

Who is saving our streamflow data? Exploring volunteer profiles and their engagement in the SIREN data rescue project

Original

Who is saving our streamflow data? Exploring volunteer profiles and their engagement in the SIREN data rescue project / Mazzoglio, Paola; Bertola, Miriam; Listo, Tommaso; Princivalle, Luca; Lombardo, Luca; Viglione, Alberto; Claps, Pierluigi; Mattozzi, Alvisè. - In: PLOS ONE. - ISSN 1932-6203. - ELETTRONICO. - 20:(2025), pp. 1-19. [10.1371/journal.pone.0333091]

Availability:

This version is available at: 11583/3003855 since: 2025-10-10T13:33:43Z

Publisher:

PLOS

Published

DOI:10.1371/journal.pone.0333091

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

RESEARCH ARTICLE

Who is saving our streamflow data? Exploring volunteer profiles and their engagement in the SIREN data rescue project

Paola Mazzoglio^{1*}, Miriam Bertola², Tommaso Listo¹, Luca Princivalle¹, Luca Lombardo¹, Alberto Viglione¹, Pierluigi Claps¹, Alvise Mattozzi¹

1 Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, Torino, Italy, **2** Institute of Hydraulic Engineering and Water Resources Management, Vienna University of Technology, Wien, Austria

* paola.mazzoglio@polito.it



Abstract

SIREN is a citizen science project that involves lay people in the digitization of historical daily discharge measurements from Italian rivers. Such data, largely available only in printed yearbooks, hinders scientific progress in hydrological studies and water resource management. In this article, we examine the motivations behind citizen engagement in SIREN. Our multi-step approach combines quantitative analysis of online contributions, pilot interviews with selected volunteers, and a comprehensive questionnaire collecting basic demographic data and subjective impressions of the experience. Through these approaches, we identify three participant profiles: two driven primarily by the activity itself and one by the scientific content. The first profile values the straightforward nature of data entry, seeing it as an easy way to contribute with existing skills. The second profile treats participation as a leisure activity, readily fitting into brief intervals of free time. The third profile stems from deeper engagement, encompassing volunteers with professional or personal interests in hydrology, Italian geography, or both. The study also highlights the significant role of retired individuals (an underrepresented group in the citizen science literature) who often contribute using skills developed during their careers. This work highlights the importance of creating citizen science projects that are accessible, meaningful, and connected to volunteers' lives and interests.

OPEN ACCESS

Citation: Mazzoglio P, Bertola M, Listo T, Princivalle L, Lombardo L, Viglione A, et al. (2025) Who is saving our streamflow data? Exploring volunteer profiles and their engagement in the SIREN data rescue project. *PLoS One* 20(10): e0333091. <https://doi.org/10.1371/journal.pone.0333091>

Editor: Naveen Joseph, Radford University Artistic College of Science and Technology, UNITED STATES OF AMERICA

Received: April 11, 2025

Accepted: September 9, 2025

Published: October 9, 2025

Copyright: © 2025 Mazzoglio et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All relevant data are within the paper and its [Supporting Information](#) files. "S1_Survey.pdf" contains the questions of the survey while "S2_Survey_69answers.xlsx" contains the answers.

1. Introduction

In Italy, since the early 1900s, the National Hydrological and Mareographic Service (SIMN) has been in charge of the management of the Italian monitoring network, whose records were made available in a series of printed hydrological yearbooks. Around 30 years ago, the SIMN was dismantled, and the data collection task was transferred to the regional level [1]. While data collected in recent years is usually

Due to privacy concerns, the video recordings of the pilot interviews cannot be shared, as the participants are clearly identifiable and did not provide consent for public distribution of the recordings.

Funding: IMPETUS is supporting our project. IMPETUS is funded by the European Union's Horizon Europe research and innovation programme under grant agreement number 101058677. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Competing interests: The authors have declared that no competing interests exist.

available in digital format, historical measurements are often only available in the printed volumes and scanned images of the hydrological yearbooks published by the SIMN. In the past decades, limited efforts have been dedicated to digitize the historical hydrological measurements, even though they would be extremely valuable for supporting water resource management, e.g., by offering insights into past and present hydrological conditions, for the detection of temporal trends in river flows, and for the development and calibration of hydrological models such as the ones used in flood monitoring and forecasting. A few independent digitization initiatives have attempted to recover this information in the past [1–5] making it available in machine-readable formats. Unfortunately, their focus has been limited to specific variables, specific regions and/or targeted years.

The SIREN (Saving Italian hydrological measurements) project was thus launched as a concrete attempt to digitize the historical daily discharge data measured along the Italian rivers by leveraging citizen science to crowdsource the digitization of historical daily discharge data. In order to fully understand the potential and limits of such an approach, it is important to examine how citizen science has evolved over time and how it has been applied in hydrological contexts.

The Oxford English Dictionary (O.E.D.) defines citizen science (CS) «*as scientific work undertaken by members of the general public, often in collaboration with or under the direction of professional scientists and scientific institutions*» [6], with the first use of the term (at least with this O.E.D.'s meaning) dating back to 1989 [7]. Such a definition presupposes science as a profession and assumes that professional scientists practice science while the “general public” (or lay people) contributes by assisting with tasks in scientific research. The analysis performed by Kullenberg and Kasperowski [8] records an oscillation in the meaning of the term (that goes from that of science at the service of citizens' needs and concerns to that of scientific activity participated by a potential and vast audience of non-experts, as well as initiatives more activism-oriented) and a popularity increase in the 2010s, in coincidence with projects using web platforms to reach a vast number of contributors. In this work, we stick to today's prevailing meaning of participatory activity.

Over time, diverse classifications of citizen science projects have been proposed, focusing on the kind of participation allowed or on the type of job required. According to Bonney et al. [9], public participation in scientific research can be divided into three main categories that measure an ascending involvement level: contributory projects (usually designed by scientists, for which citizen scientists primarily contribute data); collaborative projects (designed by scientists for which the public contributes with data but also helps in the refinement of the project design, in the analysis of the data, or in the dissemination phase); co-created projects (co-designed by scientists and citizens, with the latter actively involved in the scientific process). Strasser et al. [10], instead, distinguish qualitatively between five different epistemic practices: sensing (to monitor and collect data using sensors), computing (volunteers who share processing powers from their personal computers to make particularly expensive calculations more efficient), analyzing (data analysis work like image classification), self-reporting (sharing of personal data which will then be used in research, especially in

the medical field), making (through shared spaces and tools from crowdsourcing/crowdfunding). Haklay [11] proposes an intermediate perspective between the two previous classifications (Fig 1).

In addition to offering a concrete contribution to projects that in terms of time and resources would otherwise be difficult to accomplish [12], CS's practices are opportunities to disseminate scientific knowledge and awareness to the public. This awareness might not consist just in a public better informed about scientific knowledge, but in performing guidelines for Responsible Research [13], or in questioning clear and fixed separations between the epistemic authority of experts and that of lay people [14], no longer seen as a mere consumer of the scientific activity carried out by professional scientists [10]. This broad vision of CS includes discourses about social reorganization around scientific activity ([15,16]), framings of CS as an alternative to dominant knowledge [17] or as an attempt to bridge the wide gap between democratic deliberations and technical knowledge [18], although the literature seems to overlook or insufficiently explore the most promising forms of participation for this scope, those in which both the interests of participants and the attributions of epistemic authority are redefined throughout the participatory process [19]. Whether or not such an impact could be attributed to CS practices, there is little doubt that work is being done on their institutionalization [13,20,21] and organization, especially through web-portals (see, e.g., [22]).

Within the hydrological field, where examples of structured activities exist (see, e.g., [23]), the typical role of citizen scientists is the collection, the analysis and the interpretation of data on water resources, on weather conditions, or on some component of the water cycle, mainly in terms of quantity and/or quality. These activities include the monitoring of precipitation phase [24], rainfall depths [25], water levels [26,27], water temperature [28], water clarity and color [29], water pH [28] or dissolved hydrogeochemical parameters [28] using simple and low-cost tools such as meters, test kits, smartphones and computers, or collecting water samples for laboratory analyses on nutrient levels, sediment loads, and the presence of contaminants or pollution. In the last decade, several CS initiatives aiming at recovering hydrological and meteorological observations from very old printed and handwritten documents have been launched. Table 1 represents an overview of such initiatives. Most of these projects have been implemented on the Zooniverse platform and involve the transcription of records by the volunteers.

These diverse applications highlight not only the scientific potential of citizen involvement but also raise important questions about how and why individuals choose to participate. In this context, understanding the nature of engagement becomes essential. Engagement in citizen science is often defined through output measures such as the number of

Level 4. Extreme citizen science

- Collaborative science: problem definition, data collection and analysis.

Level 3. Participatory science

- Participation in problem definition and data collection.

Level 2. Distributed intelligence

- Citizens as basic interpreters.
- Volunteered thinking.

Level 1. Crowdsourcing

- Citizens as sensors.
- Volunteered computing.

Fig 1. Levels of participation and engagement in CS projects according to Haklay [11].

<https://doi.org/10.1371/journal.pone.0333091.g001>

Table 1. Citizen science initiatives on hydrological and meteorological data rescue and digitisation.

