CS TRACK
Investigating Citizen Science

Horizon 2020 / Science with and for Society Programme

Grant agreement number: 872522

Report on a survey among organisers of citizen science projects

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 872522

Title of project	CS Track
Full title of project	Expanding our knowledge on Citizen Science through analytics and analysis
Title of this document	Report on a survey among organisers of citizen science projects
Number of this document	n/a
Dissemination level	Public (PU)
Due date	n/a
Actual delivery	n/a
Versioning history	Version 1: 13. April 2022
Authors	Michael Strähle & Christine Urban (alphabetical order), Wissenschaftsladen Wien - Science Shop Vienna
Executive summary→	This report has been compiled within CS Track's Work Package 2, which, among other objectives, aims to present a first overview of citizen science activities according to geographical location, topics, activities and project coordinators' estimates on the participation of different socio-economic groups and gender distributions within by - exploring the availability of data about citizen science activities; - setting up a comprehensive database of citizen science activities that is used for web analytics in Work Package 3; - conducting a short survey on them; and - carrying out exemplary qualitative research on selected citizen science projects to generate ideas for maximizing the benefit of citizen science activities for the participants and to unveil incentives, barriers, enablers, disadvantages and benefits. This report contributes to this objective by presenting the results of a short survey on citizen science activities that has been conducted among project owners. In an online survey the authors of this report asked project owners, resp. coordinators for the most elementary data about the respective project and those who participate(d) in it: - the project objectives, - the scientific disciplines involved in the project, - the type(s) of citizen science activities, - rough estimates on the participation of different social groups, including their gender and age distributions, and - questions on practical issues, such as the availability of the respective project for further research.

	This might be an indication In line with the authors' research so far, the survey showed a potential indication that many projects do not know very much about the participants, their characteristics or even their number (or not want to admit to it) and refrain from answering. In view of the benefits that several scholars, practitioners, policy makers and others claim citizen science brings with it, this would make some of them unfounded if not even implausible. Moreover, an attempt was made to investigate – in cases where academics were among the organisers - how far their expertises match(ed) the research areas of the projects. This proved exceptionally tricky because there exists no classification scheme which mirrors the broad variety of academic educations in different regions.

Table of Contents

1	Concept and rationale	5
2	Methodology	8
3	Results	9
4	Technical aspects	21
5	Conclusions	24
6	References	25
7	Annex	26

1 Concept and rationale

This report presents results of the research project CS Track which is funded by the European Commission under the Science with and for Society Work Programme. The aim of CS Track is to broaden the knowledge about citizen science and the impact citizen science activities can have. This overall objective is achieved by understanding and characterising citizen science activities so that one can say how they can be improved in terms of maximising their benefit for all participants and stakeholders, citizen and professional scientists, policymakers, and funders, while meeting scientific standards of validity and reliability, paying attention to caveats and potential pitfalls, and respecting research integrity and ethics. The CS Track consortium investigates a large and diverse set of citizen science activities, discusses practices, and formulates knowledge-based policy recommendations to maximise the potential benefit of citizen science activities on individual citizens, organisations, and society at large.

As stated in other reports, notably Deliverables D1.1 and D1.2, what the term "citizen science" refers to depends, among other things, on science cultures, research orientations, fields of research and the kind of citizen participation in the respective research activities (Eitzel et al., 2017; Kullenberg et al., 2016; Riesch et al., 2014; Heigl & Dörler, 2017). Among other activities, it can refer to crowdsourcing activities such as collecting weather data, to spotting animals in an online video, deciphering handwritten historic documents, solving scientific puzzles, or making experiments in your garden, but also to formulating research questions and even to setting research agendas, developing robotic prototypes or conducting practical science projects in schools. For the Science with and for Society Work Programme the European Commission offers a description of citizen science that includes activities ranging from school education through citizen participation in scientist-led research projects to fab labs and citizen engagement in science policy. In the framework of CS Track the consortium uses the explanation of citizen science the European Commission gives in the Science with and for Society Work Programme 2018-2020:

> *(…) citizen science should be understood broadly, covering a range of different levels of participation, from raising public knowledge of science, encouraging citizens to participate in the scientific process by observing, gathering, and processing data, right up to setting scientific agenda and co-designing and implementing science-related policies. It could also involve publication of results and teaching science. (European Commission, 2018, p. 41)*

The Science with and for Society call topic under which CS Track received a grant, and particularly the rationale of CS Track, aim at an integrated investigation of participation patterns; societal, democratic, and economic benefits of citizen science; incentives, disincentives, barriers, and enablers to involving and engaging citizens and scientists in citizen science activities. Equal access and absence of

discrimination are important desiderata for this endeavour. This brings in the questions of social conditions for access, gender equity, and world-wide accessibility.

This report has been compiled within CS Track's Work Package 2, which, among other objectives, aims to present a first overview of citizen science activities according to geographical location, topics, activities and project coordinators' estimates on the participation of different socio-economic groups and gender distributions within by

- exploring the availability of data about citizen science activities;
- setting up a comprehensive database of citizen science activities that is used for web analytics in Work Package 3;
- conducting a short survey on them; and
- carrying out exemplary qualitative research on selected citizen science projects to generate ideas for maximizing the benefit of citizen science activities for the participants and to unveil incentives, barriers, enablers, disadvantages and benefits.

This report contributes to this objective by presenting the results of a short survey on citizen science activities that has been conducted among project owners.

Evolution of research objectives and methodology

According to the original workplan, the survey was scheduled for Spring or early Summer 2020. The objective of the survey was two-fold: Firstly, to make it more attractive for project organisers to agree that the CS Track consortium conducts qualitative research on their citizen science projects. They should be ready to provide material for in-depth analysis and to inform project participants about the opportunity to become interview partners and conditions for these interviews. Instead of handpicking projects, it was planned to list them in a pool for selection via a lottery based public draw. The pool would be made transparent by publishing it on the web. Projects would have been selected according to a simple algorithm into which numbers of a public lottery draw are inserted. All information for the selection, including the date of the lottery draw, would have been public and thus transparent and eligible to the scrutiny of interested citizens. The launch of the survey required a lot of questions to be answered beforehand: Would the research be confidential or would the project organisers receive an occasion to present their projects? Who would be interviewed? Which kind of cooperation would be required? What degree of anonymity could be promised to participants? Communicating to project organisers details of the planned public lottery was equally important, because it promised to attract different types of projects than usual.

