Horizon 2020 / Science with and for Society Programme
Grant agreement number: 872522

Investigating the potential of citizen science to respond to emerging challenges. A COVID-19 review and case study analysis

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	Investigating the potential of citizen science to respond to emerging challenges. A COVID-19 review and case study analysis
Actual delivery	November 2021
Author	Lead author: Anne Turbe, Ecoscope Contributors: Yaela Golumbic, Tslil Farchi and Reuma De-Groot.
Executive summary	Anne Turbe, Tel: +(972) 54 3106710, aturbe@gmail.com
Acknowledgements	All contributors to the structured interviews for the case studies have made a major contribution to this project. We are immensely appreciative of their time and honest reflections, which have added considerably to this report.
Citation	Turbe, A., Farchi, T., De Groot, R. and Golumbic, Y. (2021) Investigating the potential of citizen science to respond to emerging challenges: a review and case study analysis. Final report for CS Track (H2020 funded project).

This page is left intentionally blank

Abstract

Citizen science offers a huge potential to complement official responses to the Coronavirus pandemic, both in terms of facilitating scientific advances and of improving public engagement.

Here we provide an analysis of Covid-19 citizen science projects using a two-phase research approach. Website content analysis was used to understand the outstanding characteristics of citizen science projects that endeavoured to respond to the pandemic, in terms of their geographical distribution, aims, design and characteristics of citizen engagement. Followed by detailed case studies of seven citizen science initiatives which provided in-depth understanding of the development of projects, and practices used across a range of approaches.

The website review revealed that the covid-19 citizen science landscape was dominated by contributory crowd-sourcing projects, mainly related to health sciences. Most of the projects were purpose-developed by scientists to respond to the global Covid-19 crisis, and leveraged basic information technologies (internet, smartphones) to collect self-reported data on people's symptoms, behaviour or well-being from questionnaires. More than half of all the projects were global in scope, but projects were found at all geographical scales. Often, the effort required by participants was non-negligible, involving repeated inputs at frequent intervals, though of limited duration. The review also suggested data accessibility and validation procedures were limited.

The case studies revealed the importance of preparedness to respond to new challenges, building on existing experience, collaborations and modular software infrastructure. It was interesting to notice that simple designs and partnerships with NGOs seemed to improve the likelihood that projects would reach high levels of engagement. Additionally, all case studies highlight the importance of regular and honest feedback, to secure the trust and engagement of the participants. Recommendations from several projects include not being afraid to plan for success, thinking big, open and collaboratively when designing a project.

This page is left intentionally blank

Table of Contents

This page is left intentionally blank

1 Introduction

The coronavirus pandemic has shed light on the gap between the knowledge we need and the knowledge we have, emphasising the rapidity and flexibility with which we need to create knowledge to respond to emerging challenges.

Citizen science offers a huge potential to complement official responses to the Coronavirus pandemic, by engaging and activating the public to bring in more data, and to increase the utility and impact of information. In citizen science, the public participates voluntarily in the scientific process, addressing real-world problems to contribute in collecting or analysing data, conducting experiments or solving complex problems, and possibly designing studies (Bonney et al., 2009; Shirk et al., 2012).

Recently the use of citizen science has been booming as a result of the need for mass data, growing confidence and valuation of the input of lay people and technological development (Roy et al., 2012; Socientize consortium, 2013). Until recently, citizen science was carried out mostly in the fields of environmental sciences and remained rare in public health (Wiggins and Wilbanks, 2019). Yet, citizen science has much to offer as a way to collect data, in particular in the field of infectious diseases. Lay people's input can help detect more rapidly the first signs of outbreaks, or identify sets of symptoms. It can also inform policy-makers about citizens perceptions and views, in response to certain policies (Turbé et al., 2019). But most importantly, it has the potential to increase health literacy, empower communities, create new attitudes and values (Den Broeder et al., 2016).

Few research domains are as meaningful to the public as health and medicine, and crowdsourced citizen science in the health domain has exhibited exponential growth over the past decade (English et al. 2018). Perhaps due to the ease of self-tracking, documenting and experimenting with interventions, the role of participants is expanding more quickly in health sciences compared to other citizen science areas (Swan 2012). Crowdsourced tasks can be quite wide-ranging, but mainly rely on the use of digital tools and platforms for use by the public in undertaking specific tasks. Volunteers may classify images of bacteria, or help solve protein folding problems with the use of gamification. Crowdsourced health research projects also increasingly target health conditions, in which participants undertake self-reporting, responding to surveys, using mobile app or wearable sensors, leading for instance to the capture of qualitative and quantitative epidemiological information. Crowdsourcing has many advantages, in terms of improved cost, speed, flexibility and scalability of data. It allows for access to a large pool of participants, saves time to collect data, lowers costs and speeds up innovations (Créquit et al., 2018). For instance, the use of daily self-reported symptoms has emerged as tremendously useful means to track disease spread, predict outbreak locations, guide population measures and help in the allocation of healthcare resources, to complement official statistics (Segal et al., 2020). However, some points that remain controversial include the validity of crowdsourcing studies results and ethical aspects, related to data privacy issues notably. Additionally, crowdsourced health studies where the participants are only the subjects of enquiry and do not contribute to the scientific process in an expanded role, and receive little support or benefit, in terms of science literacy or data enquiry, may not always be considered citizen science (Wiggins and Wilbanks, 2019). Yet, there is a developing notion that sharing health information could be considered acts of citizenship.

In this context, it is interesting to investigate how citizen science can help respond to emerging health research challenges, through the example of the recent Covid-19 pandemic, both in terms of facilitating scientific advances, and of improving public engagement. Demonstrating flexibility and innovation, some existing citizen science and crowdsourcing projects have pivoted to respond to the Covid-19 crisis, whereas others have emerged from scratch in a very short time. Outstanding questions involve the flexibility of citizen science to adapt to new research questions and effectively engage the public, and the design of project that may be most conducive to this. Additionally, issues surrounding data quality and usefulness of results, robustness of research design and how participation may impact, positively or negatively, scientific literacy or trust in science are also important to investigate, to understand how to better foster future useful models.

2 Methods

2.1 Identification and selection of citizen science projects

The selection of projects to be included in this report was conducted using two methods (see figure 1):

1. Searching the CS-track project database (metabase), using the keywords “Corona” and “Covid-19”. This process resulted in the identification of 13 citizen science projects.
2. Identifying projects related to citizen science and covid-19 from projects lists produced by citizen science associations and research institutes globally (e.g. CSA - https://www.citizenscience.org/covid-19, ACSA - https://theconversation.com/citizen-science-how-you-can-contribute-to-coronavirus-research-without-leaving-the-house-134238, Wilson institute - https://www.wilsoncenter.org/blog-post/citizen-science-and-covid-19-power-distanced-crowd). This process resulted in the identification of 22 citizen science projects, of which 5 overlapped with the projects found above.

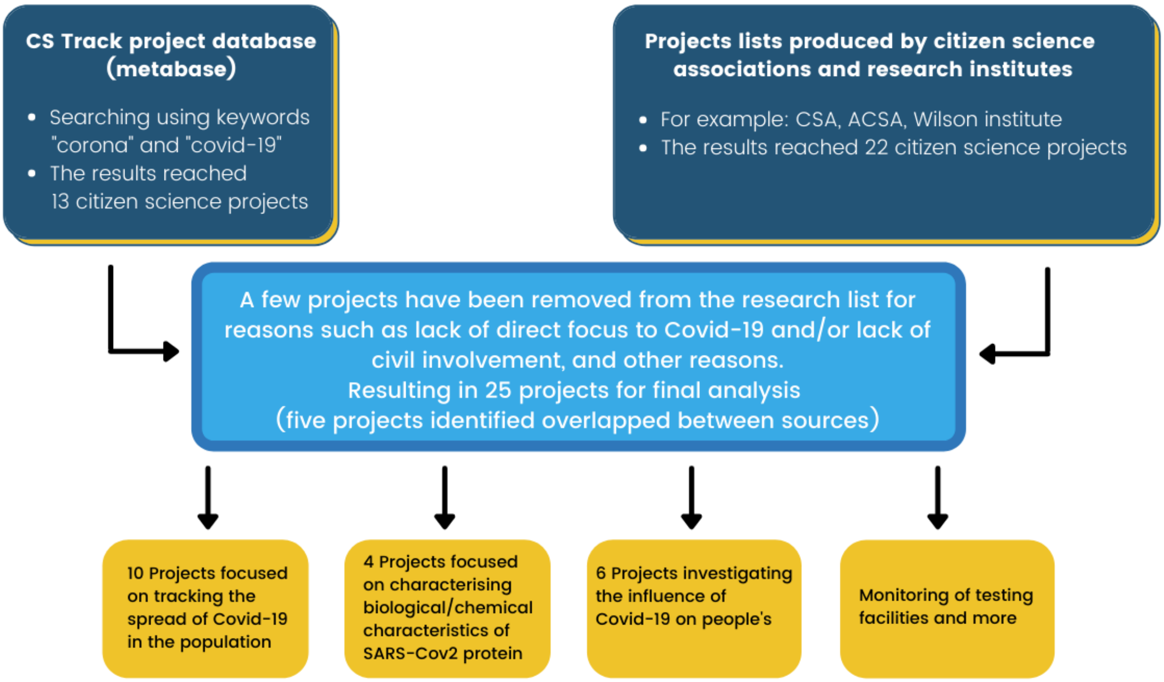


Figure 1. Selection and characterization of citizen science projects related to COVID-19 research

Following this initial compilation, projects that were not directly focused on Covid-19 (n=3 projects, two focused on influenza, one on environmental changes during Covid), projects that did not involve citizen participation (n=1), nor contribute to science (n=1, self-assessment tool to assess when to seek medical care) were removed. Therefore, a total of 25 projects remained for final analysis.

2.1.1 Website content analysis of Covid-19 citizen science projects

The following characteristics of citizen science projects were extracted from their descriptions and websites (Table 1), including aspects related to:

- Project focus and scope
    - Primary aim and novelty,
    - Geographical origin and scope,
    - Primary research domain and leadership,
- Design of citizen science
    - Level of engagement and tasks,
    - Characteristics of engagement,
    - Data required and routes to involvement,
    - Data accessibility and validation.

Attributes were further categorized into thematic groups and validated by two independent coders. The final categories are listed in Table 1 and further discussed and summarized below.