Project name	Variable(s)	Area	Volunteers involved and period of the activities (as of 30 March 2025)	Link to source and/or reference
Alpine weather news	Meteorological measurements published in local newspaper	City of Bolzano/ Bozen, South Tyrol, Italy	408 (March 2023 – April 2023)	https://www.zooniverse.org/projects/yuri-dot-brugnara/alpine-weather-news
AtmosElec – Atmospheric Electricity for Climate	Long records of atmospheric electric field measurements	Lerwick, Shetland	1158 (May 2024 – ongoing)	https://www.zooniverse.org/projects/hripsi-19/atmoselec-atmospheric-electricity-for-climate
Arctic archives	Weather observations	Greenland	1439 (April 2024 – July 2024)	https://www.zooniverse.org/projects/johandmi/arctic-archives-unraveling-greenlands-weather-history
Climate History Australia	Weather observations	Australia	2263 (September 2020 – July 2021)	https://www.zooniverse.org/projects/caitlinhowlett/climate-history-australia
Digitization Of analogue records In Time (DO IT)	Rainfall, snow and discharge observations	Germany	–	https://doi-lwi.tu-braunschweig.de/en/index.html
Jungle Weather project	Weather observations	Tropical rainforest of the Democratic Republic of the Congo	2506 (April 2020 – May 2020)	https://www.zooniverse.org/projects/khufkens/jungle-weather
Meteororum ad Extremum Terrae (Meteorology of the End of the World)	Weather observations	Argentina	–	https://www.zooniverse.org/projects/acre-ar/meteororum-ad-extremum-terrae ; [30]
Old Weather: Arctic	Sea ice and weather observations	Arctic seas	–	https://www.oldweather.org/naval_rendezvous.html
Old Weather: Whaling	Sea ice and weather observations	Arctic seas	–	https://www.oldweather.org/shipping_office.html
Old Weather - WW2	Marine weather observations	Seas and oceans	4463 (December 2020 – December 2021)	https://www.zooniverse.org/projects/krwood/old-weather-ww2
Rainfall rescue	Rainfall observations	UK	17279 (March 2020 – April 2020)	https://www.zooniverse.org/projects/edh/rainfall-rescue ; [31, 32]
Southern Weather Discovery	Weather observations	New Zealand, the Southern Ocean and Antarctica	2922 (October 2018 – April 2020)	https://www.zooniverse.org/projects/dreedeepsouth/southern-weather-discovery
Weather Rescue – Oxford	Weather observations	UK	544 (October 2023 – November 2023)	https://www.zooniverse.org/projects/edh/weather-rescue-oxford
Weather rescue at sea	Weather observations	Global	3752 (October 2021 – April 2023)	https://www.zooniverse.org/projects/p-teletti/weather-rescue-at-sea

<https://doi.org/10.1371/journal.pone.0333091.t001>

participants, the quantity of data collected, submission rates, web pages accessed, or other appraisals of recruitment, retention, and outreach [33]. However, such a definition is not able to capture the multidimensionality of engagement. In an attempt to provide a simple and clear definition of this term, [33] suggested a definition of engagement in citizen science as the “*emotional, behavioral, cognitive, and social experiences that initiate and sustain lifelong learning and that are largely influenced by motivational factors*”. Thus, engagement is strictly linked with the motivations, but the motivations themselves can not guarantee continuous engagement.

In recent years, few research works addressed the investigation of the motivations that lead people to be engaged in citizen science projects. [34] highlighted, in a review article based on peer-reviewed publications on environmental monitoring in developed countries published in the year 2005 or later, that intrinsic motivation (meaning that an activity is undertaken out of interest, enjoyment, or satisfaction) is the strongest one. [33] pointed out that this is particularly

true in contributory and collaborative CS projects. [35], however, found that intrinsic motivation mainly influences the quantity of participation, but that it did not affect the quality of participation (i.e., the accuracy and precision of contributions). [36] pointed out that the scientific activities offered in a CS project are often only sufficiently attractive for some participants: to improve the intrinsic motivation for participation and at the same time enhance participants' learning, project administrators must offer multiple engaging learning opportunities that can address the needs of all target groups. Personal interest in the topic or in the particular project also motivates people to be involved, together with recognition and attribution of their contributions [34,37–39]. In the same review paper, [34] recognized that several other extrinsic motivations emerged from the interviews carried out within several citizen science projects, such as learning, knowledge exchange, the urge to make a contribution (engagement provides individuals with a sense of purpose and contribution to a larger cause) or ego-involvement (volunteers can have a feeling that they are needed or that their friends will think positively about them if they participate in these activities). [40] pointed out from the experience gained during a case study that the motivations behind citizen scientist involvement depend on the country: citizen science in developed countries is a pastime, while citizen observers in developing countries very often earn an income from their work.

In the literature, an in-depth understanding of the motivations behind individual participation in initiatives like SIREN is still missing. Why do people choose to contribute with their time and effort to digitizing historical data? What drives their engagement in a citizen science project in the field of environmental sciences? Can we identify common profiles when analysing the different contributors?

To address these questions, in this work we conducted an anonymous survey among SIREN project participants. This survey was designed to reconstruct contributing practices, and, from there, to capture a comprehensive view of the factors influencing contributors' involvement, including personal interests, perceived benefits, and the overall appeal of the project and, more in general, of the citizen science initiatives. By analyzing survey responses, we aim to understand the key motivations that drive citizen scientists to engage with the SIREN project. Our goal is to provide actionable insights that can enhance the design and effectiveness of similar citizen science initiatives.

2. Methodology

As mentioned in the Introduction, the SIREN project aims to digitize the historical daily discharge time series from the Italian historical hydrological yearbooks and to produce a consistent and open-access dataset. The yearbooks contain more than 17.000 tables with daily discharge data (each table typically contains the measurements collected in one year by one gauging station). Fig 2 shows an example of a typical yearbook page containing discharge measurements. From this project, we expect to retrieve observations of about 1000 stations installed along the Italian rivers between the 1920s and 1990s.

After a review of all the freely available tools that can assist this type of project, we selected the user-friendly Zooniverse platform. The reasons that led us to choose this platform include its suitability for digitization projects like SIREN, without having the necessity of developing custom tools, e.g., to automatically assign pages to different volunteers. Additionally, the Zooniverse community is highly active, and the anonymous review process conducted by a small group of volunteers before the official launch (a mandatory step needed to be listed as an official Zooniverse project) allows researchers to test the workflow with individuals of various ages, nationalities, backgrounds, and experiences. We also chose the Zooniverse platform for its strong commitment to user privacy. Participants can contribute either by creating an account or anonymously, without the need to log in. For registered users, the only information visible to the project team is the chosen username and the timestamp of each digitization session. No additional personal data (such as email addresses, location, or demographic details) is shared with project managers. This privacy-by-design approach aligns well with the ethical standards of our research and ensures that participants can engage with the project while maintaining full control over their personal information.

XIV — Arno a S. Giovanni alla Vena

CARATTERISTICHE DELLA STAZIONE: a) — Bacino di dominio: kmq 8186 (parte prevalentemente permeabile: 5%; parte alluvionale: 25%); altitudine media del bacino: m 830 s.l.m.; distanza dalla foce: Km 37 circa; inizio delle misure: 27 luglio 1922. b) — Idrometro di riferimento (regolatore): « Botte dell'Emisario di Bientina » (in corrispondenza dell'asse del manufatto sp. d.); quota zero: m 6,78 s.l.m.; inizio delle osservazioni: 1 luglio 1923; massima piena: m. 8,66 (29 gennaio 1929); massima magra: m. 0,04 (15 agosto 1931). c) — Portate (1924-1942): annua media: mc/sec. 112,00 (l/sec. Amq. 13,7); massima: mc/sec. 2110,00 (l/sec. Amq. 267,7) (2 gennaio 1929); minima: mc/sec. 2,20 (l/sec. Amq. 0,3) (dal 2 al 6 e dal 12 al 25 agosto 1931).