Secondly, the objective was to collect data on some characteristics on answering projects to collect data which cannot be researched on project websites. Survey

results would be fed into a database of citizen science activities that would provide a data pool for web analytics.

The target group were organisers of citizen science projects (and similar projects), irrespectively if these are ongoing or finished. To conduct research on projects that have not been in the limelight, the plan was to target organisers of projects with a) small or no funds and thus less time resources and b) those who were not or less interested in joining citizen science communities or maybe even opposed to such networks. The latter should be attracted with a "sortition" procedure, the first one with a really short questionnaire.

The COVID-19 pandemic began in Spring 2020. At first it was not clear that the crisis would last for so many months. Hence, until Spring/Summer 2020, the original plan to use the survey as an invitation to become part of qualitative research was followed. When it became clear that the pandemic would not stop soon, the qualitative research in WP2 had to be adapted in some respects. Several alternatives were worked out and discussed in the CS Track consortium. In the end it was decided to conduct the survey independently from the qualitative research on selected projects in Work Package 2.

Originally it was planned to conduct qualitative interviews with project participants, however, due to lockdowns and other restrictions of the COVID crisis, they would have to be conducted mostly online. Anonymity is key when frankly told personal experiences and opinions are asked for. Interview partners most likely would not have had the same trust in virtual conferences as in physical meetings with the interviewer.

An abundance of case studies was found in WP1, even more than had been anticipated. In these studies, often project organisers gave their views. This was another reason that spoke against adding more general case studies.

The research questions were:

- What do CS project organisers know about participants in their projects? Are they confident to estimate gender, age and social situation?
- How do academic disciplines attributed to a project match research expertise in the team of project organisers?
- Are there other response behaviour & response patterns of interest in a survey with a very short questionnaire?

2 Methodology

In this online survey the authors of this report asked project owners, resp. coordinators for the most elementary data about the respective project and those who participate(d) in it:

- the project objectives,
- the scientific disciplines involved in the project,
- the type(s) of citizen science activities,
- rough estimates on the participation of different social groups, including their gender and age distributions, and
- questions on practical issues, such as the availability of the respective project for further research.

The survey questionnaire was restricted to few questions (using mainly tick-boxes). Otherwise, it could create a barrier for smaller projects with little or no funding. All questions but question 1 were optional questions. The obligatory question 1 served to identify the project without doubt.

The survey did not include questions on information that could be otherwise researched. The nine questions that were finally chosen, focused on estimations of organisers about participants. Moreover, the questionnaire should not build a barrier to the less funded or not at all funded projects. Time resources are presumably low if there are no employees who can fill in lengthy questionnaires. Additionally, those projects that are keener on being part of citizen science networks would be more inclined to fill in a questionnaire and thus further distort the picture. Citizen science project organisers were targeted without pre-selection.

In April 2020, in a meeting between the work package coordinator, UPF, and WLW, it was decided to use LimeSurvey for the online survey, because both institutions had the experience that non-profit organisations that are sensitive about data protection trust this tool. The authors had chosen the Blue Circle template. It was agreed that most questions should be on project organisers' estimates of numbers or characteristics of participants in their projects.

Promotion of the survey

The promotion of the survey started in December 2020 and was paused in January 2021 to avoid overlap with the distribution of the survey in CS Track's Work Package 4 and to solve the technical problems with survey software mentioned below. It was still possible to respond but since the end of January 2021 no effort was made to promote the survey. So as a record day for an interim report we chose the 18th of February 2021, roughly one month after distribution had paused.

The survey was promoted by Twitter messages and a blogpost on *Österreich forscht*, the online platform of *Citizen Science Network Austria,* at the beginning of January 2021. In December 2020 the authors promoted the survey in scientific mailing lists and by contacting research and higher education institutions by email directly.

3 Results

About the responses

Completed responses used for analysis

Only three languages were used to answer the questionnaire: English, German and Greek.

Language of answers	
English	42
German	10
Greek	4
TOTAL	**56**

(The three discarded completed responses were all in English, too.)

It is surprising that only three of the many language versions available were used. Often English was used in spite of the availability of a language version that matched the official languages of the location of the projects.

Project location	English	German	Greek	TOTAL
Austria	1	7		8
Belgium	5			5
Canada	1			1
Chile	1			1
Cyprus	1			1
Denmark	1			1
European	2	1		3
France	2			2
Germany	2	2		4
Greece	3		4	7
Ireland	3			3

Countries for which versions in the dominant official language were available, apart from English.

Netherlands	2			2
Portugal	2			2
Spain	2			2
Sweden	1			1
Switzerland	1			1
UK	3			3
USA	9			9
TOTAL	**42**	**10**	**4**	**56**

Countries for which versions for a regionally spoken language were available.

Countries with English as the dominant official language.

This allows for several cautious interpretations: The organisers of citizen science projects we reached have a habit of using English as lingua franca for global communication, especially in academic contexts. The organisers wanted to spare the analysts (the authors) the translation effort. But one cannot rule out the possibility that not all language versions worked equally well and organisers used English for such a reason.

Characteristics of responding projects according to project websites

The questionnaire could be kept very short only because respondents were not asked to answer questions that could be later retrieved from project websites. The answers can neither be regarded as representative for projects that see themselves as citizen science nor do we suppose they cover the whole spectrum of possibilities. What they do show (again) is that there is an extremely broad range of projects crowding under this umbrella term.

With classifying science shop projects as citizen science, the application of the term "citizen science" has broadened further: In a narrower sense, the two responding science shops[1] perform not one, but many citizen science projects. In the classical science shop models "citizens" - often non-profit organisations - are mostly involved by requesting research on certain topics and not so much by carrying out research by themselves.