Table 1 - Characteristics of citizen-science projects

Attribute	Categories
Aim	Tracking spread of Covid-19 in the population SARS-Cov2 biological and chemical characteristics and treatments Influence of Covid-19 on communities Influence of Covid-19 on well-being Covid-19 risk factors Contact tracing Other
New project	Entirely new Adaptation/extension of existing project N/a
Geographical scope	Global Regional National Local N/a
Lead organisation region	North America (USA and Canada) Europe Rest of the world N/a
Primary research category (based on OECD classification)	Natural sciences Engineering and technology Medical and health sciences Social sciences Other N/a
Primary field of research (based on OECD classification)	Computer and information sciences Chemical sciences Biological sciences Basic medicine Clinical medicine Health sciences - Public and environmental health Health sciences - Infectious diseases Health sciences - Epidemiology Medical biotechnology Civil engineering Materials engineering Others N/a

Attribute	**Categories**
Lead organisation type	Academic Public body, hospital Non-governmental organisation Private commercial organisation N/a
Type of participant engagement *Shirk et al. 2012*	Contributory (Scientists design projects to which members of the public contribute data) Collaborative (Scientists design projects to which members of the public contribute data + refine design, analyse data, disseminate findings Co-created (Scientists and members of the public work together and participants are actively involved in most or all aspects of the project) Other
Activity Haklay et al. 2013 ; Turbe et al. 2019	Crowd-sourcing (soliciting Web-based contributions from a crowd, volunteered computing) Distributed intelligence (People solving tasks or interpreting data) Monitoring (Regular monitoring of environmental / health variables, with a strict protocol defined by scientists) Occasional reporting (Occasional/opportunistic records, following a protocol designed by scientists or participants) Civic science (Bottom-up science, initiated and driven by a group of participants who identify a problem) N/a
Task	Problem-solving Data processing Surveillance/Monitoring Survey Other Self-assessment (text) Self-tracking data N/a
Special skills	None Limited Advanced N/a
Targeted at specific audiences	Yes No N/a

Attribute	Categories
Frequency of involvement	Once off Daily Weekly Opportunistic N/a
Different types of questions asked	1 2 to 5 > 6 N/a
Technology used	Website Smartphone application Distributed computing Game playing Citizen sensing Website + smarthpone application N/a
Type of data received	Location Questionnaire scale/description Photo or other multimedia Other Location + text/score Location + picture Text/score + picture Location+text/score+picture N/a
Data available to view	Yes No
Type of outputs shared publicly	Maps Graphes Narrative Narrative + Graphes/Visual Maps + Graphes Maps + Graphes + Narrative Other N/a
Quality assurance explicit	Yes No N/a

2.2 Approach to case studies of Covid-19 citizen science projects

Detailed case studies of Covid-19 citizen science projects were used to provide additional in-depth information of how citizen science can help respond to emergencies. It also allowed to assess alignment with the overview analysis of projects based solely on their online content. Projects were selected from the list of Covid-19 citizen science projects, using a stratified approach to ensure that the case studies represented the breadth and diversity of projects identified. For example, the case studies were selected to ensure appropriate representation of the different categories of primary aims. Additionally, we aimed to include citizen science with coverage at various spatial scales, from the community to the global level. This led to the selection of a list of 11 projects, with a backup list comprising an additional nine projects. People with a lead role in the selected citizen science initiatives were contacted by email and invited for a semi-structured interview by video conference. In case of no response after a couple of reminders, additional projects were contacted from the backup list. Although an emphasis on EU projects was sought, the response rate to projects from North America and other parts of the world was higher, such that the case studies were slightly skewed to these parts of the world.

In total 16 case studies were invited, and seven were completed (Table 2). All interviews were conducted by Anne Turbe, between June and July 2021, and lasted about one hour each. A full protocol of the semi-structured interview, including template and schedule, is available in Appendix 1, as well as the informed consent form that interviewees were asked to sign prior to their interview. The interviews were recorded and their information was reported in a standard format, following the general structure of the interview template (see Appendix 1 for details). However, care was taken not to be too prescriptive, to allow the necessary flexibility to bring out individual points relevant to a particular case study. The case studies are provided in full in Section 4.

Table 2 - Characteristics of the seven citizen-science initiatives on Covid-19 which contributed case studies.

Title	**Primary aim**	**Activity**	**Start date**	**Current status**	**New**	**Region (lead organisation)**	**Field of research**	**Task**	**Frequency of engagent**
COVID-PHYM	SARS-Cov2 biological and chemical characteristics and treatments	Crowd-sourcing	January 2020	Ended	New	Europe	Biological sciences	Data processing	Opportu-nistic
Open Pandemics	SARS-Cov2 biological and chemical characteristics and treatments	Crowd-sourcing	May 2020	Active	New	North America	Computer and information sciences	Data processing	Opportu-nistic
Safecast	Monitoring material resources	Crowd-sourcing	end March 2020	Active (dormant)	Adapted	Rest of the world	Other	Monitoring	Opportu-nistic
Quantified Flu	Tracking spread of Covid-19 in the population	Crowd-sourcing	end March 2020	Active	New	Global	Other	Self-tracking data	Opportu-nistic
Covid Open Survey.	Tracking spread of Covid-19 in the population Influence of Covid-19 on well-being	Crowd-sourcing	end March 2020	Active (dormant)	New	Europe	Other	Survey	Daily
CovidWatcher	Influence of Covid-19 on communities	Crowd-sourcing	mid March 2020	Active	New	North America	Health sciences - Infectious diseases	Survey	Opportu-nistic
Outbreaks Near Me (formerly Covid Near You)	Tracking spread of Covid-19 in the population	Crowd-sourcing	March 2020	Active	Extended	North America	Health sciences - Epidemiology	Survey	Weekly

3 Characteristics of covid-19 citizen science projects based on website content analysis

In total, 25 citizen science projects focused on Covid-19 were identified and subjected to website content analysis.

3.1 Primary aim and novelty

Most projects focused either on tracking the spread of Covid-19 in the population, based on reported symptoms (n=10, 40%), or on investigating the influence of Covid-19 on people's well-being (n=6, 24%). Additionally, 16% of projects focused on characterising the biological or chemical characteristics of SARS-Cov2 protein, to support vaccine and treatment development. Other aims, such as contact-tracing, influence of Covid-19 on communities, monitoring of testing facilities, or identification of Covid-19 risk factors were less represented (n=1 project each).

Most projects were entirely new (72%), but seven projects consisted in extending previous activities of existing citizen science projects, in particular influenza monitoring projects.

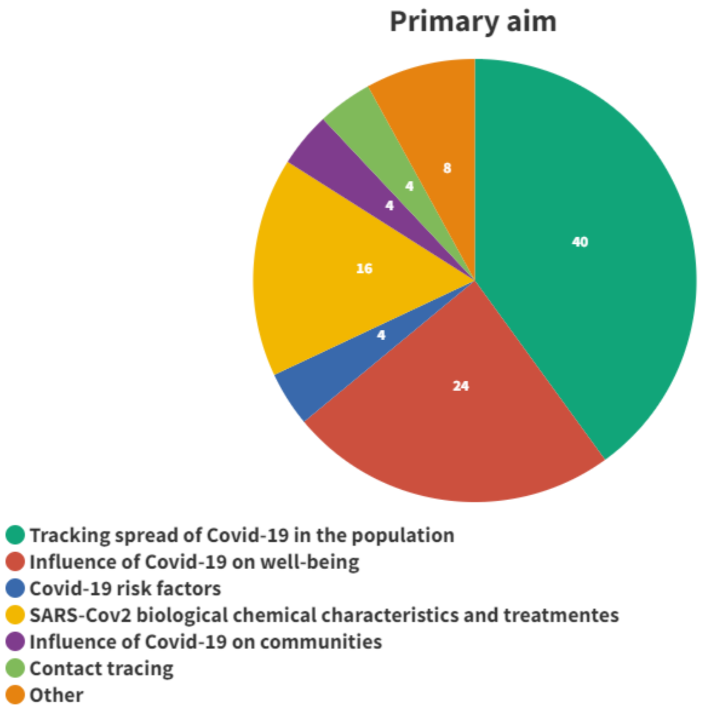
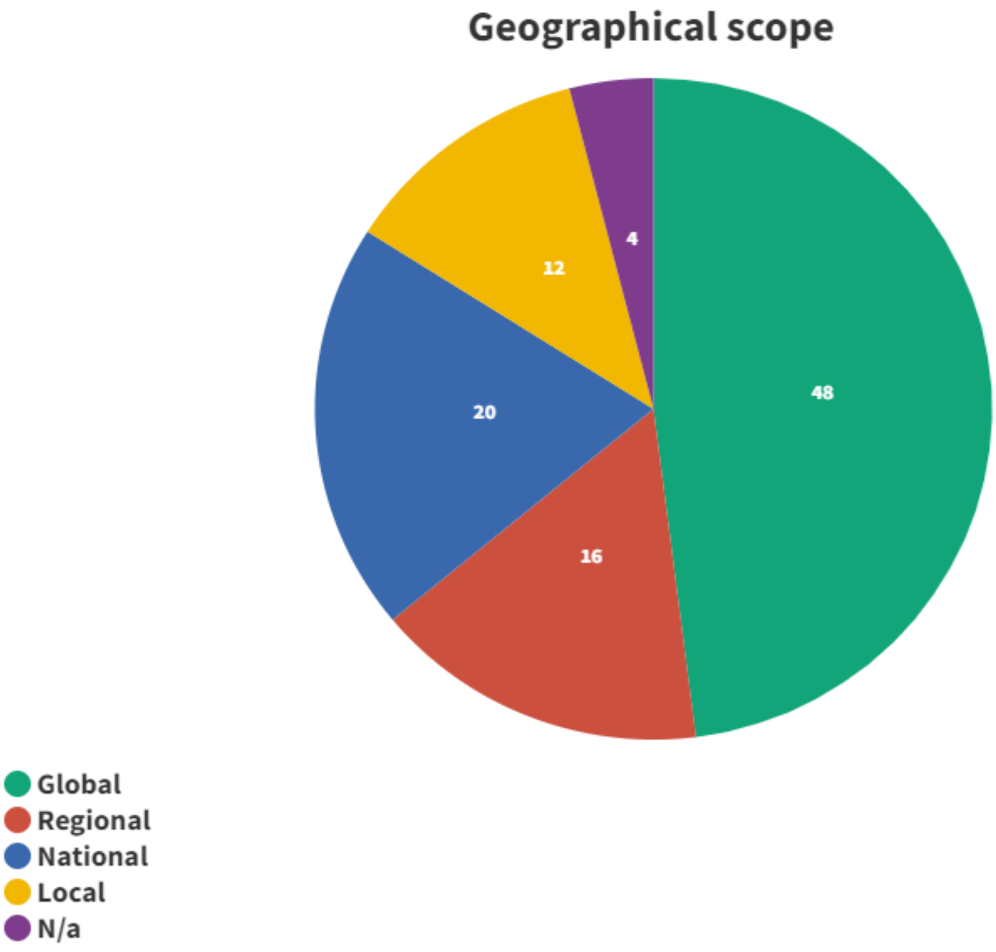


Figure 2. Primary research aim of Covid-19 citizen science project

3.2 Geographical origin and scope

Most projects were initiated either in Europe (44%) or in Northern America (36%), with only three of projects originating from other parts of the world (Australia, Singapore, Israel). Major activities across Europe were recorded from the UK (n=4) and Denmark (n=3), whereas all projects from Northern America were from the USA (n=9).

The majority of projects had a supra-national scope (n=17), with 48% of projects having a global scope and 16% of projects a regional scope (continental or several countries). A fifth of projects were focused on the national level, with a minority of additional projects relevant at local scale (12%).



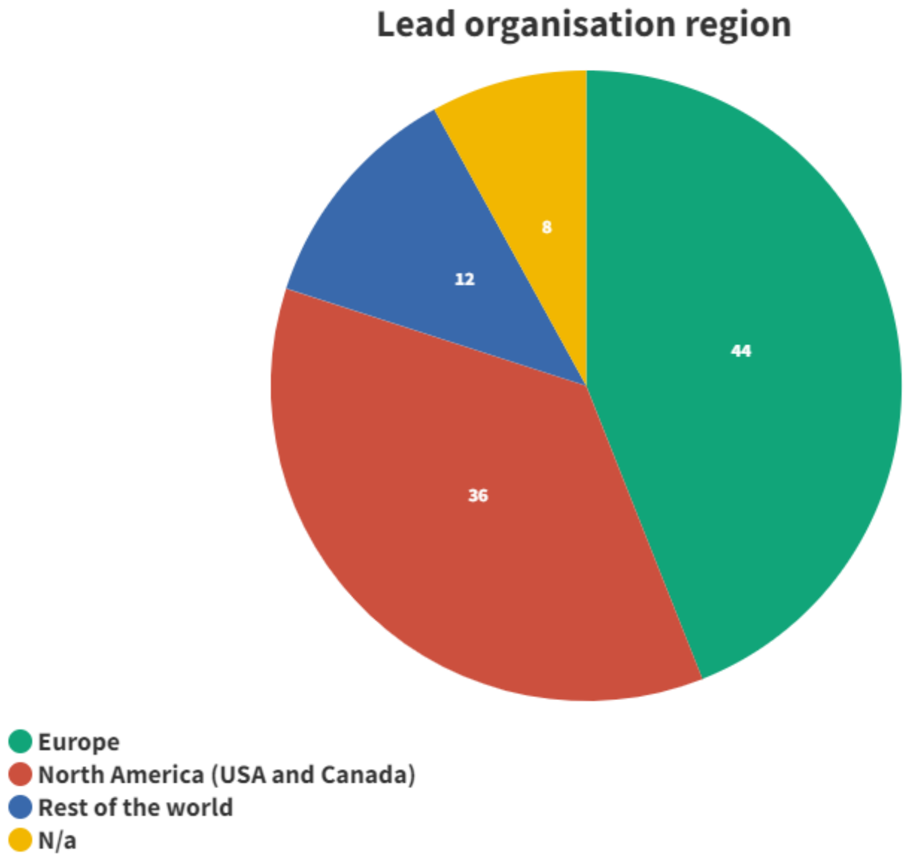


Figure 3. Geographical scope and origin of Covid-19 citizen science projects

3.3 Research field and leadership

The projects were usually interdisciplinary, and covered engineering, social sciences, natural sciences and medical sciences. However, there was a clear focus on projects originating within health sciences (n=11), in particular public and environmental health (n=7). In seven cases, the projects were truly interdisciplinary, bringing together scientists or volunteers from different fields. More than half of the projects were coordinated by academics (56%), followed by public bodies, typically a national health agency or hospital (12%) and non-governmental organisations (12%). Two projects were initiated by private companies.

