, S., Meerveld, I. van, and Seibert, J.: The Crowd-Water game: A playful way to improve the accuracy of crowd-sourced water level class data, *PLOS ONE*, 14, e0222579, <https://doi.org/10.1371/journal.pone.0222579>, 2019.
- Teacher, A. G. F., Griffiths, D. J., Hodgson, D. J., and Inger, R.: Smartphones in ecology and evolution: a guide for the app-rehensive, *Ecol. Evol.*, 3, 5268–5278, <https://doi.org/10.1002/ece3.888>, 2013.
- Tondello, G. F., Wehbe, R. R., Diamond, L., Busch, M., Marczewski, A., and Nacke, L. E.: The Gamification User Types Hexad Scale, in: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play, New York, NY, USA, 229–243, <https://doi.org/10.1145/2967934.2968082>, 2016.
- Torre, M., Nakayama, S., Tolbert, T. J., and Porfiri, M.: Producing knowledge by admitting ignorance: Enhancing data quality through an “I don’t know” option in citizen science, *PLoS ONE*, 14, e0211907, <https://doi.org/10.1371/journal.pone.0211907>, 2019.
- Trautwein, D., Raman, A., Tyson, G., Castro, I., Scott, W., Schubotz, M., Gipp, B., and Psaras, Y.: Design and evaluation of IPFS: a storage layer for the decentralized web, in: Proceedings of the ACM SIGCOMM 2022 Conference, New York, NY, USA, 739–752, <https://doi.org/10.1145/3544216.3544232>, 2022.
- Uelmen, J. A., Jr., Clark, A., Palmer, J., Kohler, J., Van Dyke, L. C., Low, R., Mapes, C. D., and Carney, R. M.: Global mosquito observations dashboard (GMOD): creating a user-friendly web interface fueled by citizen science to monitor invasive and vector mosquitoes, *Int. J. Health Geogr.*, 22, <https://doi.org/10.1186/s12942-023-00350-7>, 2023.
- von Gönner, J., Gröning, J., Grescho, V., Neuer, L., Gottfried, B., Hänsch, V. G., Molsberger-Lange, E., Wilharm, E., Liess, M., and Bonn, A.: Citizen science shows that small agricultural streams in Germany are in a poor ecological status, *Sci. Total Environ.*, 922, <https://doi.org/10.1016/j.scitotenv.2024.171183>, 2024.
- Vyzovitis, D., Napora, Y., McCormick, D., Dias, D., and Psaras, Y.: GossipSub: Attack-Resilient Message Propagation in the Filecoin and ETH2.0 Networks, *arXiv [preprint]*, <https://doi.org/10.48550/arXiv.2007.02754>, 2020.
- Weaver, N.: The Web3 Fraud, *USENIX*, 2021.
- What is libp2p: What is libp2p, <https://docs.libp2p.io/concepts/introduction/overview/>, last access: 1 October 2025.
- Wood, C. M., Kahl, S., Rahaman, A., and Klinck, H.: The machine learning-powered BirdNET App reduces barriers to global bird research by enabling citizen science participation, *PLOS Biol.*, 20, e3001670, <https://doi.org/10.1371/journal.pbio.3001670>, 2022.