First language version clicked

Several language versions were available. The survey software did not give the information which language version was eventually used by a respondent, but which language version was clicked first. This gives some indication in which language the questionnaire was found. In 50 cases the "start language", the language version that was accessed initially, was English, in six cases it was German. This does not say in which language version the questionnaire was finally filled in as respondents could

[1] Science shops are contact points for public research requests. See Strähle, Urban et al. (2021), p. 21, for a more detailed presentation of the science shop concept.

switch to another language version. It is interesting insofar as the questionnaire was sent out by different partners with respective links to different language versions. This could be cautiously interpreted as an indication that information resources in English were most often used. Promotion messages contained a link to the CS Track website. Before they filled in the questionnaire, possibly project organisers informed themselves about CS Track on the project website and accessed the online survey by clicking on the link to the version in English they have found there.

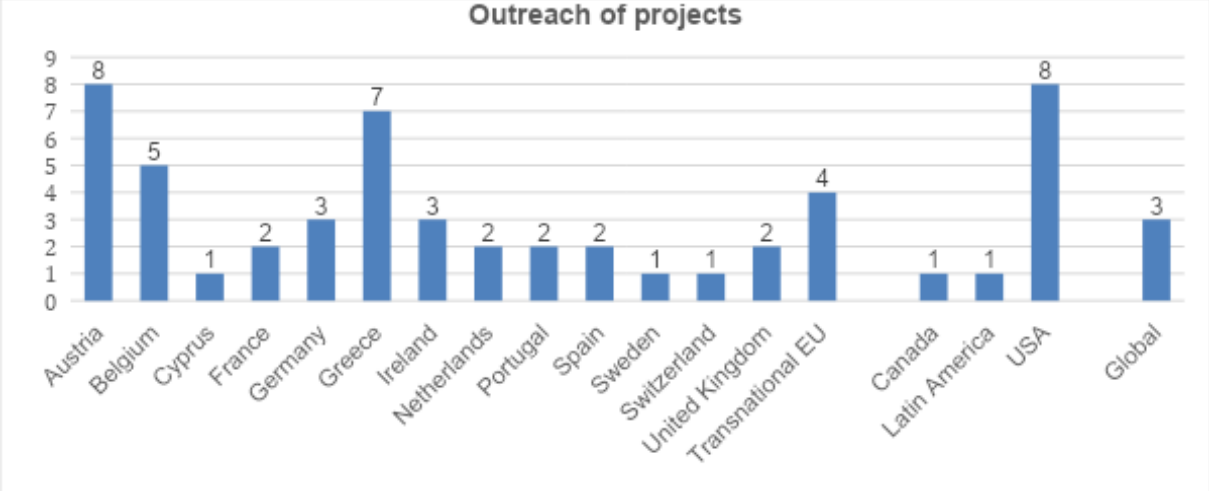
Geographic regions and sites

Where the responding citizen science projects take place is not clear at first sight. Email addresses or domain names do not always localise reliably the projects. It is necessary to visit the project website. Then one finds in most cases, but not in all, a clear-cut answer. Sometimes a differentiation between where the project is organised from and the regional outreach is necessary. One project, organised by a German research institute, is hosted on the Zooniverse platform. It was counted as "international". Another project is almost exclusively US-based but there is one UK site included on their map, thus the outreach is wider than national. If more than one European country is involved, it is counted as "European". In one case "European" means a British-Irish project. With global collaboration, platform localisation has become a little blurry, as the online tools allow for partnerships around the globe. Top-level domains (country code and generic top-level domains) according to websites and/or email addresses indicate the location of the project organisation in most but not all cases.

Country / Level Domain	.at	.be	.ch	.cy	.de	.dk	.es	.fr	.gr	.ie	.nl	.pt	.se	.uk	.io	.eu	.ca	.com	.edu	.net	.org	TOTAL
Austria	8																					**8**
Belgium		5																				**5**
Cyprus				1																		**1**
Denmark						1																**1**
France								2														**2**
Germany					2														1		1	**4**
Greece									6									1				**7**
Ireland										3												**3**
Netherlands											2											**2**
Portugal												2										**2**
Spain							2															**2**
Sweden													1									**1**
Switzerland			1																			**1**
UK														2							1	**3**
European					2											1						**3**

Canada																1					**1**
Chile																				1	**1**
USA														1			3	1	1	3	**9**
TOTAL	**8**	**5**	**1**	**1**	**4**	**1**	**2**	**2**	**6**	**3**	**2**	**2**	**1**	**2**	**1**	**1**	**1**	**4**	**2**	**1**	**6**	**56**

The location of the project organisation and geographic outreach of projects were researched separately.



Analysis of the responses

The first two questions asked for an identification of the project someone responded to and an email contact address.

Summary of goals (Question 3)

The projects responding project organisers referred to show a broad range of activities, settings, goals, involvement intensities, etc. As expected, several projects had objectives related to biodiversity, the environment and/or a combination of both, but in different ways.

Disciplines in the diverse teams (Question 4)

The questionnaire asked to type in the academic disciplines available in the organisation team and no limitation of number was given. It was an open question so that respondents could use the terminologies they were familiar with. The responses were manually allocated to the fields of science and technology according to the Frascati Manual.

The question was directed to compare the research areas in which the project operated with the academic backgrounds in the diverse teams. How far they match is one of the most difficult questions to answer. It would be necessary to do in-depth research about what are the exact contents of a study in a certain country, region or university. Disciplines are far from being universal. There are widely recognised categorisations to group the hundreds of existing and often not comparable disciplines, which come partially to different results. The **Frascati Manual (FOS)** matches in several respects with **Web of Science (WoS).** The latter is used for the CS Track database, a collection mainly of citizen science projects visible on the web in the European Union and Associated Countries to document citizen science projects (see chapter 6.2 of Strähle, Urban et al. (2022)). The Frascati Manual was used as a means to compare the overview of named disciplines to Web of Science. This allows a comparison with the scheme used in the CS Track project database. Matching the disciplines to either with the WoS classification or of the FOS proved difficult in many cases and left a lot of room for interpretation. The difficulties of classifying academic disciplines is also shown by discrepancies. Neither WoS nor FOS can do justice to interdisciplinary fields. For example, anthropology belongs to the social sciences in the FOS manual but in the WoS it is placed in "Life Sciences & Bio-medicine".