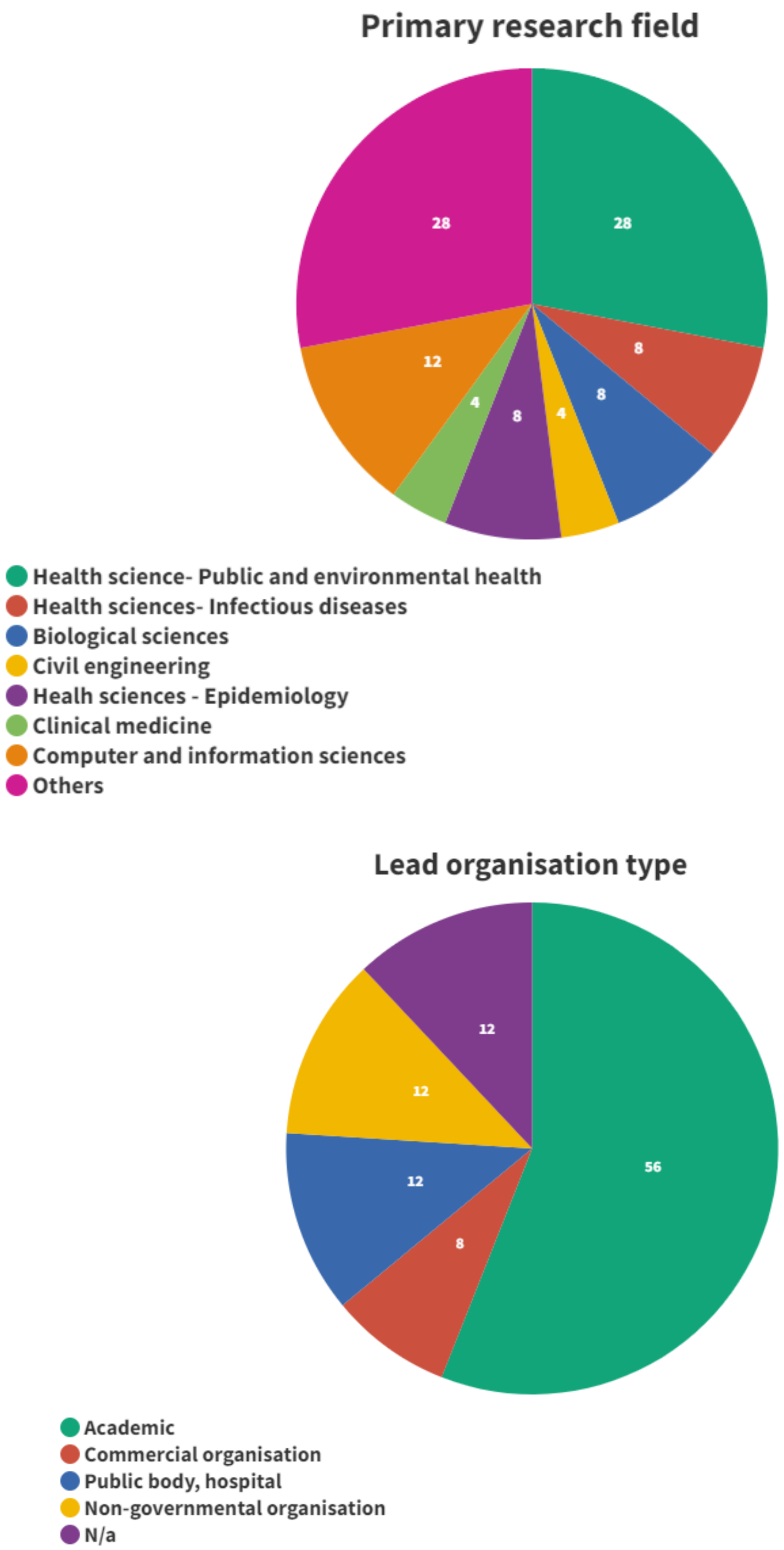


Figure 4 Primary research field and lead organisation category of Covid-19 citizen science projects

3.4 Design of citizen science and tasks

Citizens were mostly involved in the data collection, and over three fourth of projects were contributory (76%, n=19). However, in one project citizens were also given the opportunity to contribute to the project design, by developing the project tracking site, and three additional projects were co-created, whereby citizens worked together with scientists and were involved in several aspects of the research process. Two projects appeared to correspond more to extreme citizen-science, and be entirely driven by volunteer citizens.

The vast majority of projects were participatory crowd-sourced projects (92%), with only two projects relying on distributed intelligence, whereby the citizen volunteers its cognitive capacities to help scientists solve tasks. In the majority of crowd-sourcing projects, the citizens were

asked to respond to an online survey (n=15, 65%), followed by providing self-tracking data, such as from a wearable watch or Bluetooth (n=3). Three projects consisted in distributed computing, where the citizen donates unused computing power (n=3) and only few projects consisted in monitoring of the presence of testing centers (n=1) or in open self-assessment, with no pre-defined questionnaire (n=1).

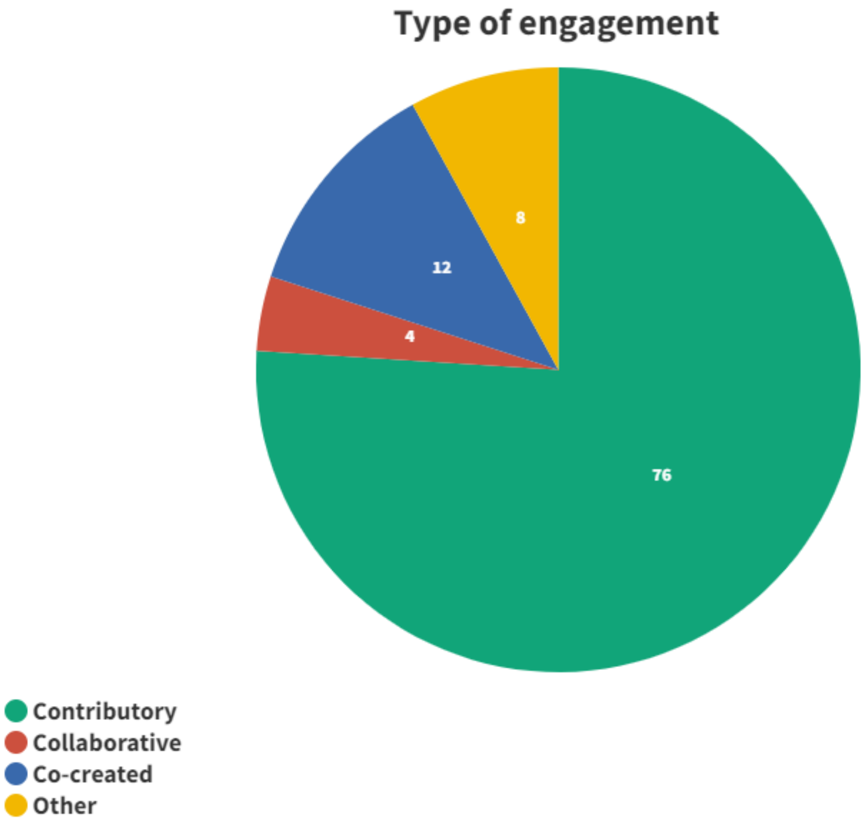


Figure 5. Design of the Covid-19 citizen science projects - level of engagement and main tasks carried out by the participants

3.5 Characteristics of engagement

Almost all projects were widely accessible, targeting a broad audience and requiring no special skills (92%). Only two projects required limited skills (n=2, 8%), or included sections dedicated to specific audiences (e.g. health care workers, schools). Most projects required at least a moderate degree of effort from participants, asking a few types of questions (71% of the 14 projects where the number of questions could be identified asked more than six questions) and many required frequent contributions at regular intervals (daily, 24%; weekly, 20%), whereas another 44% of projects allowed for opportunistic data collection. Only two projects asked for once-off information.

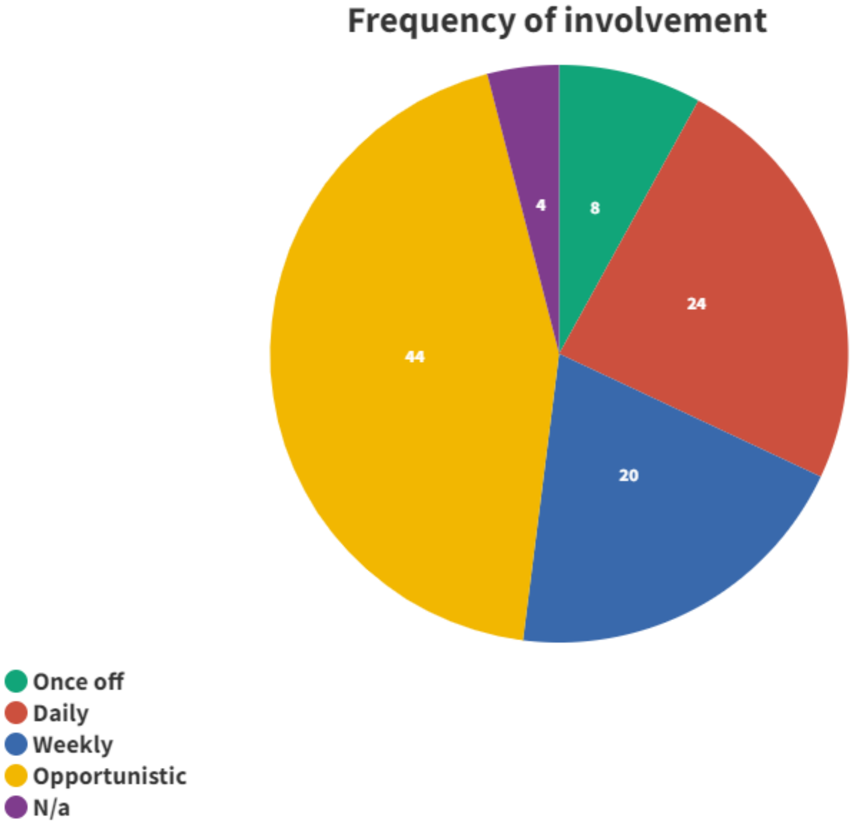


Figure 6. Characteristics of engagement of the Covid-19 citizen science projects - frequency of involvement

3.6 Data required and route to involvement

Most projects required participants to answer a questionnaire (64%), usually also including the location (tracked or from zip code, n=8 projects, 32%), or more rarely a picture (n=2, 8%). One project asked only for the data on location. Another project required only to input the physiological data tracked by wearables, and the two distributed computing projects asked the participants to contribute their computing power.