PORTATE MEDIE GIORNALIERE in mc/sec												
Mese	Gennaio	Febbraio	Marzo	Aprile	Maggio	Giugno	Luglio	Agosto	Settembre	Ottobre	Novembre	Dicembre
1	28,90	35,10	938,00	60,50	357,00	34,10	19,90	14,50	7,44	34,10	43,20	24,20
2	26,90	33,80	261,00	57,00	281,00	35,10	19,90	13,80	6,06	26,30	292,00	22,70
3	27,80	30,50	202,00	52,50	376,00	34,10	19,90	13,10	7,44	29,70	115,00	21,80
4	27,90	28,70	173,00	52,50	258,00	33,30	13,10	13,90	7,44	17,20	72,50	20,90
5	26,90	28,70	163,00	55,10	307,00	33,30	13,10	11,70	6,70	17,20	54,10	19,90
6	27,80	31,40	159,00	72,00	164,00	34,10	10,80	9,35	15,40	49,10	49,70	55,20
7	27,80	55,00	169,00	88,80	188,00	33,30	17,20	10,80	8,35	14,50	400,00	338,00
8	31,40	124,00	338,00	80,00	116,00	33,30	19,90	10,80	8,35	14,50	202,00	154,00
9	30,60	281,00	297,00	71,50	109,00	33,30	18,10	13,80	9,08	14,50	201,00	94,60
10	25,50	560,00	304,00	64,10	31,40	31,40	6,50	10,80	8,35	23,40	193,00	72,10
11	27,80	97,00	161,00	68,00	33,60	30,50	15,90	10,80	9,88	20,50	115,00	68,90
12	26,90	163,00	139,00	50,70	92,80	30,50	17,20	10,80	8,35	18,10	80,50	43,80
13	26,90	128,00	123,00	46,10	37,90	29,50	17,20	9,08	9,08	16,30	59,60	37,10
14	26,00	99,00	116,00	42,50	160,00	30,50	15,40	9,08	9,08	16,30	43,70	38,10
15	26,00	89,40	110,00	39,90	170,00	33,70	16,30	9,08	7,44	14,50	39,10	39,00
16	26,00	76,80	102,00	35,10	137,00	26,70	13,10	9,35	8,35	14,50	34,10	147,00
17	26,00	57,10	164,00	39,50	111,00	31,40	15,40	9,08	9,08	14,50	31,20	128,00
18	24,90	54,90	127,00	31,40	94,00	30,50	15,40	8,35	7,44	16,60	39,20	185,00
19	25,10	49,60	147,00	45,40	32,10	33,20	26,10	10,80	7,44	15,40	23,20	594,00
20	26,00	45,60	217,00	127,00	70,20	45,60	39,80	9,08	8,35	16,60	27,10	865,00
21	24,20	44,60	169,00	143,00	68,10	66,40	39,70	9,35	6,70	15,40	24,50	469,00
22	24,20	49,30	132,00	309,00	57,00	176,00	23,20	7,44	9,35	19,90	23,50	307,00
23	23,40	51,00	144,00	155,00	59,80	45,90	32,60	9,08	14,50	19,00	21,50	209,00
24	24,90	846,00	110,00	140,00	49,80	39,20	17,70	9,08	32,80	19,00	20,80	161,00
25	54,90	1440,00	102,00	112,00	45,60	26,70	16,30	9,88	19,90	18,10	21,80	180,00
26	120,00	697,00	89,40	254,00	44,60	32,70	16,10	9,08	16,30	16,30	19,90	141,00
27	70,00	331,00	81,00	318,00	40,80	32,70	16,30	7,44	32,60	20,90	19,90	112,00
28	40,80	546,00	73,90	173,00	39,90	25,60	15,60	7,44	34,90	37,70	19,90	37,90
29	34,10	73,60	118,00	43,70	29,20	15,60	6,70	7,30	118,00	23,90	33,90	79,20
30	37,80	71,50	604,00	40,80	30,80	15,10	7,44	69,20	71,50	26,40	49,20	88,90
31	31,40	70,20	35,10	14,90	6,70	41,10	86,60					
Tot. 1924	1012,10	6454,50	4854,90	3706,00	3704,90	1111,70	588,20	312,66	520,62	749,10	2405,20	4888,90

Portate				Frequenza		Durata	
da mc/sec	a mc/sec	Frequenza (giorni)	Durata (giorni)	da mc/sec	a mc/sec	Frequenza (giorni)	Durata (giorni)
1440,00	1301,00	1	1	80,00	70,10	14	115
1300,00	1001,00	—	1	70,00	60,10	9	124
1000,00	801,00	2	3	60,00	50,10	14	133
800,00	601,00	2	5	50,00	45,10	11	149
600,00	501,00	4	9	45,00	40,10	10	159
500,00	401,00	2	11	40,00	35,10	12	171
400,00	351,00	2	13	35,00	30,10	28	199
350,00	301,00	7	20	30,00	25,10	35	234
300,00	251,00	6	23	25,00	20,10	35	257
250,00	201,00	8	26	20,00	15,10	21	278
200,00	151,00	2	28	18,00	12,10	15	293
180,00	161,00	12	50	16,00	10,10	19	312
160,00	141,00	9	59	14,00	8,10	5	317
140,00	131,00	11	70	13,00	7,10	3	325
120,00	101,00	14	84	10,00	6,01	2	350
100,00	90,10	8	92	8,00	7,01	10	360
90,00	80,10	9	101	7,00	6,05	6	365

Scala numerica delle portate			
Altezza idrometrica		Portata mc/sec	
dal 1/1/43 al 23/9/43	dal 23/9/43 al 31/12/43	dal 1/1/43 al 23/9/43	dal 23/9/43 al 31/12/43
0,10	9,40	0,10	4,00
0,15	13,80	0,15	6,70
0,20	17,50	0,20	10,80
0,30	26,00	0,30	19,90
0,40	35,10	0,40	29,20
0,50	44,60	0,50	39,10
1,00	97,00	1,00	93,90
1,50	156,00	1,50	156,00
2,00	226,00	2,00	226,00
2,50	307,00	2,50	307,00
3,50	497,00	3,50	497,00
4,50	743,00	4,50	743,00
5,50	1030,00	5,50	1030,00
6,50	1390,00	6,50	1390,00
7,50	1950,00	7,50	1950,00

per H ≥ 7,50 Q = 117,3 (H - 1,70) 1/2

ELEMENTI CARATTERISTICI PER L'ANNO 1942													
	Anno	Gennaio	Febbraio	Marzo	Aprile	Maggio	Giugno	Luglio	Agosto	Settembre	Ottobre	Novemb.	Dicemb.
Q max: mc/sec.	1440,00	1300,00	1440,00	883,00	604,00	376,00	176,00	88,90	18,10	94,90	113,00	460,00	996,00
l/sec. Amq.	175,9	14,7	175,9	40,7	73,8	45,8	21,4	4,7	2,2	11,6	13,8	56,2	105,7
Q min: mc/sec.	6,05	19,00	23,70	70,30	81,40	85,10	80,80	14,90	6,70	6,05	18,60	19,90	19,90
l/sec. Amq.	0,7	2,3	3,5	8,6	8,8	4,3	2,5	1,8	0,8	0,7	1,7	2,4	2,4
Q med: mc/sec.	88,10	32,60	231,00	157,00	124,00	119,00	87,10	19,00	10,10	17,40	24,20	80,20	158,00
l/sec. Amq.	10,2	4,0	28,2	19,2	15,1	14,5	4,5	2,3	1,2	2,1	3,0	9,8	19,3
Deflusso: 10 ⁶ mc.	2618,5	87,4	557,7	419,4	290,2	220,1	96,1	50,8	27,0	45,0	64,7	207,8	422,3
mm.	519,8	10,7	68,1	61,3	59,1	59,1	11,7	6,3	3,3	5,5	7,9	26,4	51,6
Afflusso: 10 ⁶ mc.	8141,8	3741,6	1060,0	494,4	1081,4	470,7	594,3	505,1	188,4	939,7	612,3	353,0	1023,2
mm.	994,6	45,7	139,5	60,4	136,0	57,5	72,6	61,7	22,4	114,8	74,3	104,2	136,0
Perd. app. mm.	674,8	35,0	61,4	9,2	86,9	13,4	60,9	55,5	19,1	109,3	66,9	78,8	79,4
Coef. di def.	0,39	0,23	0,53	0,85	0,31	0,68	0,16	0,10	0,15	0,05	0,11	0,24	0,41

ELEMENTI CARATTERISTICI PER IL PERIODO 1924-1941													
	Anno	Gennaio	Febbraio	Marzo	Aprile	Maggio	Giugno	Luglio	Agosto	Settembre	Ottobre	Novemb.	Dicemb.
Q max: mc/sec.	2060,00	2060,00	1810,00	1680,00	808,00	998,00	927,00	381,00	147,00	735,00	940,00	1810,00	1980,00
l/sec. Amq.	251,6	251,6	221,1	205,2	98,1	120,0	88,8	34,3	18,0	89,3	114,8	221,1	241,9
Q min: mc/sec.	2,20	10,40	16,00	17,90	18,60	21,10	9,26	3,85	3,20	2,90	4,00	7,40	9,70
l/sec. Amq.	0,3	1,5	2,0	2,1	2,3	2,6	1,1	0,4	0,5	0,5	0,5	0,9	1,3
Q med: mc/sec.	114,00	177,00	197,00	211,00	121,00	113,00	54,40	21,90	11,90	20,10	75,80	170,00	198,00
l/sec. Amq.	13,9	21,6	24,1	25,8	14,8	13,8	6,6	2,6	1,4	2,5	9,3	20,8	24,2
Deflusso: 10 ⁶ mc.	3685	474	480	564	314	308	141	57	30	52	208	439	538
mm.	439	58	59	69	38	37	17	7	4	6	26	54	65
Afflusso: 10 ⁶ mc.	9776	745	704	819	698	635	540	378	377	647	1121	1105	917
mm.	1072	91	86	100	84	102	66	34	46	79	137	136	112
Perd. app. mm.	683	39	37	31	46	65	49	27	43	73	112	81	47
Coef. di def.	0,41	0,64	0,69	0,69	0,45	0,86	0,26	0,21	0,09	0,08	0,16	0,40	0,58

PORTATE	1942		1924-1941	
	mc/sec	l/sec. Amq.	mc/sec	l/sec. Amq.
di giorni 91	92,80	11,3	128,00	15,6
di giorni 182	33,20	4,1	52,80	6,5
di giorni 274	18,10	2,2	17,70	2,2

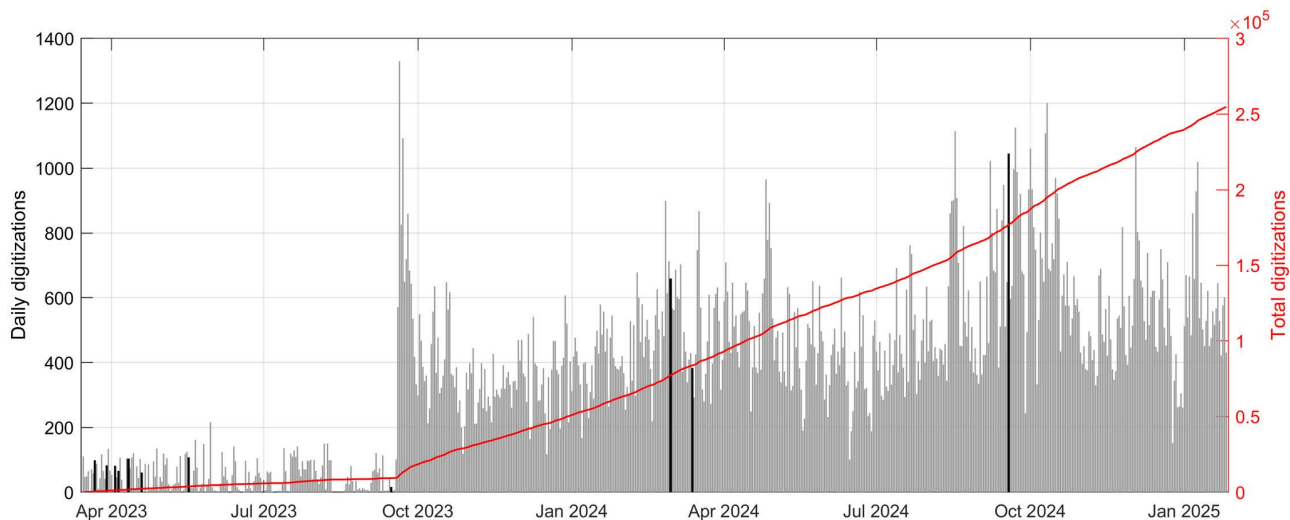


Fig 3. Number of daily contributions to the SIREN project. In black, the days in which we advertised the project on social media or on blogs of the sector. The first beta version was launched on 22nd March 2023, in the occasion of the 2023 World Water Day. The final version was released on 19th September 2023.