Of the 56 analysed responses, 53 gave an answer to the question while 3 respondents skipped it. Those who answered named 1 to 5 disciplines for their project team which resulted in a total of 162 entries. Each of these entries was manually allocated to both Web of Science (WoS) subareas and the Frascati Manual (FOS) classifications as far this was possible.

Major fields of science and technology per project according to Web of Science (WoS)

WoS (CS Track)	Total number of disciplines named by respondents	Number of projects in which discipline is named at least once

Acoustics	1	1
Anthropology	4	4
Archaeology	1	1
Architecture	1	1
Arts & Humanities Other Topics	2	1
Astronomy & Astrophysics	2	2
Behavioral Sciences	1	1
Biodiversity & Conservation	5	5
Biology	4	4
Business & Economics	8	7
Chemistry	4	4
Communication	1	1
Computer Science	14	12
Construction & Building Technology	1	1
Education & Educational Research	3	3
Energy & Fuels	1	1
Engineering	2	2
Environmental Sciences & Ecology	19	17
General & Internal Medicine	2	2
Geochemistry & Geophysics	2	2
Geography	1	1
Geology	3	3
Government & Law	1	1
Health Care Sciences & Services	2	1
History	3	3
Linguistics	2	2
Literature	3	2
Marine & Freshwater Biology	5	5
Medical Informatics	1	1
Meteorology & Atmospheric Sciences	5	4
Nutrition & Dietetics	1	1
Oceanography	1	1
Operations Research & Management Science	1	1
Physical Geography	5	4
Physics	5	5
Plant Sciences	3	2
Political science	4	4
Psychiatry	1	1
Psychology	1	1
Public, Environmental & Occupational Health	2	2
Remote Sensing	2	2
Science & Technology Other Topics	2	2
Social Sciences Other Topics	1	1
Social Work	1	1
Sociology	6	6

Transportation	1	1
uncategorised	10	8
Unspecified Arts & Humanities	1	1
Unspecified Social Sciences	2	2
Urban Studies	1	1
Veterinary Sciences	1	1
Zoology	6	6
TOTAL	**162**	**149**

Major fields of science and technology per project according to the Revised Fields of Science and Technology (FOS) classification in the Frascati Manual

FOS	Total number of disciplines named by respondents	Number of projects in which the discipline is named at least once
Art (arts, history of arts, performing arts, music)	1	1
Biological sciences	34	24
Chemical engineering	1	1
Chemical sciences	3	3
Civil Engineering	2	2
Clinical medicine	1	1
Computer and information sciences	15	12
Earth and related Environmental sciences	19	13
Economics and Business	8	6
Educational sciences	3	3
Electrical engineering, Electronic engineering, Information engineering	1	1
Environmental engineering	4	3
Health sciences	6	5
History and archaeology	4	3
Languages and literature	5	4
Law	1	1
Mechanical engineering	2	2
Media and communications	1	1
Other engineering and technologies	1	1
Other humanities	1	1
Other medical sciences	1	1
Other social sciences	2	2
Physical sciences	8	6
Political science	4	4
Psychology	1	1
Social and economic geography	4	3
Sociology	11	10
uncategorised	13	10

Unspecified Humanities	1	1
Unspecified Social sciences	2	2
Veterinary science	2	1
TOTAL	**162**	**129**

Comparison of categorising according to WoS and FOS

	Research Areas FOS							
Over-arching areas WoS	*Agri-cultural Sciences*	*Engineer-ing and Techno-logy*	*Humanities*	*Medical & Health Sciences*	*Natural Sciences*	*Social Sciences*	*Uncate-gorised*	*TOTAL*
Arts & Humanities			9		1			**10**
Life Sciences & Bio-medicine	2			8	40	5	2	**57**
Physical Sciences		2			25			**27**
Social Sciences			3			30		**33**
Techno-logy		9			13	2	1	**25**
Uncate-gorised							10	**10**
TOTAL	**2**	**11**	**12**	**8**	**79**	**37**	**13**	**162**

For the following fields of science and technology no equivalent could be found in the Frascati Manual:

Coastal environmental monitoring / environmental protection / organizational Management / cognitive science and creativity / computational thinking and learning sciences / engineering / environmental education / natural resource management / participation / product design / research principles / web design / website programming

For the following fields of science and technology no equivalent could be found in Web of Science:

Cognitive science and creativity / computational thinking and learning sciences / engineering / environmental education / natural resource management / participation / product design / research principles / web design / website programming

Three indications on social sciences and the humanities were too unspecific to be categorised in a subarea of WoS or FOS, but they were included in main areas.

According to the 6 main categories of the Revised Field of Science and Technology classification in the Frascati Manual, natural scientists were most strongly represented.

A first check does not show an obvious mismatch of disciplines in the organisation teams and the WoS-based classification of research areas of the projects named in the CS Track project database. But this may be more easily answered for disciplines in the technical and natural sciences than for the social sciences and the humanities. For the latter, there may be major differences between science traditions and curricula in the different parts of the world. Furthermore, some of the mentioned methods, like participatory action research, may in some contexts need some group dynamical and almost therapeutic psychological knowledge, which cannot be followed up with a questionnaire.

Number of participants roughly estimated by respondents (Questions 5 & 6)

Cumulative number: Estimated number since project start (Question 5)

Number of participants since project start (cumulative number)	
≤ 20	8
21 – 100	15
101 – 1000	14
> 1000	19
TOTAL	**56**

All respondents gave an estimate of the numbers of participants.

Present number: Estimated number at the time of the response (Question 6) or when the project ended.

Number of participants at present or end of the project	
≤ 20	19
21 – 100	10
101 – 1000	18
> 1000	8
TOTAL	**55**

55 respondents gave an answer to this question.