Most projects used websites as a major route to involve people in the project (48%), and a substantial minority used smartphone applications (20%), or a combination of smartphone applications and website (8%). Game playing was used in one project.

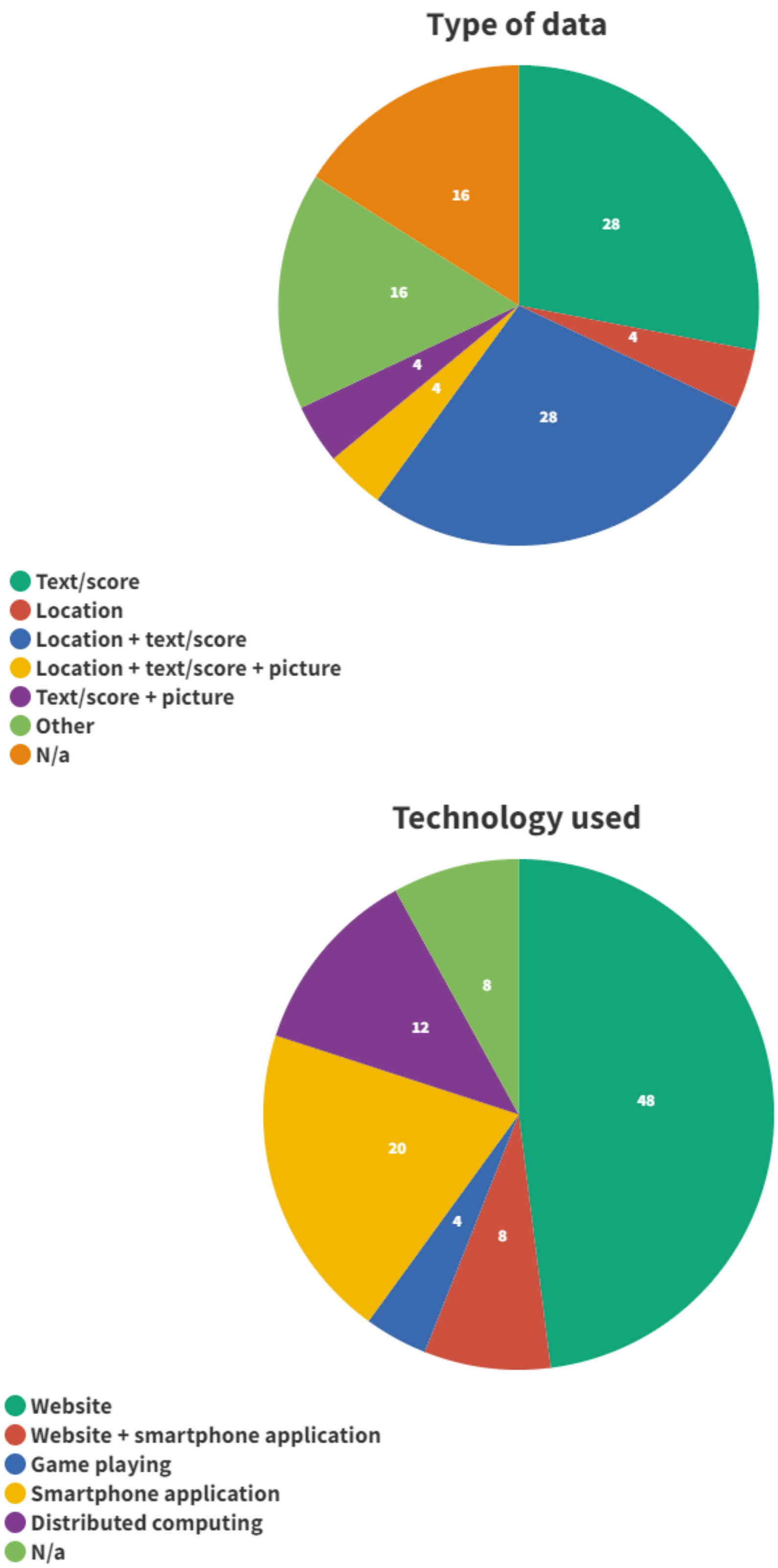


Figure 7. Data requirements and route to involvement of Covid-19 citizen science projects: type of data and technology used

3.7 Data accessibility and quality

Overall, the projects provided limited opportunities for the participants to interact with the data. The data were only available to view in the form of summary outputs in 64% of projects. For those projects, the data were most often presented visually, in the form of maps (n=4), graphs (n=4), or both (n=2), that are dynamically updated, but without additional explanations (n=10, 59%). A substantial minority of projects did provide some explanations with the visual outputs (n=5, 29%). Additionally, one project provided only a few key figures, and a single one made the raw data available for download. Only one project reported the quality assurance procedures used (in a peer-reviewed publication, and not directly on the website). In all other projects, no data quality validation was discussed.

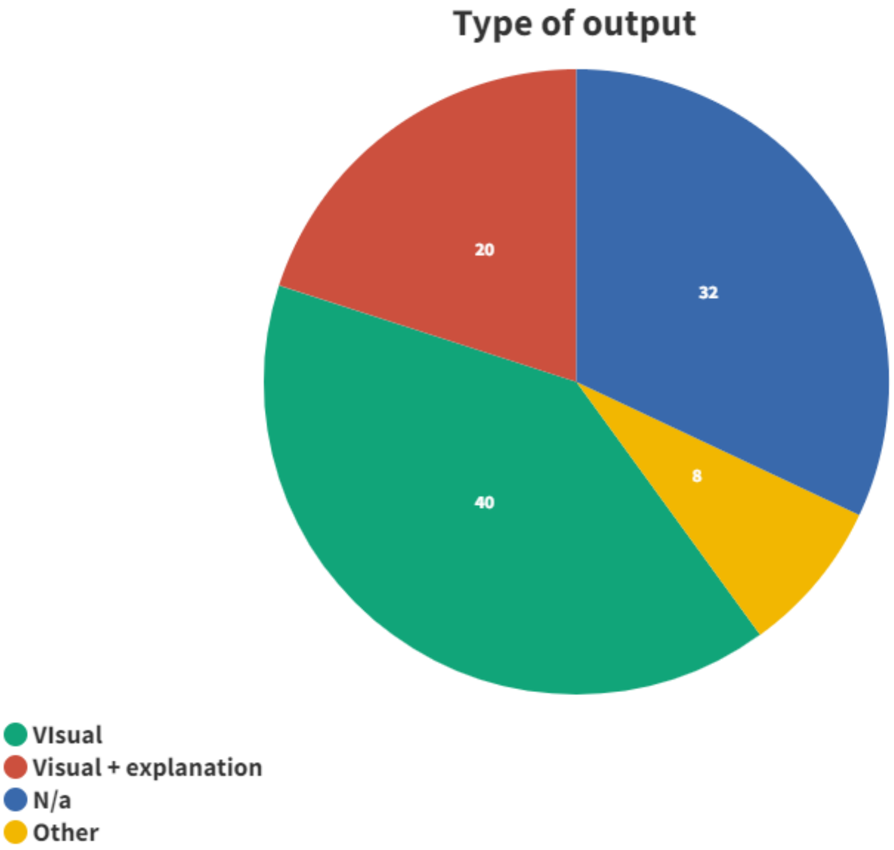


Figure 8. Type of output available to citizen scientists in Covid-19 citizen science projects

3.8 Conclusions

The Covid-19 citizen science landscape is currently dominated by contributory crowd-sourcing projects, mainly related to health sciences. Most of the projects were purpose-developed by scientists to respond to the global Covid-19 crisis, and leveraged basic information technologies (internet, smartphones) to collect data on people's symptoms, behaviour or well-being from questionnaires. A minority of more diverse projects was also found, including projects from biology to solve protein-folding problems, from engineering to explore resident satisfaction with their home under lockdown, or a bottom-up project to centralise and visualise verified Covid-19 data. More than half of all the projects were global in scope, but projects were found at all geographical scales.
In summary this review revealed:

- As expected from the increased use of crowd-sourcing in health sciences over the last few years, most projects are crowd-sourced surveys. Furthermore, crowdsourcing allows for a large number of crowd workers to be mobilized in record time and at low cost, which is a key benefit to respond to a crisis.
- All projects stated their aims clearly.
- The effort required by participants was non-negligible, with repeated inputs at frequent intervals often requested, although each input was expected to take little time.
- Limited data accessibility and absence of information about whether any procedures are used to validate the data quality before its use. In particular, the reliance on self-selected data, the absence of training or calibration may alter the validity of the data. Making data available to download in an understandable form is an intrinsic aspect of citizen science.
- Few projects appeared to have involved the public in their establishment, or to focus on the local, city or neighborhood scale, which may be due, at least in part, to a bias in our project identification, since such projects are less likely to have an online presence and to be featured in citizen science platforms.
- No crowd-sourcing projects identified included problem-solving tasks, which may represent an untapped opportunity given the complexity of medical issues and the amount of data to be considered.

These findings raise some questions about the motivations behind the creation of these projects, and how much consideration was put into the design of the citizen-science participation and in ensuring data quality, as well as about the level of engagement they achieved and their utility for science and public policy. The issue of funding and expected lifetime for these projects also remains open. These aspects will form the focus of the case study analysis.

4 Covid-19 citizen science case studies

4.1 Covid Phym

4.1.1 About the project

Effective drugs against coronavirus are essential to reduce the severity and mortality of the disease. Covid-Phym relies on computer simulations to assess the ability of drugs used against other viral diseases (Ebola, HIV infection, influenza or hepatitis B) to stop the replication of the SARS-Cov-2 virus. The simulations aim to show if any of the molecules manages to inhibit a key protein in the multiplication of the virus called 'RNA-dependent RNA polymerase'. If so, the drug would become an ideal candidate for testing in human clinical trials. Covid-Phym is a distributed computing project initiated by the CSIC Biophym lab (Spain), in collaboration with the Foundation Ibercivis (Spain).

4.1.2 Context and motivations

Research group and previous experience: The project is led by the Biophym group of the Institute for the Structure of Matter of the Spanish Higher Council for Scientific Research (CSIC) and promoted by the Higher Council for Scientific Research (CSIC) and the non-profit Foundation Ibercivis.
The Biophym group, led by Professor Javier Martinez-Salazar (JMS), relies on a combination of experimental and computational research to better understand the structure and biophysical properties of macromolecular systems, such as molecular dynamics and the assembly and functionality of synthetic polymers and biomacromolecules. Some of this research is focused on health-related topics, and the identification of drugs for diseases. For instance, eleven years ago, the research group used experimental computing, to identify monoclonal antibodies targeting the human epidermal growth factor receptor (HER) 2 that plays a central role in breast cancer. It already relied on volunteer computing, implemented in collaboration with Ibercivis.

Motivations: When the Spanish National research council (CSIC) published a general call to direct all research efforts on Covid19 in January 2020, JMS immediately thought to adapt the design his group had previously used for the treatment of breast cancer.
The motivations for this prompt reaction were two-fold. On the one hand, JMS was aware early on of the risk this new disease posed. His son, an MD at Harvard (USA) warned him of the risk of a respiratory epidemic that would lead to hospital care and to the saturation of hospital services. JMS thus understood the need to do everything in his capacity to help find a treatment for the disease. On the other hand, JMS recognized the societal need to better understand the issue and to contribute to the fight against Covid.

The underlying motivation for developing the project were altruistic: JMS was urged by a desire to help and contribute to the fight against this new disease in a useful manner, given his capacity. The team was aware that several other groups were working with the same kinds of proteins, but this did not alter their motivation. They focused on inhibitors of the RNA polymerase since the high-resolution molecular structure of the protein was available, enabling to carry out computer simulations to test the efficacy of different drugs in silico.