<https://doi.org/10.1371/journal.pone.0333091.g003>

Fig 3 shows the daily contributions (in terms of number of months digitized) with the indication of the main advertising activities in black color. The project was advertised in scientific conferences, media channels and blogs of hydrological sciences. From Fig 3 we can not see an increase in the digitizations following the advertisements. We hypothesized that people from Academia, who were mainly reached by these dissemination activities, rarely have the time or interest to contribute as anonymous volunteers to other research projects. On the contrary, we noticed a marked increase in the number of digitization starting from the day of the official launch on the Zooniverse platform in late September 2023, confirming the value of the Zooniverse community. From Fig 3 we can also see a modest decrease in the number of daily digitization after one week from the launch, probably due to new projects launched on the platform that divert the attention of the users. Thus, we decided to investigate the motivations that drive volunteers to be constantly involved in these initiatives.

Our analysis of engagement in the SIREN project employed a multi-step methodology: i) data analysis; ii) pilot interviews; iii) an anonymous survey.

In the first step, we approached the investigation with a quantitative data analysis: we collected and analyzed the digitization data from the Zooniverse platform to identify the contributions of each registered anonymous volunteer (with this term we identify the contribution performed by volunteers with a Zooniverse account; the term anonymous is used to highlight that for those users we only know the username).

Based on the findings of the data analysis, we conducted a few pilot interviews with selected individuals identified from the contribution's statistics. The selection of subjects for interviews (whose only known information was their username on Zooniverse) was guided by the criterion of including at least one intensive user and a couple of users with a lower number of entries. Recruitment for the interviews lasted from 23rd December 2023–15th January 2024. Recruitment was carried out by contacting SIREN participants via the Zooniverse platform, where users are present anonymously. Three people agreed to be interviewed. The three participants gave their consent to the processing of personal data, under the guarantee of the European GDPR, by signing an informed consent document.

We scheduled an hour-long interview with each of them to gain deeper insights into participant motivations and experiences. We chose the method of semi-structured interviews [41], presenting ourselves to the interviewees as social scientists trying to understand how to gain a better understanding of how CS projects function, with the goal of enhancing the user experience of activities like those on Zooniverse – which was also a point of the research.

The interviews were structured around three main narrative blocks [42]. The first block included questions about their experience with data entry for the SIREN project, their interaction with the web interface, the consistency of these operations over time, and the amount of time they dedicated to this activity. The second block focused on citizen science projects in general, asking whether they had participated in other projects (either on Zooniverse or elsewhere) and if those were similar to SIREN. The final block of questions delved into the respondents' interests in participating in SIREN and similar initiatives, whether their interest stemmed from the specific topic, environmental issues, or other factors.

The interviews have been mainly analyzed through grounded theory [43].

Based on the findings of the previous step, we structured an anonymous questionnaire (S1 File) around two hypotheses: 1) that a significant portion of SIREN contributors shared aspects of this profile, and 2) that there was another segment with different profiles that we had not yet identified and needed to uncover.

The survey was launched on the 20th of March 2024 and closed on the 10th of April: a banner that advertised it was inserted in the Zooniverse platform. To incentivize participation, 284 Zooniverse users among the most active were contacted via private message (277 out of 284 digitized more than 10 months of data), asking them to fill out the form in their spare time. The questionnaire collected information anonymously, an aspect explained to participants in a recruiting message on Zooniverse's SIREN project page.

3. Results

3.1 Data analysis

As of the 26th of January 2025, 1570 unique registered volunteers contributed to the digitization of at least one month of data. Users who did not log in before entering the digitization workflow can not be tracked. It is interesting to notice that the 6 most active users digitized about 50% of the data, and the 100 most active users digitized about 90% of the data (Fig 4). If we focus on the lower parts of the ranking, we notice that 729 registered users digitized 3 or fewer months. These data suggest that most of the work is performed by few loyal volunteers. The temporal evolution of the digitizations is visible in Fig 5. Fig 5a highlights heterogeneity in user activity, ranging from short-term engagement to sustained

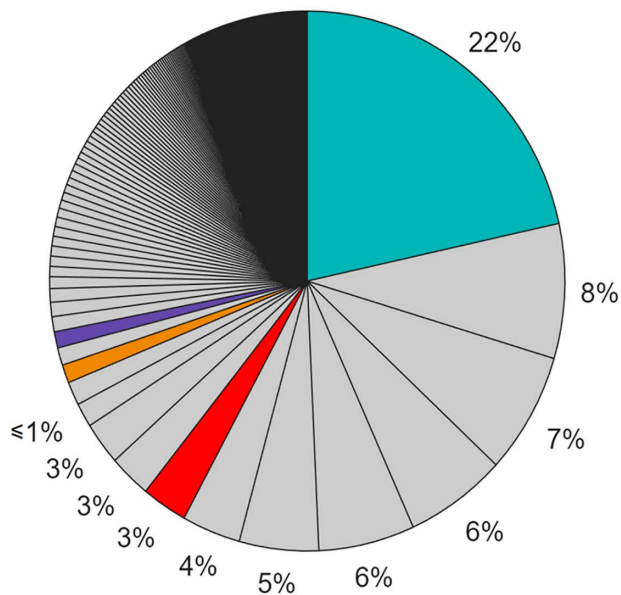


Fig 4. Contributions of each volunteer, expressed in percentage. In red, the 3% of the contributions that were made by users who did not log in. In light blue, orange, and purple, the contributions of three selected volunteers further analyzed in Fig 5.