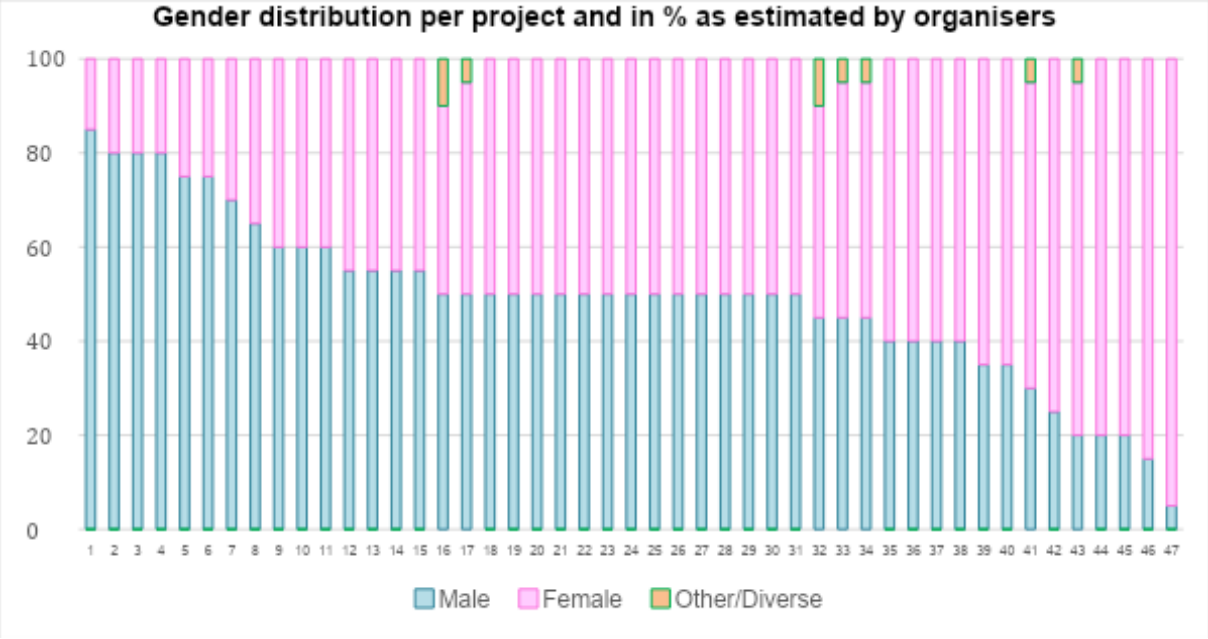
Comparison cumulative number and present number of participants as estimated by organisers

	Present number				
	≤ 20	21 – 100	101 – 1000	> 1000	TOTAL
Cumulative number:					
≤ 20	7		1		**8**
21 – 100	9	6			**15**
101 – 1000	2	3	9		**14**
> 1000	1	1	8	8	**18**
TOTAL	**19**	**10**	**18**	**8**	***55***

Asking for two estimates gives a safer impression of the project size than asking for only one as projects can change considerably over time. The answers mirror the very broad range of the size of such projects that one can find in literature as well.

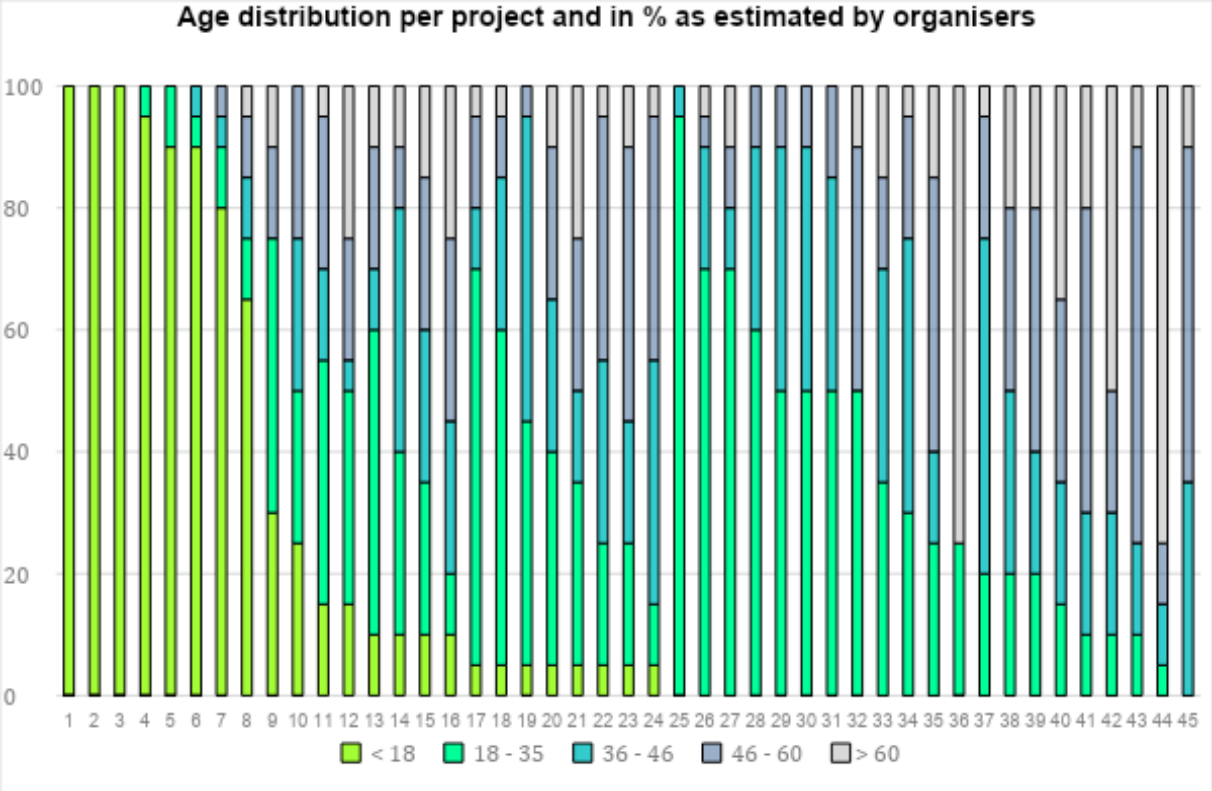
Estimations of gender distribution (Question 7)

47 respondents gave a rough estimation of how many percent of the participants would be male, female or of diverse/other gender. The chart below shows that 13 respondents believe in a female majority of at least 60% of all participants, while 11 respondents thought that 60% or more were males. A few more see a slightly higher proportion of males. The 24 respondents who estimated an exactly equal distribution of males and females may have expressed that they were at least not aware of a gender imbalance. Because gender was one of the slider questions, they had to finish the answer once they started to move the slider. While most of the respondents indicated a rough gender balance, there are a few projects that involve mostly men or women. The percentage of diverse/other gendered participants was estimated in 7 cases.