Partners: The Biophym lab partnered with Ibercivis in January 2020 to develop the project. In their initial collaboration in 2011, the Ibercivis foundation was just establishing itself as a platform for distributed computing initiatives, based on BOINC infrastructure. The BOINC infrastructure allows to support scientific research from a wide range of domains, ranging from nuclear fusion to protein folding. Today, Ibercivis is a well-established and recognised platform promoting all kinds of citizen science initiatives, with a strong volunteer community behind it. Its BOINC-based infrastructure has more than 20,000 volunteers who give the computing power of their computers and has supported more than 15 research projects. With Ibercivis on-board, the team was able to develop the Covid Phym platform within a couple of weeks and launch the project in March 2020.

Landscape: The project team was aware of several other groups working on the same protein, and kept a close eye on ongoing developments. This enabled it to benefit from the technical advances and databases produced by other research teams.

4.1.3 Design choices and resources

Why citizen science? JMS turned to citizen science for several reasons. First, he knew that with the help of citizens, the team would be able to rapidly mobilise the needed computing power, allowing it to address the problem in a record time, a decisive factor in a time of crisis. Additionally, citizen science allows the team to get started even with low resources and leaves some flexibility for the research approach. More importantly, it was fundamental for JMS and his team to provide an opportunity for the Spanish people to contribute to the fight against Covid. JMS and his group thus naturally applied their previous experience in designing citizen science projects, and turned once more to Ibercivis for support.

Approach: Rapid response is of the essence in a time of crisis, and the project team uses a two-fold approach to try and speed up the process to find solutions to control the pandemic. As is common in the fight against new diseases, the project team decided to use computer methods to test the potential to repurpose drugs with proven effective action for the treatment of other RNA virus diseases. This can enable precious time saving in the steps for developing a treatment, since the drugs tested have already been shown

to be safe enough for human health, and would therefore be available to treat COVID-19 patients much earlier than a newly created compound.
Second, by resorting to the help of the computers of thousands of volunteers connected through Ibercivis' distributed computing platform (Ibercivis BIOINC), the project team can perform a much higher number of simulations in a much shorter amount of time than would have been possible based on its own resources. Volunteers only need to download the BOINC program, an open-source application developed by the University of Berkeley, and join 'Ibercivis BOINC' at the time of installation. By doing so, they are able to easily choose when and how to participate. For example, whoever does not want the computational capacity of their computer to be affected while using it, just leave the default configuration so that the program runs only during pause times, when the screen saver pops up. Immediately afterwards, the computer receives the collection of data that it will analyze and the instructions to do it.

Achieving fast response: Given the sense of urgency of the situation, the team did not have time to wait for governmental support. They thus started developing the project using their own resources, to avoid bureaucratic delays and hurdles. Building on their previous experience and collaborations established in setting up a similar project, enabled the team to abridge the testing and consultation stages. Consequently, the team was able to put together the methodology and set up the infrastructure for collecting the data in a record two-week time. Building on the established volunteer community of Ibercivis meant that the project could be rapidly launched and was fast to recruit volunteers. The project ran for 3 months (March-May 2020), with the whole team working around the clock from February to May , and the first results were published in June[1].

Principles: the project focused on drugs that can inhibit the viral enzyme that enables the replication of RNA viruses and that is well conserved across RNA-virus families, the RNA-dependent RNA polymerase. The three-dimensional molecular structure of the SARS-Cov-2 RNA-Polymerase became known at the beginning of 2020, enabling to conduct virtual assays between various drugs and the target in that viral protein. Candidate molecules were antiviral drugs that interfere with the action of RNA-Polymerase, such as Remdesivir (used in the treatment of Ebola), Favipiravir (used to treat influenza) and Tenofovir (used in the treatment of AIDS and hepatitis B). Initially, the project considered 8-9 candidate molecules, to finally focus its attention on two molecules that are cheap to produce, Remdesivir and Tenofovir. Computer simulations were then used to find the most stable compounds with the lowest binding energies, which have the highest affinities with the RNA-dependent polymerase.

---

[1] (https://www.authorea.com/doi/full/10.22541/au.160133726.63184055).

Extensive searching in the space of protein-ligand docking conformations is needed to measure the strength of the interaction of each of the possible associations between the drug and the protein. This process requires hundreds of thousands of calculations, which would take several years to perform with a conventional computer. This is where the power of distributed computing was leveraged, with the support of Ibercivis, which developed the Covid-Phym platform. The platform was implemented using BOINC (Berkeley Open Infrastructure for Network Computing), whereby the operations are divided into small packets that are in turn sent to each of the volunteered personal computers, achieving a calculation capacity similar to that of a supercomputer.

4.1.4 Participation

Participation: The citizen participation is an essential part of this project. The project is designed as a vehicle for information and involvement of the society. Most of the participants were regular members of the BOINC community, who participate from one project to another. Although efforts were made to try and get new people to contribute, the majority of the job was done within this community. In total, around 3,000 volunteers contributed to Covid Phym, with more than 7,000 computers and 4 million computing hours. The majority of volunteers came from Spain, but many joined from all around the world as well (e.g. DE, USA, FR, AUS, …). In total, about 3.2 million potential binding modes were collected for each protein-molecular target pair.

Outreach: Substantial efforts were made to communicate the project and promote citizen engagement. The project was endorsed by CSIC, through an official project launch and press release, as well as wider media outreach, leading to radio and tv interviews. This support was significant as CSIC became a trusted source of information on Covid in Spain, by means of their Covid-19 platform and blog which provided real time information on ongoing developments, and debunked false information.
Ibercivis also actively promoted the project through its network, leveraging the volunteer computing community. In particular, the German BOINC team organised a Pentathlon for Covid-related volunteer computing projects in which Covid-Phym was selected. This attracted significant computing resources to the project, as the volunteer computing teams competed to contribute to the selected projects over the course of two weeks[2]. This created an important boost for the project, helping both in raising the computing time and in further disseminating the project. Additionally, JMS contacted different public institutions, notably in Andalucia, several of which were happy to volunteer their computing servers, since these were often left unused because of the pandemic and confinements. The participation of these governmental organisations received a wide

---

[2] https://www.seti-germany.de/forum/content/1231-BOINC-Pentathlon-2020-5th-project-in-the-discipline-Sprint

coverage from the local press, in turn generating a high citizen response. The project thus gained popularity in a relatively short time.

Engagement: During the project, citizen engagement and information was managed through a forum, allowing participants to communicate between themselves and to raise any questions. The forum was moderated by Ibercivis, who was in charge of relaying the information to the scientific team and of providing regular updates on the progress of the project. In particular, the scientific team explained the aims, lowered expectations when needed. A short project summary with the main results achieved was shared with the participants before the project ended[3].

4.1.5 Outcomes and lessons learnt

Outcomes: All expected data were obtained by the end of May 2020, when the project ended. The work of the scientific team on the other hand is continuing, in order to analyse these results and to complete some clinical studies. Additional funding may be sought to this end. A couple of scientific publications already showcase some of the results obtained by the project[4].

Successes: A key success factor of the project were the volunteers. The team felt that people responded to this call in a very generous way, always raising to the challenge each time they were asked for a contribution. The project team was both impressed and emotional about the great response from people. It was a strong source of motivation throughout the project.

Lessons learnt: Covid Phym underscores the importance to involve all citizens, and to give them an opportunity to do something to help fight the global and personal crisis they are experiencing. It helps to have a starting point, and to be able to take advantage of the community already engaged in order to leverage more people. In plenty of citizen science projects, a lot of efforts are often deployed to engage people, which could be optimized by going to the communities already working on these issues, or with these methods.

We did not expect that this project would be as successful and widespread as it got, and it was launched only in Spanish. In the end, other people were interested to join, but it was too late to change the infrastructure.

The main limitation for the continuation of the project was the lack of financial support. All the project team volunteered its time during the project, but it takes a lot of time to support the community of volunteers, and to solve bugs that can appear on the platform.

---

[3] https://www.csic.es/sites/default/files/proyecto_covid_ibercivis.pdf

[4] https://digital.csic.es/bitstream/10261/220537/1/483674.pdf ; https://d197for5662m48.cloudfront.net/documents/publicationstatus/66195/preprint_pdf/5b9749603e7bc22cd4f9519a6610583e.pdf

4.2 OpenPandemics

4.2.1 About the project

OpenPandemics is a distributed computing project developed to speed up the identification of candidate molecules for Covid-19 treatment. It is leveraging molecular docking techniques developed in the Forli lab at Scripps Research (USA) at scale, thanks to World Community Grid's massive computational power, to virtually screen millions of chemical compounds to see which might be capable of binding to proteins of the SARS-Cov2, as well as other coronaviruses proteins. OpenPandemics goes the extra mile to then test the most promising compounds identified by simulations in the laboratory, so as to provide the biological validation of the results. In a long-term perspective, the project also hopes to build a fast-response, open-source toolkit that will help all scientists quickly search for treatments in the event of future pandemics.

4.2.2 Context and motivations

Research group and background: OpenPandemics was developed by the Forli lab at Scripps Research, in partnership with IBM's World Community Grid Initiative. The Forli lab develops new computational methods to predict how chemical compounds fit together with macro-molecules, such as viral proteins and DNA/RNA. This enables it to screen for potential candidate drugs that would be effective in the treatment of diseases, such as HIV and malaria. Such molecular simulations require vast amounts of computational processing power that can be met through distributed computing, whereby volunteers on the grid lend their processing power to a common scientific cause, effectively forming a virtual supercomputer.

The research group already relied on distributed computing efforts to screen libraries of commercially available therapeutic agents in the search for potential HIV treatments[5]. Already then, it partnered with the World Community Grid, an IBM social impact initiative, to implement the distributed computing infrastructure that allows anyone with a computer and an internet connection to donate their device's computing power to help scientists study the world's biggest problems in health and sustainability.

As the Covid-19 pandemic unfolded, and in light of their long-standing collaboration, IBM approached the Forli lab to develop a project to help fight Covid-19. Despite their experience in developing similar projects, Open Pandemics presented specific challenges. By then, all Scripps researchers and members of Stefano Forli's lab were working remotely, some of them at home abroad. Moreover, usually a fairly long incubation time is needed in order to set up the project resources for crowdsourcing projects, such as the software infrastructure and the team to receive and process the

---

[5] FightAIDS@Home; http://fightaidsathome.scripps.edu/

data. The software tools were readily available from previous projects, says Stefano Forli, but if you plan for success, you need to anticipate the need for a computer server with a large storage capacity in order to pull and upload all the crowd-sourced data, and manage your resources optimally.

Project motivation and goals: The motivation for the project was to provide tools to respond to the current pandemic, but also to prepare for the ones that may follow.
The primary goal for the project was scientific, to study proteins from the SARS-CoV2 virus and identify new molecules that could help find treatments for COVID-19. Since future pandemics could stem from a progressive accumulation of mutations, which can eventually lead to a new virus variant, as happened with SARS-Cov2, the researchers also focused on other coronavirus proteins, in order to assess how difficult it would be to identify molecules capable of overcoming mutations. A collateral objective, in keeping with World Community Grid's open data policy, was to create an open-source toolkit that could be used as a basis for seeking treatments in the event of future pandemics. As a result, all data, tools, and processes that are developed through OpenPandemics will be made freely available and accessible to the scientific community.

Landscape: At the time when OpenPandemics unfolded, several other similar volunteer computing projects with similar overarching aims were in preparation, but it was not considered a hindrance to the development of OpenPandemics, since the questions to be answered were so broad, from understanding the structure of the Corononavirus, to searching for treatments, and the urgency so prodigious, that the more hands on deck the better. The OpenPandemics team reached out to the two or three other major initiatives underway to try and find a common ground. It turned out that each project approached the issue in a slightly different way, or with different techniques, creating a healthy diversity, and such that there was little overlap. Molecular docking is indeed a complex and multifaceted problem, and there are many computational tools for this task, which are often complementary. IBM also collaborated with other national labs and super-computing centres on the issue, not solely with Forli's team.