<https://doi.org/10.1371/journal.pone.0333091.g004>

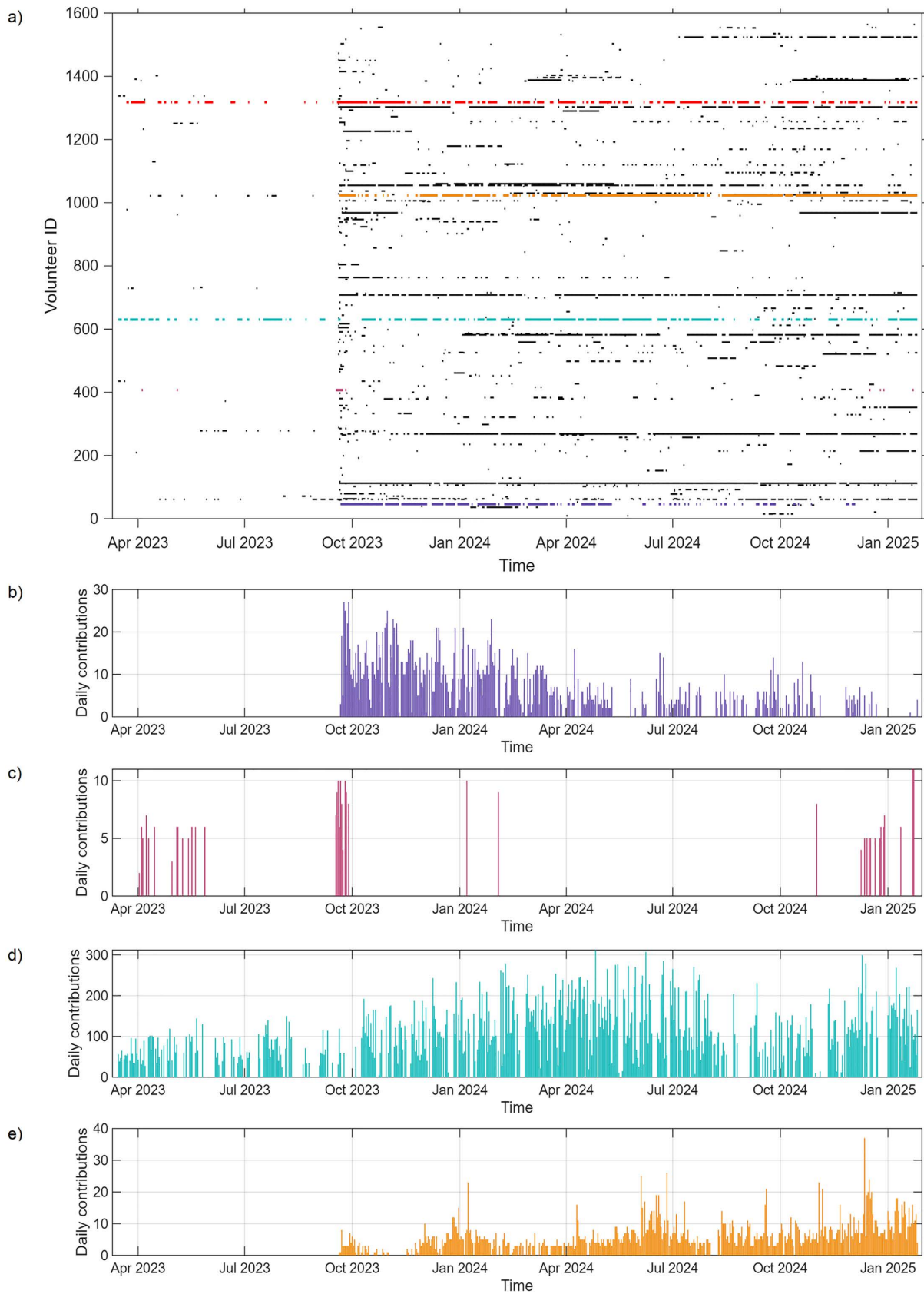


Fig 5. Temporal patterns of user activities in the SIREN project. Panel a shows daily engagement across all volunteers (each horizontal line represents one volunteer, and each vertical column corresponds to a day in the project timeline). Cells are colored black when the corresponding volunteer contributed to at least one digitization task on that day, and white when no contribution was made. Contributions made by non-logged-in users are indicated in red, distinguishing anonymous inputs from those traceable to registered accounts. Panels b–e present the daily number of digitized months for four selected volunteers, illustrating different patterns of contribution intensity and consistency.

<https://doi.org/10.1371/journal.pone.0333091.g005>

participation over long periods. [Fig 5b–5e](#) show instead individual time series and allow for a closer inspection of diverse engagement trajectories, including regular versus sporadic participation.

3.2 Pilot interviews

From the transcriptions of the interviews, we highlighted thematic cores and noted some interesting and unexpected patterns. First, the profiles of the three interviewees showed commonalities: they were women, retired, and had jobs related to data entry. It is worth remembering that the Zooniverse profiles from which they were selected are anonymous. Second, none of the three expressed particular interest in hydrology as a subject. As we explored topics that might align more closely with their interests, we noticed a recurring theme: rather than the subject matter itself, it was the practical tasks required by SIREN, such as copying and data entry, that contributed to the enjoyment they derived from participating in the project. All three interviewees found rewarding in their roles as contributors that they were offering concrete value to the project, as SIREN activities required skills they had developed in their professional lives.

We found these insights relevant and worthy of further investigation. However, we also acknowledged the possibility that these profiles, while similar to one another, might not be fully representative of broader participation in SIREN.

3.3 Anonymous survey

After a few weeks from the survey publication, we received 116 responses, comprising 49 partial and 67 complete responses. Some partial responses were deemed potentially useful for a first investigation since they replied to most of the questions and stopped answering only in the last phases, while all the others were discarded since the volunteer stopped answering after a few initial questions. In total, we obtained 69 useful responses (67 complete and 2 incomplete, see [S2 File](#)).

The gender identity of the volunteers was: 38 out of 69 (55%) female, 24 out of 69 (35%) male, 2 out of 69 (3%) non-binary, while 5 preferred not to answer. The female component is thus clearly predominant.

The dataset includes 58 valid responses to the question related to the age, with missing values and entries coded as –1 excluded from the analysis. The average age of respondents is 46.7 years, with a median age of 45.5 years and a standard deviation of 19.4 years. The youngest participant is 16 years old, while the oldest is 81. Overall, the sample covers a broad age range, encompassing almost the entire adult life course. The age distribution is visible in [Fig 6](#).

If we focus on the employment status:

- 29 out of 69 (42%) are employed full-time or part-time or self-employed;
- 19 out of 69 (28%) are retired;
- 8 out of 69 (12%) are students;
- 1 out of 69 is unemployed;
- 6 out of 69 (9%) fall in other categories;
- 6 out of 69 (9%) preferred not to answer.

The majority of the volunteers are thus employed or students, while about a quarter are retired.

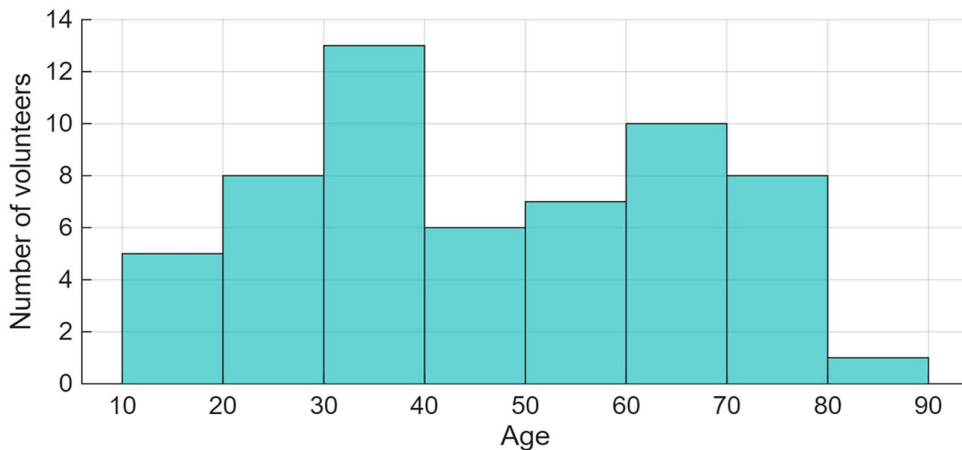


Fig 6. Age distribution of the volunteers.

<https://doi.org/10.1371/journal.pone.0333091.g006>

Almost all the volunteers usually work on SIREN from their home (62 out of 69, i.e., 90%) or from work (18 out of 69, i.e., 26%), while only a few of them work from public spaces (like libraries, cafes, or other recreational spaces) or from public means of transportation. Here, the total exceeds 100% because respondents were allowed to select more than one location, indicating that some volunteers contribute from multiple places depending on the context or availability. 36 out of 69 (52%) of them usually do other activities on the side or in the background when working on SIREN, like watching TV, listening to music, or talking with someone. 5 out of 69 (7%) of them found the activities requested during the digitization workflow akin to activities that they are/were used to perform for their job, while 26 out of 69 (38%) of them found a partial similarity.

Participants were drawn to the SIREN project for a variety of reasons, reflecting a range of personal interests and professional backgrounds. Here are the main points of interest that motivated their participation:

- interest in data digitization and climate studies: many participants were attracted to the project due to their interest or expertise in digitizing data, while others found it interesting since in the past they were already involved in climate and hydrological data digitization. They appreciated the straightforward, easy-to-understand tasks with clear instructions, which allowed them to contribute meaningfully to scientific research;
- connection to personal and professional backgrounds: some contributors had backgrounds in environmental science, hydrology, or related fields. These participants were motivated by the relevance of the project to their professional interests and research areas. Others were drawn by past experiences with similar projects, such as digitizing weather data or contributing to environmental data projects;
- citizen science and volunteering: the value of citizen science and the opportunity to contribute to important scientific research were significant motivators. Participants enjoyed the feeling of making a positive impact, even in a small way, on understanding climate change and preserving environmental data;
- personal interests and hobbies: some contributors were motivated by their personal interests, such as a fascination with weather data, natural disasters, or Italian geography. The project offered a way to engage with these interests while contributing to scientific research;
- ease and simplicity of the task: the simplicity and accessibility of the tasks were appealing to many. Participants liked that the tasks were easy to complete, required minimal effort, and could be done in short bursts of free time. This made the project a convenient and enjoyable way to contribute;

- invitation and word of mouth: a few participants joined the project after being invited by family members or learning about it from professors or online platforms like Zooniverse. This highlights the importance of personal networks and community engagement in attracting volunteers;
- nostalgia and personal connection: for some, the project evoked nostalgia for past experiences or personal connections to environmental studies. This emotional connection provided additional motivation to contribute.

Most users do not seem overly interested in counting their entries, suggesting that gamification factors are unlikely to be a significant motivation.

We asked participants how they felt after leaving the SIREN digitization workflow, and we received a range of different responses. The majority of them expressed positive feelings, such as accomplished, satisfied, relaxed, fulfilled, content, and glad. A few noted feelings of fatigue and exhaustion. Some participants expressed mixed emotions, combining satisfaction with other feelings such as curiosity, competitiveness, or frustration. Other notable responses included those who stopped due to external reasons like family time or physical discomfort. A small number of respondents felt neutral or indifferent. When answering this question, respondents were not constrained by predefined options and could freely express their thoughts in their own words. This unrestricted format allowed for diverse, personal responses. It is particularly noteworthy that, despite this freedom, many participants provided similar answers, often describing feelings of accomplishment, satisfaction, or fulfillment. This consistency across responses suggests a shared positive experience among contributors and highlights the intrinsic rewards associated with participating in the SIREN project.

Participants were asked to assess how much a series of terms applied to their experience with the SIREN project (Fig 7), using a 5-point Likert scale from 1 (the term does not apply) to 5 (the term applies very much). Overall, the most strongly associated terms were “accessible” (mean=4.51), “understandable” (4.30), “satisfactory” (4.19), and “easy” (4.12), indicating that volunteers generally found the platform clear, usable, and rewarding. Conversely, the lowest scores were recorded for “difficult” (1.74), “demanding” (2.30), and “boring” (2.29), suggesting that the task was not perceived as particularly fatiguing or cognitively heavy. Intermediate scores were assigned to terms like “engaging” (3.43), “fun” (3.17), and “relaxing” (3.55), indicating that while the activity was generally positive and pleasant, it was not necessarily considered exciting or entertaining by all.