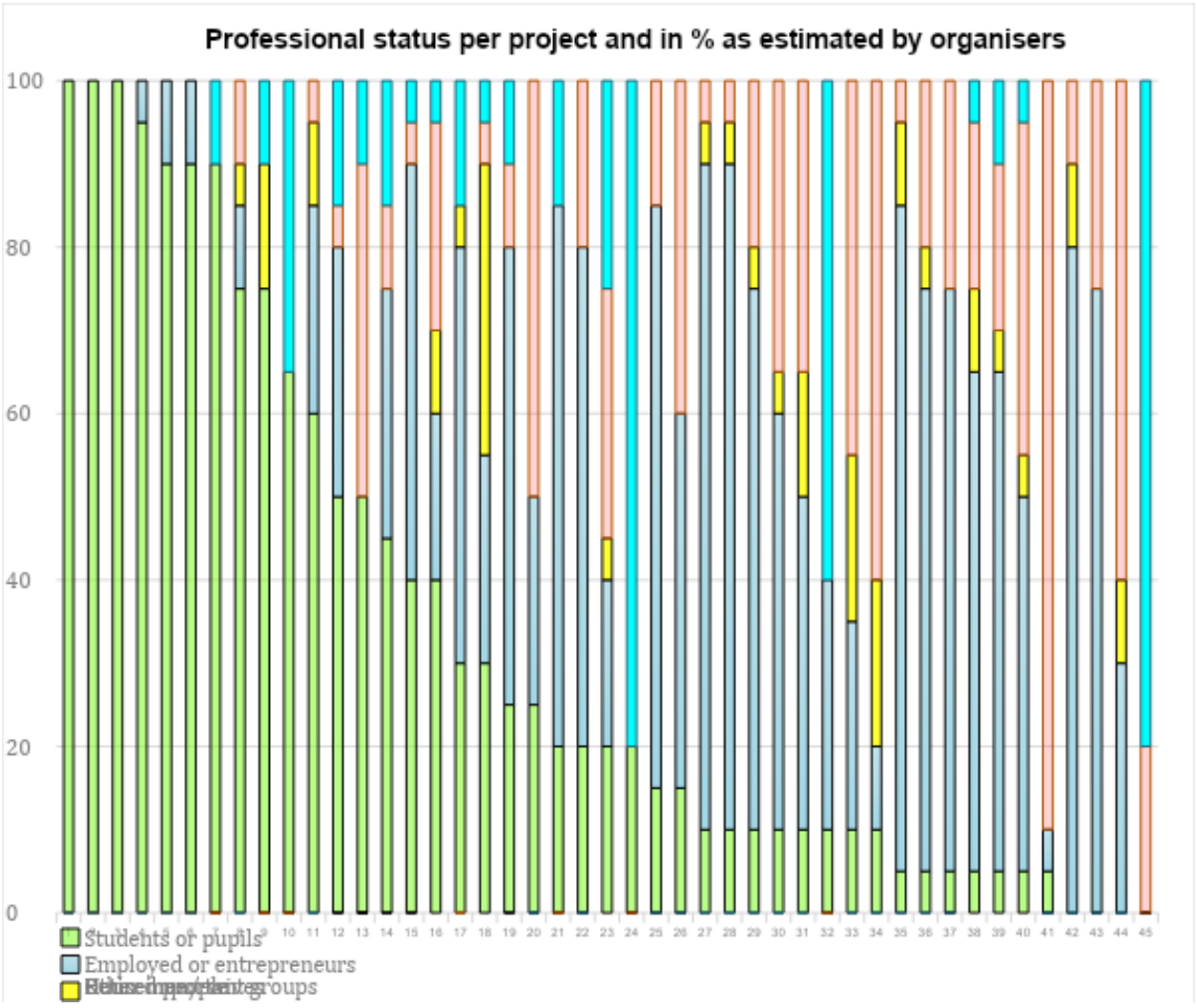
Estimations of age distribution (Question 8)

45 responses contain an answer to this question. There are only three projects which indicate 100% for one age group, namely below 18 years. The youngest age group is also highly dominant in 4 additional projects (80% or more of the participants are estimated as being younger than 18 years old) and moderately dominant in another project (65%). The second youngest group (18 – 35 years) is seen as very present, too: They are estimated between 65% and 95% of the participants by 4 organisers. At the other side of the spectrum, we find 2 respondents who estimate that 75% of their projects' participants are older than 60 years.



Estimations of professional status of participants (Question 9)

45 respondents gave feedback to this question, and it most likely can only be answered if a project is targeted to a specific group (i. e. pupils, students) or if the project is small enough that people know each other quite well. In larger projects it would be necessary to ask such information from the participants which might go against privacy rules and create a barrier for participation. For example, the same three organisers of school projects, who place 100% of the participants in the youngest age group, indicate that they are students or pupils.



Response patterns

Project size and estimates of participants' characteristics

Did organisers of "smaller" projects indicate more characteristics of participants than those of projects with large number of participants?