Partners: The team collaborated with other labs, in particular for the experimental part of the work, and received support from chemical companies. The OpenPandemics researchers had a limited budget to carry out the experiments to test the candidate molecules identified. The group was fortunate to receive the chemical compounds to be tested for free from chemical companies. It was approached by numerous other scientific teams in a very collaborative fashion, but the researchers chose to work mostly with colleagues within Scripps Research Institute, or with whom they already had established collaborations.

4.2.3 Design choices and resources

Principles: The Forli lab leveraged the molecular-docking technique that they developed in previous projects at scale, in order to rise to the challenge posed by a little understood and rapidly spreading coronavirus. The approach was to virtually screen different chemical compounds to see which might be capable of binding to proteins of the SARS-Cov2, or other coronavirus proteins, in the quest for promising drug candidates against Covid-19. Thanks to the World Community Grid's massive computing power, the team could plan to simultaneously screen millions of chemical compounds from a variety of sources, including in-house libraries of compounds available at Scripps Research, existing commercially available and clinically tested compounds for repurposing, and libraries generated in silico. The benefit of using compounds from libraries, instead of from the entire realm of possibilities, is that they have a very high likelihood (> 85%) to be synthesizable; some of these molecules also include drugs for repurposing or molecules that failed clinical trials, which are good candidates because it was demonstrated that they are not too toxic for humans and the approval process can be speedier.

A singularity of OpenPandemics is that it goes beyond in-silico research, and proceeds to validate the effectiveness of the most promising compounds experimentally, by sending them to collaborators for laboratory testing and chemical optimizations. OpenPandemics is generating results at the same pace that the research team is analysing them, and then sending them out for experimental validation. This gives the contributors a sense being at the heart of the Covid-19 fight. This aspect made the OpenPandemics particularly valuable for IBM, differentiating it from the other crowdsourcing projects it supports.

Resources: Stefano quickly devoted all his lab resources and efforts to OpenPandemics, succeeding to establish the project within three months. In total, 10-12 people worked on the project full time. NIH also allowed to divert resources from other projects to Covid-19 research, allowing the lab to recruit experimental biologists willing to test the compounds. Enamine, a chemical vendor company, partnered with the team to support the orders and synthesis of compounds for the experiments. Besides providing the computing power, IBM provided support in terms of funding to set up the software infrastructure, and purchase molecules for experimental testing.

Although the technical feasibility of making the project work remotely, with new collaborators joining the team, was an initial concern, Stefano Forli acknowledges that his concerns vanished rapidly. Everybody in the lab was incredibly motivated, working late hours, and around the clock, benefiting from the fact that team members were scattered in confinement with their families in different time zones.

4.2.4 Participation and engagement

Participants: Anyone equipped with a laptop or desktop and an internet connection can volunteer to OpenPandemics. The volunteer simply needs to download a secure software program to his computer, such than when he is going about his regular activities, typing emails or browsing the internet, if his computing power is not fully used, the computer will automatically run a simulated experiment in the background. The simulations help predict the effectiveness of a particular chemical compound as a possible treatment for COVID-19. When the simulation is completed, the volunteer's computer contacts the World Community Grid server, which combines them with hundreds of thousands of results from other volunteers all over the world, and sends the data back to the research team. The whole process occurs seamlessly, automatically and securely.

Data visualisation: Participants were provided with the opportunity to visualize what they were contributing to. The right portion of the screen saver shows both the target protein and drug candidate molecule that your device is currently working on. The progress bar, towards the bottom of the screen saver, represents approximately how much of the current research task your device has processed. When it reaches 100%, the computation is complete and the results will then be sent back to World Community Grid and delivered to the OpenPandemics researchers.

Engagement and outreach: Over 600,000 people worldwide contributed their computing power to OpenPandemics. This remarkable engagement was due in large part to the terrific efforts IBM made in advertising the project, but also to the existing volunteer community. OpenPandemics was able to leverage the existing community of the World Community Grid. Surprisingly, only about half of the participants were people that were currently active on other World Community Grid projects. The other half were people that had been inactive members of the community for some time, and that resumed participation, possibly in light of the urgency of Covid-19 and the fact that they were at home. In addition, IBM created a buzz for the project launch, through an active media coverage and social media presence, that was actively relayed by Stefano's group at Scripps Research. IBM's president himself endorsed the project and tweeted about it. IBM also communicated widely internally, throughout its different sister companies, such as Red Hat and Linux. In a snowballing effect, the project team started getting several solicitations from the media, even an Italian newspaper, to learn about the project.
In March 2021, in response to a demand from the community, the World Community Grid released a faster version of the molecular docking application, running on GPU (Graphics Processing Unit). Within a week, the research team developed a new code for it and carried out a stress test in order to assess the upper capacity of the system to support the load if it was provided with enough GPU work. This created a huge buzz in the community, and computing teams tried to put together as many of their resources as possible to win the challenge. People want to contribute, acknowledges Stefano. This dynamic was always there, we experienced with HIV research ten years ago, but has been exacerbated with the Covid-19 pandemic.

What was in it for the participants? People are enthusiastic about contributing, but eager to know that their effort is not in vain. Forli's group was well aware of the importance to keep the community of contributors happy and active, even though this sometimes represented a challenge. The team of researchers was pressed for time and resources, since it was simultaneously updating the molecular docking software at the heart of the project, preparing the input datasets and performing the analyses of the results. But IBM was always mindful of the social dimension of the project, and constantly poked the researchers to provide regular updates to the participants about the project progress. They requested regular progress reports, with content that they could update the volunteers with. The team fast realised the importance of these updates to secure engagement and interest. The responses to the updates were very positive and appreciative, with volunteers encouraging each other, and each update led to boosts in participation. IBM also created a forum for people to post their questions and invited the researchers to answer the most interesting ones in a weekly digest. The forum was a very good venue for interacting with the community. The research team was surprised to discover that the volunteers were interested about every single detail of the research, including very technical ones. But people also shared much more personal stories, such as a testimonial of someone who lost his mother to Covid-19 asking people to volunteer for OpenPandemics. Others directly contacted the researchers to express their appreciation and hoping to have been of assistance. The community has been very positive, very appreciative of the researchers efforts, and was both moving and a source of motivation. It seems that OpenPandemics gave the contributors a sense of purpose during the Covid-19 epidemics, admits Stefano Forli. People were locked down, and contributing to the science was empowering, giving people the feeling that they are helping society

4.2.5 Outcomes and lessons-learnt

Outcomes: The remarkable collective effort allowed the project to complete more than 415 million computational assignments, or a rate of about one million assignments a day[6]. It would have taken a single processor computer the equivalent of 70,000 years to perform these computations.

By November 2020, the project identified 70 virtual compounds, from a group of over 100, 000 screened products, that showed promise for inhibiting SARS-Cov-2. That identification process took over fifty billions of energy calculations to check how well the shape of compounds for a potential treatment "fit" five different viral protein targets when they were digitally "docked" together during a simulation. About one billion dockings were

---

[6] Project statistics on 28/06/2021

processed, allowing to reduce the number of results from hundreds of millions of molecules to a few dozen that showed the most interesting interaction patterns with the viral enzymes. The researchers identified those compounds that were the most accessible through synthetic chemistry, ultimately selecting molecules that could target two of the main proteases of the SARS-Cov2 virus.
To carry out the experimental validation, the researchers looked for a company that could produce the compounds already in October 2020, so that these would be ready to be purchased quickly for lab testing of potential treatment. The team partnered with Enamine, a large chemical vendor, who was able to rapidly synthesize, purify, and ship the molecules to the team's experimental collaborators at Scripps Florida and Emory University and at the University of Oxford. A few dozen dockings have been confirmed experimentally in the lab so far, but the researchers still need to confirm the mechanism and validate that they are binding where they expected, before making the results public. Not all data that has been made public during the Covid were reliable, explains Stefano Forli. Our partners are very cautious and want to validate the results before going public with them. Unfortunately, this is a side project for them, so things are not moving as fast as if it had been their main project and we had the funding to ensure that, he adds.

Stefano declares that his team really broke some barriers scientifically, and that this has only been possible thanks to crowdsourcing. The team used a protocol that they developed just before Covid-19 for irreversible binders. These binders were thought to be toxic until now and little studied, but with proper chemistry, it is possible to reduce their toxicity to humans and benefit from all the advantages of irreversible binding on viral proteins. The team could not have tested the new protocol without the massive computer power obtained from crowdsourcing.

So far, no results from the project were published in the general media, or in scientific publications. However, the researchers recently submitted two papers related to the World Community Grid for peer-review. Scientifically, the data generated in the project could represent an ideal starting point for the design of new classes of drugs to target COVID-19 and other coronaviruses. Since all data, tools, and processes that are developed through OpenPandemics will be shared freely, in keeping with World Community Grid's open data policy, the findings and tools will be freely available to the scientific community. the project will benefit the scientific community at large. In particular, OpenPandemics researchers are creating a software infrastructure to streamline the computational research process of finding potential treatments for other diseases. OpenPandemics is thus directly helping to build a model for the fast deployment of computational resources to identify potential treatments to help address future pandemics.

Sustainability: There's a good possibility that OpenPandemics may have additional phases, says Stefano Forli. From what can be learnt from past outbreaks, it is likely that pandemics caused by newly emerging pathogens will become more and more common. That's why the project is being designed to be rapidly deployed to fight future diseases–ideally before they reach a critical stage. We hope to have enough preliminary results to be able to secure funding for the project continuation in the long-term.

Did you meet your goals? The project was incredibly helpful in many ways. On the more technical side, Forli's group was able to test the code and molecular docking software it developed on a scale that was not possible before. This allowed the team to identify unforeseen problems and fix them, optimizing the entire process.
From a societal perspective, OpenPandemics agreed with IBM to make the code infrastructure for managing the project publicly available. This should make it much easier to launch any new project based on these methods in the future.
From a broader scientific perspective, the research group is in a good position to propose something really different. The researchers switched to a different, much more powerful, methodology that is less specific and can have much broader targets on the viral proteins. Deploying this method on such a large scale took a big leap of faith: numerous of parameters had not been tested and the model could not be trained before.
OpenPandemics demonstrated the efficiency of this method: the sheer quantity of molecules screened is impressive, and the success rate of identifying fit was very good. Usually, the success rate of computational predictions when tested in the lab lies around 2 or 3%. Here, out of 20 candidate molecules, five turned out to be effective when tested in the lab, corresponding to a tenfold improvement. These advances are important not only for the current pandemic, but also in terms of preparedness for future epidemics. The group will publish the method soon, and is in a good position to propose something really different and innovative. This allowed the researchers to carve a research niche, which tackles the problem from a different perspective than most other research. For Covid-19, the group is hoping to publish its findings by the end of the year (2021).

Lessons learnt: Stefano Forli recognises that a large part of the success of OpenPandemics came from the fact that IBM did a fantastic job in advertising the project. Any new project, even a conventional science project, should think about proper media coverage. It is much easier to recruit collaborators at the beginning, and the process then self-reinforces itself, attracting more collaborators or even support. In the case of OpenPandemics, Enamine was happy to work with the team because IBM was on board, and they knew that the molecules would get tested. But whereas for the researchers, the main benefit of involvement are the scientific findings, for industry partners or others, the media coverage is the reward. It helps convey an improved reputation and image of the company to the world.

Additionally, ensuring engagement is crucial for success, and requires an active effort on the project's side. People are happy to contribute, but they want to know that their efforts are not in vain. The efforts it takes to provide regular feedback to the community should not be underestimated, but the value gained, both from a human and data perspective is immense.