The volunteers acknowledged the use of various tricks to enhance efficiency during the digitization process. It is interesting to notice that most of them dedicated time to detail the tricks they are using, probably in an attempt to provide

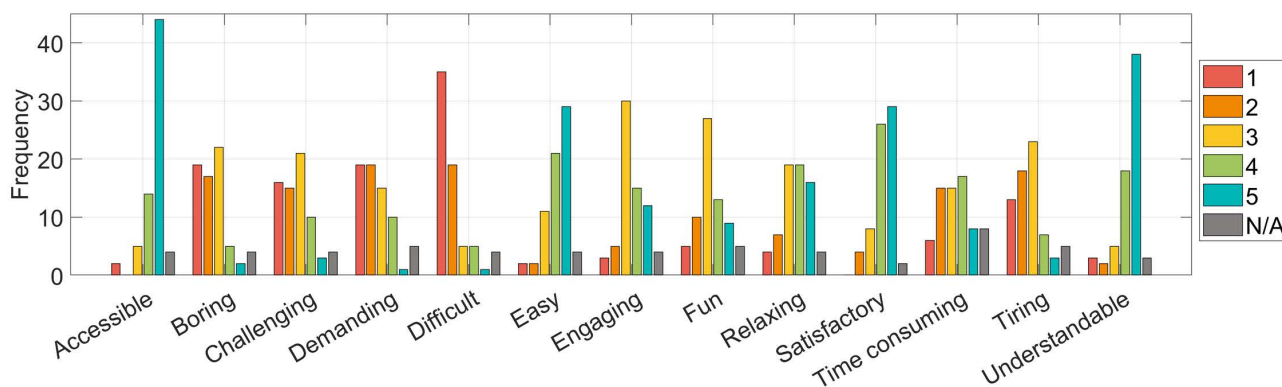


Fig 7. Answers to the question “How much do these terms apply to SIREN? Please rate, on a scale from 1 to 5, where 1 means that the term does not apply and 5 means that applies a lot, the following terms.”.

<https://doi.org/10.1371/journal.pone.0333091.g007>

project administrators with potentially useful information that can be shared with the community, or because they are proud of having identified them. A common approach was using copy-and-paste for repeated numbers to save time and reduce errors. Many are used to zooming in on the image, aligning the month to the screen's edge, or showing only 10 numbers at a time to maintain focus and accuracy. Some used optical character recognition (OCR) tools, although this often required significant corrections. Visual aids like holding half sheets of paper over the screen or dividing the image into sections with lines drawn in software like Microsoft Paint were also mentioned to avoid mistakes and keep track of progress. Others verified their entries by counting and comparing blocks of ten numbers, ensuring no entries were missing or duplicated.

Participants also employed various tools to enhance their speed and accuracy during the digitization process. A significant number relied on the number pad for quicker numerical entry, finding it faster and more convenient than the horizontal number keys above the letters on the keyboard. This method allowed them to focus on the screen without needing to look down at the keyboard. Some participants used dual screens or split their screen to manage the data entry process better. Additional tools included the mouse for navigating and zooming in and out, as well as pointing with fingers or a pencil to keep track of their position in the data.

Participants in the SIREN project reported a variety of learning outcomes. Many individuals improved their familiarity with Italian geography, including the names of months, places, and rivers. Some participants, particularly those learning Italian, found the project beneficial for language practice. Several contributors became more efficient with data entry techniques, learning to utilize tools like copy-paste functions and the number pad for faster input. They also gained insight into the methods of digitizing historical data and the importance of such efforts for scientific research. Participants also appreciated the historical aspect of the data, noting the meticulous records kept and the variability in river flows. They expressed admiration for the dedication of past scientists and the extensive data collection on even small rivers. The project also highlighted the differences in river dynamics and flow rates across various locations and times, which was something volunteers were not aware of.

Through the analysis of the questionnaires, we were able to identify three distinct participant profiles. Two of them are mainly activity-related, while the third is content-related.

The first one is composed of individuals who are not so much interested in the research topic, but are willing to contribute to a project considered useful through their skills without too much effort. Sentences like *"It was a simple task of transcribing numbers that I could do with little effort"* and *"I was looking for a transcription project on Zooniverse and SIREN looked interesting"* were common in this group. This group consists also of former data-entry employees (*"As several of my Zooniverse project contributions, I found I was drawn by the act of data entry. Many years ago, I was a data entry operator and perhaps that is what grabbed my attention. I also appreciate that I don't have to draw conclusions - just transfer the information!"*). This group values the straightforward nature of the work and the chance to remain involved without significant cognitive load.

The second one consists of those who use it as a leisure activity in their daily lives, especially during spare time or breaks in the working hours, recognising the ease of use and speed with which they can contribute (one respondent noted that SIREN is *"easy to use, useful, a good break"*, while another one said *"I work Zooniverse projects during lulls at work. I look for projects that can be completed in short bursts, that provide clear instructions, and that can be picked up at intervals without losing the thread of what I was doing"*).

The third is made by those who are professionally interested and contribute on the basis of an understanding of the research topics and objectives, underscoring a deeper engagement with the scientific objectives (*"I'm interested in hydrological data"*, *"I study hydrology and am learning Italian so it seemed pretty cool and similar to my situation currently!"*). In this case, factors in professional solidarity also emerge. In this third profile, we also included a few people who feel a relation to Italy (*"I'm often volunteering for many causes on Zooniverse, but SIREN is the first one that is strictly related to Italy (my country) so I hope to have been useful in helping control river conditions"*).

Statistical methodologies were used to analyse the results of the questionnaire to explore the relationship between sociodemographic characteristics and a set of subjective indicators related to the participants' experiences in the project.

The focus was placed on the three independent variables (gender, age, and employment condition), and only the 58 answers with complete information regarding the independent variables were considered. Within this subset, the gender distribution is: 34 out of 58 (59%) female, 22 out of 58 (38%) male and 2 out of 58 (3%) non-binary. As for the complete sample described at the beginning of this section, the average age of respondents is 46.7 years, with a median age of 45.5 years and a standard deviation of 19.4 years. The youngest participant is 16 years old, while the oldest is 81. If we focus on the employment status:

- 27 out of 58 (47%) are employed full-time or part-time or self-employed;
- 17 out of 58 (29%) are retired;
- 7 out of 58 (12%) are students;
- 1 out of 58 is unemployed;
- 6 out of 58 (10%) fall in other categories.

The three independent variables were tested against 13 dependent variables measuring participants' subjective experience with the project (e.g., "accessible", "boring", "challenging", "demanding", "difficult", "easy", "engaging", "fun", "relaxing", "satisfactory", "time consuming", "tiring", "understandable", as shown in Fig 7) evaluated on a 5-point Likert scale. In addition to analyzing the continuous responses, a binary (dummy) version of each variable was also created. This resulted in a total of 26 dependent variables (13 continuous and 13 binary) used in the analysis.

To explore potential differences in user perceptions based on demographic characteristics, we conducted a series of independent samples t-tests comparing male and female respondents. Specifically, we tested whether gender was associated with differences in how volunteers rated various aspects of the digitization activity (e.g., whether it was perceived as "engaging," "relaxing," "challenging," etc.). For each variable, we compared the mean responses between the two groups and computed the t-statistic and corresponding p-value (Table 2).

Table 2. Results of the t-test analysis performed to investigate user perception (evaluated on a 5-point Likert scale) based on demographic characteristics.

Variable	Mean value of the women sample	Mean value of the male sample	p-value	Test passed
Accessible (Likert scale)	4.72	4.23	0.0466	Yes
Boring (Likert scale)	2.00	2.43	0.1628	No
Challenging (Likert scale)	2.27	2.57	0.3585	No
Demanding (Likert scale)	2.15	2.30	0.6518	No
Difficult (Likert scale)	1.48	1.95	0.1218	No
Easy (Likert scale)	4.45	3.71	0.0162	Yes
Engaging (Likert scale)	3.79	3.15	0.0267	Yes
Fun (Likert scale)	3.39	3.05	0.2743	No
Relaxing (Likert scale)	3.94	3.15	0.0146	Yes
Satisfactory (Likert scale)	4.36	3.86	0.0539	No
Time-consuming (Likert scale)	2.97	3.26	0.4095	No
Tiring (Likert scale)	2.06	2.75	0.0197	Yes
Understandable (Likert scale)	4.24	4.41	0.5275	No

<https://doi.org/10.1371/journal.pone.0333091.t002>

From [Table 2](#) emerges that women evaluated the experience as more relaxing, engaging, easy and accessible, while men found it more tiring. No significant differences were found for the other variables. Results obtained with the binary (dummy) version of each variable are not reported since they do not provide relevant additional information.

To explore potential associations between age and the subjective experience of volunteers, we computed Pearson correlation coefficients between the age of the respondents and each of the variables describing their experience with the digitization task ([Table 3](#)). Before performing the correlation analysis, missing values were removed.

From [Table 3](#) emerges that as age increases, participants tend to perceive the experience as less challenging and less difficult.

To investigate whether subjective experiences differed across employment categories, we performed a one-way Analysis of Variance (ANOVA) for each of the dependent variables. Prior to conducting the ANOVA, we excluded any groups with fewer than 5 valid responses to ensure statistical robustness. For each variable, an ANOVA model was fitted using employment status as the independent factor. When the ANOVA yielded a statistically significant result ($p < 0.05$), a Tukey HSD test was applied to identify which specific employment categories differed significantly from one another. Overall, none of the ANOVA tests produced statistically significant results at a 5% level. In most cases, the p-values were well above this threshold, indicating no significant differences between groups.