Rough estimations by project organisers: cumulative participant number & gender distribution

	Gender distribution		
Participant number since project start	no answer	Estimation	TOTAL
≤ 20	1	7	**8**
21 – 100	2	13	**15**
101 – 1000	4	10	**14**
> 1000	2	17	**19**

TOTAL	**9**	**47**	**56**

In this small survey we see no stark difference between estimations of gender distribution and rough cumulative number of participants.

Rough estimations by project organisers: cumulative participant number & age distribution

	Age distribution		
Participant number since project start	no answer	Estimation	TOTAL
≤ 20	1	7	**8**
21 – 100	1	14	**15**
101 – 1000	6	8	**14**
> 1000	3	16	**19**
TOTAL	**11**	**45**	**56**

Rough estimations by project organisers: cumulative participant number & distribution of professional status

	Distribution of professional status		
Participant number since project start	no answer	Estimation	TOTAL
≤20	1	7	**8**
21 – 100	1	14	**15**
101 – 1000	6	8	**14**
> 1000	3	16	**19**
TOTAL	**11**	**45**	**56**

The authors had expected to see a stronger connection between the number of participants and organisers' tendency to give rough estimations of their characteristics. As expected, almost all responding organisers of projects with manageable numbers of participants (less than 21) answered the three questions. It is plausible that in smaller project those involved know each other personally. But we also see a surprisingly high number of estimations from very large projects (more than 1000 participants) who made a rough estimation. Hence it would be interesting how project organisers of larger projects retrieve information about participants, for example, if they send out questionnaires or if some questions have to be answered during the registration.

4 Technical aspects

Empty, completed, discarded and test responses saved by the survey software

Overview of all "responses" saved by the survey software

All entries	Number
Empty responses	841
Tests	30
Discarded	3
Used for analysis	56
TOTAL	**930**

Empty responses

Until the last test on 18th January 2021 there were 841 entries that were empty. There are several possibilities why there are empty answers:

- The answers were produced automatically by crawlers and bots
- Interested persons clicked on the questionnaire, but decided that they could not answer the questions or did not want to answer them.
- Technical problems like a time-out.

It is not possible to find out which combination of events yielded the high number of empty answers.

Some efforts were made to find out what could be the cause of the high number of answers that lacked any information at all, how they came about. Interestingly, the empty answers show a higher variety of languages in which respondents started the survey than the completed answers that could be used for analysis.

Start language versions in empty, completed, tests and excluded responses

Start language	Empty	Completed	Excluded	Test	TOTAL
Arabic	3			6	9
German	51	6		7	64
Greek	5			1	6
English	763	50	3	9	825
Spanish	4			2	6
Finnish	10			2	12

French	2				2
Hebrew	2			2	4
Dutch	1			1	2
TOTAL	**841**	**56**	**3**	**30**	**930**

Discarded responses

Of 59 completed responses three were excluded. One project had not started yet, another respondent answered twice and a third referred to the theoretical work of an academic on involving the public. 56 responses remained for descriptive analysis.

Technical difficulties with the survey software

The authors experienced several unexpected technical difficulties. Why these occurred could not be conclusively clarified.

* Frequently, changes to the questionnaire could not be saved, even if the authors or the host of the survey software tried several times. (This is the reason why the survey had to be launched with slightly different formats in 3 slider questions).

* The software did not allow to make copies of similar parts of the questionnaire, so they had to be set up one by one.

* In general, the survey software was slow.

* The feature "quick translation" allowed only for a limited number of languages. Further languages could have been added by revising the HTML and PHP codes, but the software seemed already quite overloaded, hence the risk was not taken.

* After the survey had been launched, new issues appeared: The survey contained altogether three questions on age groups, gender distribution and professional status which could be answered by moving sliders.

To give an example: For a very rough estimate of the gender distribution, respondents could move one slider each for "male", "female" and "diverse/other". Answering the questions with sliders was not obligatory. If a respondent skipped one of the three questions without moving a single slider, the answers could be submitted without an issue. However, if a respondent moved a slider, then decided not to answer the question - e.g., because s/he was not sure if the answer is correct - and moved the slider back to 0, the question became an obligatory one. Then all the answers could not be submitted without receiving an error messaging asking to answer this particular question. (According to the test results, the software would have saved the answers in the questionnaire nonetheless, but the respondent could not know this and may have believed they could not submit without completing the respective question).

Plausibility checks were made. For example, if someone wanted to express that s/he really has no idea about the age distribution or professional status, s/he had several ways to signal this, for example, by moving the slider to 100% in one age or

professional status group. The authors have found only three answers with 100% for one age group and 100% for one professional group, namely the youngest (< 18 years) and "students or pupils". The three projects were checked and that they are school projects verifies the answers.

* To make sure that the survey was running without error, the authors made numerous test submissions before launching it. When hundreds of empty answers came back after each dissemination campaign, 30 additional tests were carried out to find out how the software reacted to different respondent behaviours, e. g. how the software would save responses brought about by different behaviours. Because the survey software was set up not to save the date of a response, some tests functioned as time markers. According to the try-outs, the software would save all answers in a response even if a respondent received an error message but did not correct the respective answer. On the other hand, if submitting answers was followed by a time-out, only a blank questionnaire was saved that contained no information who had tried to submit it. It is noticeable that apparently no respondent lost patience with the questionnaire after an error message had been received: No unfinished questionnaire with answers were saved.

It is possible that the technical problems were at least partially due to the pandemic: Servers could have been overloaded, as during the lockdowns many activities had been shifted onto the internet and many meetings took place as video-conferences. In spite of these issues, the Linux-based survey software is still one of the most trusted survey tools among people who are worried about data and privacy protection.

It was decided to pause the promotion of the survey until another CS Track survey was finished, the survey in Work Package 4 mentioned at the beginning of this report. During this temporary interruption of the promotion campaign, it was tried to solve some of the technical difficulties mentioned above.

5 Conclusions

The replies to the very short questionnaire confirm that there is a broad diversity of projects and initiatives that can fall under this blurry term, although they may not even cover the whole spectrum of what is called citizen science.