Challenges: With crowdsourcing projects, there is a strong asymmetry in how fast it is to produce a lot of data and how long it takes to analyse it and test it in the lab. The project put a lot of strain on the team. It was a huge time crunch, with pressure on the one hand to produce results and not being too late for the game, since many other projects had the same objective. On the other hand, the team needed to keep the community of contributors and collaborators happy. Finding the right balance is tough, concedes Stefano Forli. At some point OpenPandemics did not have enough data to process for the computing power available, which required adapting the code to increase the capacity of the infrastructure. But once you have the data you need to be able to process it fast enough. The entire team worked around the clock, across multiple time zones and including weekends. It is important to find compatible people that keep up with the pace to strike the right balance between the two objectives and manage your resources optimally.

4.3 Safecast Covid-19 map

4.3.1 About the project

Safecast was developed in Japan, in response to the Fukushima natural disaster, and has become a leading and established radiation monitoring project. Safecast's belief is that people deserve to know what they have been exposed to and what they might be exposing others to. It thus decided to build on its crowdsourcing experience to develop a Covid-19 testing map. With the help and input of people around the world, the map would hopefully begin to provide a more accurate picture of the relative difficulty of obtaining testing in various locations and become a useful tool with which to better target resources and hold governments and officials accountable. As more people contribute their stories to the map, the larger picture will become clear and the need for comprehensive widespread testing will become obvious. The map provides coloured markers according to the type of experience ("Refused Testing" "Testing Unavailable" and "Successfully Tested"), to which a detailed narrative can be appended.

4.3.2 Context and motivations

Research group and background:
Azby Brown is an associate professor of Architectural Design, and after 12 years as the founding director of the Tokyo Future Design Institute, he now retired and primarily focuses on his work as lead Researcher for Safecast. Safecast is a global volunteer-based citizen-science organization devoted to developing new technology platforms for crowdsourced environmental monitoring that promote open-source and open-data principles. Azby has been a core member of Safecast since its creation, after the Fukushima nuclear disaster in March 2011. Azby is involved in Safecast's day to day operations and responsible for aspects related to education and outreach. He is involved in assisting affected communities and reporting the issues they face, and represents Safecast at international events.

Background
Safecast was created after the Fukushima nuclear meltdown in 2011, in an aim to answer the public need for accurate, trustworthy radiation information. One of Safecast's founding beliefs is that people should have access to reliable and accurate environmental information in order to make decisions about their own safety. It developed an innovative model involving the rapid, agile development of new, affordable technologies so that citizens could monitor and share information on environmental radiation in a participatory fashion. The project rapidly gained credibility, reach, and grew in scope, expanding to air quality monitoring as well. It is now a respected source of information, valued for its neutrality.

As they observed the global spread of Covid-19 and the responses to it across the globe, the Safecast team could not help but see striking similarities with the Fukushima nuclear meltdown. In both cases, fear of the unknown and a rapidly changing situation fueled general anxiety. This anxiety was further heightened by the lack of transparency from public authorities, misinformation and the politicisation of the debate. But the landscape had shifted a lot over the past nine years. The environment for online citizen science activities has matured, information is more widely available, and it is easier to find trusted sources of information. Yet, it is not always simple to make sense out of this flood of information.

It thus followed to identify ways in which the Safecast's experience could be put to help in this new situation. Coronavirus is very different, and in many ways a more complex issue than radiation. The team did not believe that they had the expertise or know-how to offer anything medically (although they did support FabCafe's activities to create personal protective equipment, such as facials and ventilators). But they did have expertise in outreach, crisis communication and getting communities together.

Specifically, noting the ever-growing collection of dashboards tracking the numbers of confirmed cases, tests and deaths, the Safecast team came to realize the human element was missing. It saw the gap between the availability of testing claimed by some governments and what was actually available to people. It also observed that the stories of the people who were suffering were getting lost behind the statistics.

Project motivation and goals: The trigger came from the situation in Japan, where most of the team is based. They were startled by the lack of clarity about testing, and numerous reports of acquaintances that wanted to get tested but were unable to. Upon investigation, they uncovered that there was a national (non-transparent) policy decision to limit testing, promoted by the medical advisory board. The medical advisory board was concerned that the healthcare system would get overwhelmed, since at the time, in Japan, anybody who tested positive for Covid-19 needed to be hospitalised, regardless of the seriousness of his condition. This blanket policy created a bottleneck, such that people had to jump through numerous hoops to get tested in case they experienced symptoms, even when they were sick for days. This made the testing data less than robust. The country had also developed a contact tracing methodology to identify and isolate clusters that were in contact with an infected person. While the method worked quite well in the beginning, it was strongly criticised by experts. The team thus thought that it would help if they could map the testing facilities and let people inform others of their experience in getting tested. In this emotionally traumatic time, it would let people share their concerns, and put faces and stories behind the numbers.

4.3.3 Design choices and resources

Approach: Taking the appropriate steps to stop the pandemic relies on people having access to information about their health and environment. A core value of Safecast is to provide credible sources of information to enable people to free themselves from the dependence on government-based information. In response to the Covid-19 emergency, Safecast thus ventured to provide an alternative source of crowdsourced information on Covid-19 testing experiences. It originated from the observation that in many places around the world, testing information was ambiguous and incomplete, and people were dependent on single sources of official information, which were not always relevant or trustworthy. Due to the failure of test kits to arrive where needed, delivery of incomplete test kits, overly complicated approval processes, or favoritism and discrimination, it became apparent that there was a gap between the availability of testing claimed by some governments and what is actually available. With the help and input of people around the world, the crowdsourced information can be mapped and begin to provide a more accurate picture of the relative difficulty of obtaining testing in various locations, and hopefully become a useful tool for people and with which to better target resources and hold governments and officials accountable, so that people get easy access to Covid-19 testing options.

From the start, Safecast has embraced open-source and open-data methodologies, along with the development of affordable technologies. It has enabled people to easily monitor their homes and environment, and free themselves of the dependence on government-based information.

Data accessibility: All Safecast data are openly shared and accessible in the public domain on the Covid-19 testing map. Since the contributions to the map are crowd-sourced, it would be impossible for Safecast to verify the validity of the information and there are no data validation procedures. This is openly recognised on the project website.

Data visualisation: By design, the Covid-19 testing map provides a visualisation of the locations where people shared a testing experience, within a 10-km radius to protect the privacy of the contributors. The map uses a traffic light system to rapidly indicate the testing experience (green - successful testing; yellow - refused testing, asymptomatic; orange - refused testing, symptomatic; red - testing unavailable). The map offers the possibilities to zoom in and out, and to show the specific data associated with each entry. It is available in nine different languages.

Resources: The Safecast team got the Covid-19 testing map up and running within a week, an accomplishment in itself. It was able to rely on the existing team - the core team of a handful of people responsible for the development of the API and hardware issues;

the outreach team, comprising about 20 people around the world active on a weekly basis and volunteering their time on outreach issues, and the slack channel, which involves around 200 people. For the development of the Covid-19 testing map, Safecast used the Ushahidi open-source software platform. No additional funding was received. In fact, in 2020, the project suffered from lack of funding since most of its donors were focused on Covid.

4.3.4 Participation and engagement

Participation and engagement: The Safecast team actively promoted the Covid-19 testing map, frequently speaking at public events, in webinars, or through its blog posts. The project gathered interest in the beginning, but participation levelled off by June 2020, less than three months after its launch. In total, 129 participants contributed to the Safecast Covid-19 testing map, including about 100 from Japan, mostly foreign residents. Safecast's testing map all received a fair amount of contributions from Europe and the US, demonstrating that there was interest in the network. The participation levels reached were reasonable, considering that Safecast's active community consists of a couple of hundred people, out of its thousands of users. But the participation level was much less than the team had hoped.

There are several reasons that could explain this limited uptake. From a cultural point of view, Covid-19 infection was taboo in Japan, and people were unwilling to associate themselves with it. It did not matter that the participants were granted full anonymity. Overall, Japanese people were very reluctant to say to others, even family, that they were going to be tested. According to the Shinto religion, anyone exposed was now impure. The first infected person was typically considered like a criminal, and anyone it infected as the victims. There were examples of harassment, such as with nurses asked to remove their children from school. Similar reactions had happened after Fukushima and were also founded on a lot of misinformation and ignorance of the medical aspects.

What did the users gain? The value of the Covid-19 testing map depended on getting sufficient participation, so that it could provide à representative picture of the situation in a given location. Without sufficient input, it failed to become a useful resource. Azby Brown reckons in hindsight that it is likely that if the project had been able to provide a service, for instance to help people identify where they could get tested, it would have gathered a wider interest.

4.3.5 Outcomes and lessons-learnt

Outcome: Azby Brown admits that Safecast Covid-19 testing map did not achieve the uptake they had hoped for. Back in June 2020, the whole team was very active on the map, but gradually felt it was not being effective. The Safecast team realised that they had probably moved too far out of what made the core identity of the network.

Challenges: The Covid issue was too big for them to leverage their network and expertise in the successful way they had been able to for radiation. The growth of Safecast came from the fact that it provides people an opportunity to do actual science, and to contribute to an open and progressive agenda. But with the Covid-19 pandemic, it was beyond the team's expertise to tackle medical or scientific issues frontally. Sticking to the DNA of Safecast, by providing compact test kits for instance, is not something they could have developed on their own, without support from a bigger institution, and involves more complex ethical and regulatory issues than what they had dealt with so far.

Furthermore, although Safecast is a virtual organisation, in the sense that all its functions are organised and developed online, through virtual teams and a slack channel, its activities involve a strong direct contact dimension. Face to face outreach and workshopping have always been essential for generating interest and to demonstrate to people how things work. This was impossible under the Covid restrictions.

Other aspects probably did not help, such as the fact that Covid was very much a personal issue, in contrast to Fukushima, which was an external environmental issue, meant there was reduced interest in participating and contributing data.

Why not branch out? Safecast has limited resources, and the directions it takes are tightly linked to the core's team interest and availability. "It is true that it did not really occur to us to focus on opportunities for vaccination", recognises Azby Brown. But at the time, team members behind the testing programme had personal priorities and did not have the time to initiate something new. Similarly, the group never considered developing surveys further, since Safecast is more about providing sensors that are affordable and accessible.

Lessons learnt: The Safecast Covid-19 map was a big lesson for us, recognises Azby Brown. "It helped us understand where the limits of the Safecast method are". Safecast is about providing environmental sensors, and being a neutral source of information. Contrary to Covid-19, there are well-known technologies with known reliability to measure radiation, or air quality. Safecast is also about outreach, education and crisis communication, to increase understanding of environmental data. In that regard, although there were strong similarities between the Covid-19 and Fukushima discussions, in terms of disinformation, and flood of information, Covid-19 was a personally difficult issue, and the network may have underestimated the reticence of people to associate themselves with the illness.

The Safecast team observed a number of similarities between the Covid-19 pandemics and Fukushima meltdown, which unfortunately suggest we did not yet learn from the lessons of the past. Both radiation and viruses are invisible threats, making them a source of anxiety as people have almost no way to determine whether they have come into contact with either, and must rely on experts, measurements and governmental information, relayed by media reports. This underscores the importance of how governments behave in the initial days of the crisis, since it will determine public trust going forward, and whether the public will seek alternative sources of information, often in the form of online rumors. This was very obvious after Fukushima, and the same occurred with the Covid-19 pandemic, where misinformation sometimes came from the highest administration offices, putting experts in the difficult position of having to choose between speaking up or their jobs. In 2011, the way governments behaved during the initial days and weeks of the nuclear crisis determined whether or not the public could trust what they said later. As the Fukushima disaster unfolded, the Japanese government did not provide timely or credible information about the spread of radiation or their response to it. The media failed to hold representatives accountable, to vet their own sources, and to provide reliable information when it was most needed. The government and media lost the public trust within days and still have not regained it, nine years later. We are seeing the same patterns emerge with regards to COVID-19. Governments are inadequately prepared to handle the incident, and official statements have been vague, contradictory and/or misleading. This information vacuum or confusion opens the door for rumors about the virus and its effects on many aspects of daily life, resulting in panic-driven actions, as witnessed with the sequential shortages in face masks, hand sanitizers or toilet paper. The lack of trustworthy information has spawned conspiracy theories, and scams offering supposed cures. Although these can most often be easily debunked, they remain in circulation, and are pervasive, and even scientifically educated people can be vulnerable to it.