Although some statistically significant differences emerged from the analysis, it is important to interpret these results with caution. The relatively small sample size limits the generalizability and statistical power of the findings, particularly in detecting complex relationships among variables. While the analyses provide valuable preliminary insights, they should be considered exploratory rather than definitive. Future research involving larger and more diverse samples would be essential to validate these initial observations, strengthen the reliability of the results, and allow for more nuanced statistical modeling. Such studies could also explore additional factors that may influence participants' engagement and experiences in citizen science projects.

4. Discussion

These profiles confirm what has already been observed in [\[13\]](#), namely that citizen science is linked to broader social trends such as the expansion of education, increased leisure time, attention to personal well-being, and, above all, widespread access to digital technologies and web applications. These developments have opened new avenues for participation in scientific research by enabling individuals to contribute from home or other informal settings. However, it is important to note that these opportunities are not equally accessible to the entire population, but rather to more economically affluent social groups (this, in itself, serves as a criterion for access to CS activities, which has not been explored here, although it may be partially inferred from certain data collected through the questionnaire).

This study also revealed the significant presence of what we defined as an activity-related profile and, in part, highlights a potentially original aspect of it in comparison to the reviewed literature. More specifically, in terms of activities, we observed that participant engagement was motivated not so much by the opportunity to learn new skills as an investment for the future, but rather by the desire to refine and enhance skills and competencies learned in one's professional life.

Related to motivations, participation is often analyzed as a form of investment in the future, as participants gain new skills, knowledge, or insights through their involvement. In contrast, for the group of individuals who engage in this project to leverage their existing expertise, SIREN is viewed as an investment in the present. Rather than focusing on learning

Table 3. Results of correlation analysis between age and user perception. Only the applications that resulted to be statistically significant are here reported.

Variable	Correlation	p-value	Test passed
Challenging (binary scale)	-0.357	0.007	Yes
Difficult (binary scale)	-0.293	0.0286	Yes

<https://doi.org/10.1371/journal.pone.0333091.t003>

something new, their motivation lies in applying and refining the skills they already possess, deriving satisfaction from contributing to a meaningful initiative in real-time. This distinction highlights the appeal of SIREN for participants who seek immediate purpose and productivity, rather than long-term personal development. By offering opportunities to utilize pre-existing competencies, the project serves as a platform for participants to channel their expertise into impactful scientific work, enabling also participants to remain professionally and intellectually active.

Furthermore, this work sheds light on a demographic that is often overlooked in the design of CS initiatives: retirees. A review of the literature reveals limited evidence that retirees are significantly engaged in citizen science activities, particularly online. Indeed, most of the CS projects consider youth as one of the main targets [44–46]. However, this project highlights a notable exception, as a substantial proportion of contributors are retired individuals, many of whom bring with them decades of work experience in relevant fields. This finding suggests that online citizen science projects, such as SIREN, may hold appeal for retirees, offering them opportunities to apply their skills, stay mentally active, and contribute to scientific initiatives. This piece of evidence underscores the importance of considering retirees as a valuable and underutilized resource in the design and promotion of future citizen science projects.

In sum, within this work, the responses have highlighted some previously unseen participant profiles and patterns of engagement, which can inform future iterations of the project and similar citizen science initiatives. These findings serve as a new hypothesis and starting point for exploring the diversity of motivations and experiences among contributors. Despite these new insights provide a new point of view for understanding the motivations and characteristics of participants in the SIREN project, the scope of the analysis is limited to some qualitative insights by the relatively small sample size of 69 valid responses, which is insufficient for conducting in-depth statistical analyses since it lacks the statistical power needed to draw robust conclusions or identify broader trends with high confidence. Future work could focus on expanding the sample size to strengthen the reliability of the findings and allow for a more comprehensive statistical evaluation.

Despite the limitations, this analysis represents an important step toward understanding the complex and multifaceted dynamics of volunteer engagement in citizen science projects. It underscores the importance of designing projects that not only welcome broad participation, but also recognize and support the diverse ways in which different people (depending on their background, experience, and life stage) may wish to contribute to science. Further research building upon these initial observations will be critical for deepening our understanding of how citizen science can inclusively engage and sustain a wide range of participants.

5. Conclusions

In this work we investigated the motivations behind citizen engagement and the possible presence of volunteer profiles in the SIREN project, a citizen science initiative that aims at digitizing the historical daily discharge measurements collected along the Italian rivers, available only in printed format, thus not easily accessible by researchers, practitioners, and policymakers.

The investigation consists of a multi-step methodology: i) analysis of quantitative data on citizen contributions; ii) three pilot interviews with selected individuals identified from the contributions statistics obtained in the previous step; iii) a comprehensive questionnaire open to a broader audience to verify and expand upon the insights gained from the pilot interviews.

Overall, the SIREN project attracted a diverse group of participants, each bringing their own unique motivations and interests to address the same common objective. Through the analysis of the questionnaires, we identified three distinct participant profiles in SIREN, two primarily activity-related and one content-related. The first group leverages its data entry skills with minimal effort and shows limited interest in the research topic, while the second views digitization as a leisure activity easily integrated into short breaks. The third profile demonstrates deeper engagement, with volunteers who possess professional or personal interests in hydrology or Italian geography. Many participants regard SIREN as an investment in the present, focusing on applying existing skills rather than acquiring new ones.

Contrary to literature suggesting limited retiree engagement in online citizen science, a notable proportion of SIREN's contributors are retired individuals, highlighting their potential as an underutilized resource.

To further investigate relationships between demographic characteristics and subjective experiences, we conducted statistical analyses on survey responses. T-tests, correlation analyses, and ANOVA tests were applied to assess the influence of gender, age, and employment status on users' perception of the digitization activity. While some statistically significant results emerged (such as women reporting the experience as more relaxing and engaging, and older users finding it less challenging) the overall explanatory power of demographic factors was limited. Nevertheless, these analyses provided useful quantitative support to the qualitative insights, reinforcing the importance of considering both perspectives when designing inclusive citizen science projects. The results presented in this work highlight the importance of creating citizen science projects that are not only scientifically meaningful but also tailored to the needs, interests, and capacities of diverse participant profiles.

To conclude, we would like to emphasize that exploring the relevance of the activity-related profile, as it has emerged in this context, holds particular promise for future projects. This approach could offer a fresh perspective on organizing participatory activities between scientific research and society at large. It goes beyond the traditional, linear model of citizen science, where expertise is confined to scientific institutions. Instead, it diversifies and enriches the process by highlighting specific activities, thereby incorporating expert contributions in both directions.

Supporting information

S1 File. The file contains the questionnaire that we shared with the citizen scientists.

(PDF)

S2 File. The file contains the answers to the questionnaire.

(XLSX)

Acknowledgments

We sincerely thank the 1570 volunteers for the transcription of the observations, Gail Erickson, Silvana Naselli and Sue Kennedy who agreed to be interviewed, and the 69 volunteers who agreed to answer all the questions of the survey.

This publication uses data generated via the Zooniverse.org platform, development of which is funded by generous support, including a Global Impact Award from Google, and by a grant from the Alfred P. Sloan Foundation.

Author contributions

Conceptualization: Paola Mazzoglio, Miriam Bertola, Tommaso Listo, Alvisè Mattozzi.

Formal analysis: Paola Mazzoglio, Miriam Bertola, Tommaso Listo, Luca Princivalle, Alvisè Mattozzi.

Funding acquisition: Paola Mazzoglio, Miriam Bertola.

Investigation: Paola Mazzoglio, Miriam Bertola, Tommaso Listo, Luca Princivalle, Alvisè Mattozzi.

Methodology: Paola Mazzoglio, Miriam Bertola, Tommaso Listo, Alvisè Mattozzi.

Project administration: Paola Mazzoglio, Miriam Bertola.

Software: Paola Mazzoglio, Miriam Bertola.

Supervision: Paola Mazzoglio, Miriam Bertola, Tommaso Listo, Alvisè Mattozzi.

Visualization: Paola Mazzoglio, Miriam Bertola.

Writing – original draft: Paola Mazzoglio, Miriam Bertola, Tommaso Listo, Alvisè Mattozzi.

Writing – review & editing: Paola Mazzoglio, Miriam Bertola, Tommaso Listo, Luca Princivalle, Luca Lombardo, Alberto Viglione, Pierluigi Claps, Alvisè Mattozzi.