Even for the time of promoting the survey (before Christmas), the rate of respondents seems low. This might be an indication that many projects do not know very much about the participants, their characteristics or even their number (or do not want to admit to it) and refrain from answering. This is in line with the authors' research so far. Literature reviews (Strähle, Urban et al., 2021) showed a lack of knowledge concerning those who take on the role of "citizen scientists" in a citizen science project. Whether the individual participants should be known to the organisers cannot be answered in general terms. The answer to this question depends on the specific conditions under which a citizen science activity takes place, among other things, its setting and goals. The Activities & Dimension Grid of Citizen Science (see Strähle & Urban (2021), page 96 - 125) shows how complex these conditions can be. Also, the more influence a single participant has on the research carried out in a project or on research policies (e.g., when s/he participates in deliberations on policy issues), the more important and/or sensitive the question of her or his identity becomes. The trade-off between open and anonymous participation has to be regarded separately for different activities and dimensions. The impact this has on benefits, caveats, the creation of enablers, barriers and (dis-)incentives for all types of participants of CS warrants research of its own.

Characteristics and backgrounds of "citizen scientists" constitute probably the biggest black box in the field although "citizens" are an essential part of the whole concept. Assessing some of the claimed benefits, caveats, perceived barriers, enablers, incentives and disincentives of, respectively for, citizen science depends on sufficient knowledge about who the "citizen scientists" actually are. In view of the many benefits that several scholars, practitioners, policy makers and other people claim citizen science brings with it, this would make some of them unfounded if not even implausible. Although there are examples of benefits a citizen science approach had for science (Strähle, Urban et al., 2021), it cannot be said with sufficient certainty that, in general, participation in citizen science activities creates a benefit in terms of science literacy and educational value for volunteers. For example, claims that citizen science can democratise science, improve the education of lay participants and raise awareness on science depend directly on who can be involved and how. Some benefits will suffer less from not knowing who the participants are than others. For example, crowdsourcing of activities in knowledge production in which results are non-ideologic, provable and/or if results of different persons can be compared to each other, may not require to know who actually takes part.

An attempt was made to investigate – in cases where academics were among the organisers - how far their expertises match(ed) the research areas of the projects. This proved exceptionally tricky because there exists no classification scheme which mirrors the broad variety of academic educations in different regions. Depending on science cultures and research traditions, seemingly similar academic studies can follow quite different curricula. Language is an issue, too. To find out if the academic competences are in line with the scientific and educational activities of a project would require an assessment scheme for the acquired competences of academics organising citizen science activities. The information what somebody has studied could be complemented with more concrete ideas what s/he has actually learned.

Hence it could be done on a very rough level only at best. This is an important question because it is often demanded that academics steering citizen science projects or activities have all necessary competences in the field. Also, it would be the absolute minimum requirement for the frequent claim that citizen science would improve "scientific literacy". Formal higher education in the research fields citizen science activities take place is part of the concept as it serves as a core characteristic that separates the involved "scientist" from the "citizen scientist".

6 References

Eitzel, M. V. et al. (2017). Citizen Science Terminology Matters: Exploring Key Terms. Citizen Science: Theory and Practice, 2(1), 1. https://doi.org/10.5334/cstp.96.

European Commission (2018). Horizon 2020: Work Programme 2018–2020: 16. Science with and for Society. https://ec.europa.eu/research/participants/data/ref/h2020/wp/2018-2020/main/h2020-wp1820-swfs_en.pdf.

Heigl, F. & Dörler, D. (2017). Public participation: Time for a definition of citizen science. Nature 551: 168. https://doi.org/10.1038/d41586-017-05745-8.

Kullenberg C. & Kasperowski D. (2016). What Is Citizen Science? – A Scientometric Meta-Analysis. PLoS ONE 11(1): e0147152. doi:10.1371/journal.pone.0147152.

OECD Working Party of National Experts on Science and Technology Indicators (2007). Revised Field of Science and Technology (FoS) Classification in the Frascati Manual. DSTI/EAS/STP/NESTI(2006)19/FINAL. https://www.oecd.org/science/inno/38235147.pdf.

Riesch, H. & Potter, C. (2014). Citizen science as seen by scientists: Methodological, epistemological and ethical dimensions. Public Understanding of Science 23(1) 107-120.

Strähle, M., Urban, C. et al. (2021). Framework Conceptual Model D1.1. Zenodo. https://doi.org/10.5281/zenodo.5589618.

Strähle, M., Urban, C. et al. (2022). Conceptual Framework for Analytics Tools D1.2. Zenodo. https://doi.org/10.5281/zenodo.6045639.

7 Annex

Questionnaire

Date submitted Last page Start language Seed

1. What project do you refer to in the following?

[Name of project]

[Website]

2. Project contact information [Email address of the project]

3. Please write just a few words about the main objectives of your project.

4. Which academic disciplines are available in the organisation team?

[Discipline 1]

[Discipline 2]

[Discipline 3]

[Discipline 4]

[Comments]

Please read before you proceed: As participants, in this context, we understand unpaid contributors who are not part of the project organization team and not graduates in the disciplines relevant to your project.

5. Please give us a rough picture of how many participants have been involved in your project since its very beginning.

6. Please indicate how many participants are presently active in your project. For finished/pausing projects: Use the latest number of participants instead.

In the following we try to get a picture of who are the participants in your project. Please give us an idea of percentages by moving the sliders. If you cannot or do not want to give even a very rough estimate, just skip to next question.

7. Can you give a very rough estimate of the gender distribution among the participants in your project? Please move the slider until it feels right.

[male]

[female]

[diverse/other]

8. Please give us a very rough overview of the age distribution of participants. Just move the slider bar until it feels right.

[< 18 years]

[18 - 35 years]

[36 - 45 years]

[46 - 60 years]

[> 60 years]

9. Please give us a very rough estimate of the professional status of the participants. Just move the slider until it feels right.

[Students or pupils]

[Employed or entrepreneurs]

[Housemen/-wives]

[Retired people]

[Other important groups]

Thank you very much for taking your time to answer this questionnaire. Please find more information on CS Track here.

Note: All questions but Question 1 were optional questions. The obligatory question 1 served to identify the project without doubt.