References

- Mazzoglio P, Butera I, Claps P. I2-RED: a massive update and quality control of the Italian annual extreme rainfall dataset. *Water*. 2020;12(12):3308. <https://doi.org/10.3390/w12123308>
- Claps P, Ganora D, Apostolo A, Brignolo I, Monforte I. *Catalogo delle Piene dei Corsi d'acqua Italiani*. Vol. 1. CINID; 2020a.
- Claps P, Ganora D, Apostolo A, Brignolo I, Monforte I. *Catalogo delle Piene dei Corsi d'acqua Italiani*. Vol. 2. CINID; 2020b.
- Claps P, Ganora D, Apostolo A, Brignolo I, Monforte I. *Catalogo delle Piene dei Corsi d'acqua Italiani*. Vol. 3. CINID; 2020c.
- Claps P, Evangelista G, Ganora D, Mazzoglio P, Monforte I. FOCA: a new quality-controlled database of floods and catchment descriptors in Italy. *Earth Syst Sci Data*. 2024;16(3):1503–22. <https://doi.org/10.5194/essd-16-1503-2024>
- O.E.D. Citizen science. September 2024. <https://doi.org/10.1093/OED/6784505301>
- Kerson R. Lab for the Environment. *MIT Technol Rev*. 1989; 92(1):11–12.
- Kullenberg C, Kasperowski D. What is citizen science? A scientometric meta-analysis. *PLoS One*. 2016;11(1):e0147152. <https://doi.org/10.1371/journal.pone.0147152> PMID: 26766577
- Bonney R, Ballard H, Jordan R, McCallie E, Phillips T, Shirk J, Wilderman CC. Public participation in scientific research: defining the field and assessing its potential for informal science education - a CAISE inquiry group report. [Internet]. Washington, DC: Center for Advancement of Informal Science Education; 2009. <https://files.eric.ed.gov/fulltext/ED519688.pdf>
- Strasser BJ, Baudry J, Mahr D, Sanchez G, Tancoigne E. "Citizen Science"? Rethinking science and public participation. *S&TS*. 2018;;52–76. <https://doi.org/10.23987/sts.60425>
- Haklay M. Citizen science and volunteered geographic information: overview and typology of participation. In: Sui D, Elwood S, Goodchild M, editors. *Crowdsourcing geographic knowledge*. Dordrecht: Springer, 2013. p. 105–122. https://doi.org/10.1007/978-94-007-4587-2_7
- Halberda J, Ly R, Wilmer JB, Naiman DQ, Germine L. Number sense across the lifespan as revealed by a massive Internet-based sample. *Proc Natl Acad Sci U S A*. 2012;109(28):11116–20. <https://doi.org/10.1073/pnas.1200196109> PMID: 22733748
- Haklay M. *Citizen science and policy: a European perspective*. Washington, DC: Woodrow Wilson International Center for Scholars; 2015.
- Mayer EC. Citizen science and scientific authority: have you checked the boundary work? *Citizen Science: Theory and Practice*. 2022;7(1). <https://doi.org/10.5334/cstp.519>
- Jasanoff S. *States of knowledge: the co-production of science and the social order*. London: Routledge; 2004.
- Vohland K, Land-Zandstra A, Ceccaroni L, Lemmens R, Perelló J, Ponti M, et al. *The Science of Citizen Science*. Springer International Publishing; 2021. <https://doi.org/10.1007/978-3-030-58278-4>
- Kullenberg C. Citizen science as resistance: crossing the boundary between reference and representation. *J Resistance Studies*. 2025;1(1):50. <https://doi.org/10.63961/2025.017>
- European Commission, Directorate-General for Research and Innovation, Iagher R, Monachello R, Warin C, Delaney N, Tornasi Z. *Science with and for society in Horizon 2020: achievements and recommendations for Horizon Europe*. Luxembourg: Publications Office of the European Union; 2020. <https://doi.org/10.2777/32018>
- Minniti S, Magaudda P. The 'obligatory passage point' in knowledge co-production: Italy's participatory environmental monitoring platform. *Science as Culture*. 2024;34(1):13–30. <https://doi.org/10.1080/09505431.2024.2317236>
- AAPS – Association for Advancing Participatory Sciences. <https://participatorysciences.org/>
- Citizen Science: Theory and Practice. <https://theoryandpractice.citizenscienceassociation.org/>
- Zooniverse. <https://www.zooniverse.org/>
- Nardi F, Cudennec C, Abrate T, Allouch C, Annis A, Assumpção T, et al. Citizens AND Hydrology (CANDHY): conceptualizing a transdisciplinary framework for citizen science addressing hydrological challenges. *Hydrolog Sci J*. 2021;67(16):2534–51. <https://doi.org/10.1080/02626667.2020.1849707>
- Arienzo MM, Collins M, Jennings KS. Enhancing engagement of citizen scientists to monitor precipitation phase. *Front Earth Sci*. 2021;9. <https://doi.org/10.3389/feart.2021.617594>
- O'Hara T, McClean F, Villalobos Herrera R, Lewis E, Fowler HJ. Filling observational gaps with crowdsourced citizen science rainfall data from the Met Office Weather Observation Website. *Hydrol Res*. 2023;54(4):547–56. <https://doi.org/10.2166/nh.2023.136>
- Balbo A, Galimberti G. Citizen hydrology in River Contracts for water management and people engagement at basin scale. COWM2016 - International Conference on Citizen Observatories for Water Management. 2016 Jun 7-9; Venice, Italy. http://www.cithyd.com/images/spsimpleportfolio/citizen-observatory-water-management-2016-giugno-venezia/COWM16_abstract_template-paperABGG-persito.pdf
- Seibert J, Blanco S, Scheller M, Schwarzenbach F, Ze W, van Meerveld I. Engaging the public for water data collection – experiences from the CrowdWater project. *EGU General Assembly 2022*. 2022 May 23-27; Vienna, Austria. EGU22-6319, <https://doi.org/10.5194/egusphere-egu22-6319>
- Yevenes MA, Pereira H, Bermudez R. Citizen science as a co-creative measure to water quality: chemical data and local participation in a rural territory. *Front Environ Sci*. 2022;10. <https://doi.org/10.3389/fenvs.2022.940778>
- Brewin RJW, Brewin TG, Phillips J, Rose S, Abdulaziz A, Wimmer W, et al. A printable device for measuring clarity and colour in lake and near-shore waters. *Sensors (Basel)*. 2019;19(4):936. <https://doi.org/10.3390/s19040936> PMID: 30813342

30. Lakkis SG, Canziani PO, Rodriguez JO, Yuchechen AE. Early meteorological records from Corrientes and Bahía Blanca, Argentina: Initial ACRE-Argentina data rescue and related activities. *Geosci Data J.* 2022;10(3):328–46. <https://doi.org/10.1002/gdj3.176>
31. Hawkins E, Burt S, McCarthy M, Murphy C, Ross C, Baldock M, et al. Millions of historical monthly rainfall observations taken in the UK and Ireland rescued by citizen scientists. *Geosci Data J.* 2022;10(2):246–61. <https://doi.org/10.1002/gdj3.157>
32. WIRED. Citizen scientists digitized centuries of handwritten rain data [Internet]. WIRED; 2021. <https://www.wired.com/story/citizen-scientists-digitized-centuries-of-handwritten-rain-data/>
33. Phillips TB, Ballard HL, Lewenstein BV, Bonney R. Engagement in science through citizen science: moving beyond data collection. *Science Education.* 2019;103(3):665–90. <https://doi.org/10.1002/sce.21501>
34. Rutten M, Minkman E, van der Sanden M. How to get and keep citizens involved in mobile crowd sensing for water management? A review of key success factors and motivational aspects. *WIREs Water.* 2017;4(4). <https://doi.org/10.1002/wat2.1218>
35. Nov O, Arazy O, Anderson D. Scientists@Home: what drives the quantity and quality of online citizen science participation? *PLoS One.* 2014;9(4):e90375. <https://doi.org/10.1371/journal.pone.0090375> PMID: 24690612
36. Kaplan Mintz K, Arazy O, Malkinson D. Multiple forms of engagement and motivation in ecological citizen science. *Environ Educ Res.* 2022;29(1):27–44. <https://doi.org/10.1080/13504622.2022.2120186>
37. Raddick MJ, Bracey G, Gay PL, Lintott CJ, Murray P, Schawinski K, et al. Galaxy Zoo: Exploring the Motivations of Citizen Science Volunteers. *Astronomy Education Review.* 2010;9(1). <https://doi.org/10.3847/aer2009036>
38. Rotman D, Preece J, Hammock J, Procita K, Hansen D, Parr C, et al. Dynamic changes in motivation in collaborative citizen-science projects. In: *CSCW '12: Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work.* 2012.
39. Etter S, Strobl B, Seibert J, van Meerveld HJ, Niebert K, Stepenuk KF. Why do people participate in app-based environment-focused citizen science projects?. *Front Environ Sci.* 2023;11. <https://doi.org/10.3389/fenvs.2023.1105682>
40. Buytaert W, Zulkafli Z, Grainger S, Acosta L, Alemie TC, Bastiaensen J, et al. Citizen science in hydrology and water resources: opportunities for knowledge generation, ecosystem service management, and sustainable development. *Front Earth Sci.* 2014;2. <https://doi.org/10.3389/feart.2014.00026>
41. Moerman G. Probing behaviour in open interviews. a field experiment on the effects of probing tactics on quality and content of the received information. Amsterdam: VU University Amsterdam; 2010.
42. Bryman A. *Social research methods.* Don Mills, Ontario, Canada: Oxford University Press; 2016.
43. Charmaz K. *Constructing grounded theory.* 2nd ed. Los Angeles: Sage; 2014.
44. Ballard HL, Dixon CGH, Harris EM. Youth-focused citizen science: Examining the role of environmental science learning and agency for conservation. *Biol Conserv.* 2017;208:65–75. <https://doi.org/10.1016/j.biocon.2016.05.024>
45. Harris EM, Dixon CGH, Bird EB, Ballard HL. For science and self: youth interactions with data in community and citizen science. *J Learn Sci.* 2019;29(2):224–63. <https://doi.org/10.1080/10508406.2019.1693379>
46. Khanaposhtani MG, Ballard HL, Lorke J, Miller AE, Pratt-Tawah S, Jennewein J, et al. Examining youth participation in ongoing community and citizen science programs in 3 different out-of-school settings. *Environ Educ Res.* 2022;28(12):1730–54. <https://doi.org/10.1080/13504622.2022.2078480> PMID: 36217414