This underscores the importance of more prepared and competent governance. Good crisis communication relies on providing clear, trustworthy information quickly, and to ensure actions are consistent with the messaging. The role of organisations like Safecast can be in helping them understand the consequences of lack of information and transparency. The network is currently supporting efforts to include better governance in policy or legislation. Achieving this change is a very long-term process, in which the European Commission has so far been the most receptive and progressive.

Safecast can be part of such processes for radiation, because it has succeeded to establish its credibility. The Safecast network succeeded in getting participation from the expert community, and provided a neutral objective source of information which quickly helped it gain credibility. As a testimony, the network is now regularly solicited by official organisations. This allows it to raise its voice.

4.4 Quantified Flu

4.4.1 About the project

Quantified Flu is a collaborative, community science project which explores how wearable devices and symptom self-tracking can help predict and understand when we are getting sick. While this question seems more timely than ever with the Covid-19 pandemic influencing all our lives, it applies much more broadly to other infections, such as flu and common cold.

In the longer term, the project team hopes to use these data to learn whether physiological variables measured by wearables can be used to train algorithms to predict when someone is falling sick, before the person is even consciously aware of it.

4.4.2 Context and motivations

Research group and background: The Quantified Flu project was initiated by the non-profit Open Humans Foundation (USA), where Bastian Greshake Tzovaras is acting director of research. As Director of Research for Open Humans, Bastian works to facilitate a wider re-use of personal data in learning and research through citizen and community science.

Doctor Bastian Greshake Tzovaras has a long-standing interest in community-driven and open science. With a PhD in bioinformatics, and a background in biology and evolution, he has been involved in citizen science projects for over 10 years. During his Master's degree, he co-founded the openSNP project, enabling people to put their personal genomic and phenotypic data into the public domain. The project relied on crowdsourcing for the data collection process, and applied open science principles by ensuring the data was publicly available, so that it could be reused and its value could be shared by everyone. With over 5,000 public data sets, openSNP is one of the largest public personal genome databases so far.

This initiative spurred Bastian's interest in open and participatory science. He went on to join the Open Humans Foundation, a non-profit organization dedicated to empowering individuals and communities to explore and share their personal data for the purposes of education, health, and research. The organization operates and manages Open Humans, a project and community that enables individuals to collect their own data, from their personal wearables, social media, geolocation, or even genome in one place, and analyse or share it as they see fit.

Bastian is also a research fellow at the Centre for Research Interdisciplinarity (CRI) at the University of Paris, INSERM, where he heads the Peer-Produced Research Lab, working on how to enable the co-creation of citizen science projects. His research focuses on understanding how to peer-produce knowledge and citizen science where people are more involved in the goal setting and design process, and not just data collection and data processing.

Although Quantified Flu aligns very closely with the research interests of Bastian's research group, focused on the co-creation of peer-generated knowledge, Quantified Flu originated directly from the Open Humans community. The academic researcher's team played mostly a supporting and facilitating role.

Project motivation: The trigger was the beginning of the Covid-19 pandemic, and the rising individual interest among Open Humans community members in understanding how useful wearables data might be in the context of this pandemic. The project itself started out from a single community call in early March 2020. The community generated the hypothesis that wearable devices might be able to predict when people are falling sick, even before they are consciously aware of symptoms. This hypothesis is backed by some preliminary research findings indicating that physiological wearable data might be predictive of infection, for example the resting heart rate was found to increase when coming down with a cold.

Landscape: Spurred by the COVID-19 pandemic, research teams around the globe are now trying to use wearable data for predicting infections on a more individual level. Several teams working on this can be found at the Universities of California San Francisco, San Diego, Stanford, Scripps, the Robert-Koch-Institute (the German CDC equivalent) and in many more places. But all of these initiatives use a more traditional, top-down, academic-led research approach, whereby participants typically provide data but are not engaged in other steps of the research, and are not provided access to either raw or aggregated data.

Quantified Flu thus stands in stark contrast to these initiatives, by being community-driven, based on the large interest of the Quantified self and Open Humans communities in doing this kind of research, and by adopting a strong open source, open science stance, with full commitment to making the data available to individual participants and the larger community.

4.4.3 Design choices and resources

Approach: Quantified Flu starts from the simple premise that we can collectively learn a lot about how our physiological signals respond to infections and which symptoms we display when having an infection. The project finds its roots at the nexus of the rise of a global pandemic, the growing interest in the possibilities afforded by wearable technology data, and a focus on bottom-up, personal science. Quantified Flu uses a co-creation process to develop an application where regular users of wearable technology can provide their physiological data and report on their symptoms. The data are shared within the community of participants and made publicly available for research purposes. Specifically, the app can be used retrospectively, if a participant wants to report on its past illness, allowing it to visualise his/her wearable device data in the context of a potential infection. Alternatively, the participant can use a prospective approach and choose to track symptoms going forward.

Quantified Flu was developed iteratively, following community suggestions and ideation processes. Following an initial brainstorming, the planning and coordination stages to develop Quantified Flu were continued through a dedicated communication channel of the Open Humans community slack. The software development happened through GitHub and the git repository of Quantified Flu.

Quantified Flu stands in contrast to other crowd-sourcing initiatives that aim to predict illnesses at the population level, by focusing on the individual scale and by engaging people in co-creating symptom tracking efforts and allowing them to explore and learn from their own data.

Principles: Quantified Flu uses Open Humans for the data storage and management of the wearable device data. Open Humans provides a digital toolkit, which offers support for self-research, tools for personal data access, aggregation, and analysis. It already provides integrations for a number of wearable devices (Fitbit, Oura Ring, Withings, Google Fit) and was extended within the project to support additional devices (Apple Watches, Garmin). For instance, a feature that was added following participants discussions was to allow users to opt-in to receive daily symptoms report reminders sent by email at a user-selected time.

For the collection of the symptoms and additional metadata, the project team designed the QuantifiedFlu website. The website can pull in the individual wearable data and allows people to easily add their daily symptom recordings once per day. Users can also report currently experienced symptoms through QuantifiedFlu at any moment in time by selecting symptoms and their experienced strengths from a list (prospective symptom tracking update was released on March 24, 2020). Users can set up their own reminder time slots in which they will get a notification email to that end. The website also allows people to annotate their sickness events to perform retrospective analyses and provide data visualizations.

Contributors to this project can store their aggregated data privately in Open Humans and then consent to these data being used more widely. The website gives all contributors the choice to make data public at any point to allow others to reuse this data. Data made public that way can be aggregated and re-used through APIs which are currently in active development. As for the Open Humans platform, all the code for Quantified Flu is open source at https://github.com/openhumans/quantified-flu

Data accessibility: Quantified Flu emerged from the Open Humans community-based platform, and accordingly was designed according to Open Humans policy, notably regarding ethics, data protection and data sharing issues. Notably, to be publicly shared, projects first need to be approved by the community, based on a communal validation and negotiated coordination review process, from members who choose to participate. The process usually results in a dialogue with the project leaders, who may need to adapt their project to answer the concerns expressed by the community. Additionally, projects on Open Humans need to adhere to community guidelines, whereby projects need to inform participants about the level of data access they would request, how the data would be used, and security precautions they have in place. Data-sharing decisions remain in the hands of individual members.

Data protection: data privacy issues are effectively addressed through the Open Humans platform, where all the data are stored anonymously and encrypted on the cloud. Contributors can always review and access their data, and have the option to publicly share it via random ID. Access to data requires individual consent each time. Consent can be revoked at any time, even for data that have already been shared. End-users would then receive a notification and the corresponding data would be deleted.

Data validation: In most cases, the data is directly downloaded from the wearable device, and there is no data validation process on the project end. Open Humans provide the raw data, as it is directly downloaded from the wearable or other device, without any data validation step. There are two reasons for refraining from editing the data: first, people have different interests, and any processing would result in some data loss. Furthermore, the project team does not necessarily have the expertise for all the different data types that can be uploaded through Open Humans, and would not necessarily know how to validate them. So although this means a little more work may be required by the end-user to process the data, in this way they are sure that no information has been inadvertently lost or filtered out. Currently, Quantified Flu does not provide an option to correct or edit past entries, but this is planned in future updates.

Resources: About 50 people were highly engaged in the co-creation process and forum discussions, with 10-15 people actually contributing to its development (writing the codes, setting up the software, making the data visualisations). The development team comprised students, who had time on their hands and were looking to help out, as well as software engineers and a minority of academics, interested in the overarching research questions. Overall, it can be estimated that about 3 FTE were needed over a two months period to set up the project.

The project team did not have any specific funding to start the project, and benefited from in-kind support from the CRI institute (Centre de Recherches Interdisciplinaires) in Paris (France) and Open Humans, allowing them to work on this project. Infrastructure costs were covered by Open Humans, and an additional micro-grant from the Open Covid-19 Initiative was secured for software development.

4.4.4 Participation and engagement

Participants: Participants to Quantified Flu comprise a mix of academics, communities, and individuals. Open Humans came out of the Harvard Open genome project, which bestowed academic links, but in recent years, it has opened more to individual communities, such as patients communities. Given the anonymity granted to its users, little information is available to describe the volunteer and user community.

Data visualization: Providing users feedback and ways to better understand, explore and learn from their data is a cornerstone of the Quantified Flu project. For this reason, dedicated efforts were made to create interactive visualisations, presenting the evolution of the physiological data points through time, in the context of the symptoms report.

In the retrospective analysis, users can select the date at which they fell sick or had specific symptoms, and Quantified Flu will extract the wearable data for the period before and after. To facilitate the interpretation of changes and outliers in the graphs, both the first and second standard deviations are presented as well.

In the prospective analysis, following entering their symptom reports, users are automatically taken to their data visualization, which provides the same details as the retrospective analysis. Additionally, this latter view aligns a heatmap of each daily symptom report to the wearable data timeline, allowing the identification of patterns within the reported symptoms themselves, as well as for visual cross-comparisons between the physiological data and the symptom reports. Furthermore, users can also access their comments for each symptom report from this visualization, allowing them to understand the contexts in which they made those reports.

Engagement and outreach: Quantified Flu achieved unusually high engagement levels for an online community. In December 2020, around half of the total 200 users of the platform reported their symptoms regularly, leading to 45 symptom reports per day on average (±5). Over 50 users reported their symptoms consistently for more than three months, and up to a year.

The interest and engagement in Quantified Flu snowballed from the Open Humans community. No particular outreach efforts were made, beyond sharing progress on social media and in the Open Humans forum. Links were also made with other related communities, such as the Open Covid 19 community that grew in the past year from the biohacking space. Quantified Flu thus grew from a core community with shared interest, where people brought in their non-professional interests. The ongoing discussions in the forum are evidence that participants appear satisfied from Quantified Flu.

Usually, participant retention is quite low in crowd-sourced online projects, with most people downloading the app, trying it once or twice, and never using it again. In contrast, Quantified Flu achieved relatively high user engagement, sustained over time. Bastian reckons that this is a result of the community co-creation process and the collaborative design of the app. Retention rates are much higher if you work with users to design an app that fits their needs, instead of assuming that as a researcher, you know what they are looking for. Additionally, the bottom-up approach engenders trust, because the research comes from the community and not from an institution that is primarily interested in collecting the data.

What did the participants gain? However, the patients community is manifesting increasing interest in both goal setting and implementing research. Patients, in particular those with chronic diseases who have several years of lived experience and very strong incentives and drive to experiment, bring innovative ideas and expertise which are not prevalent in traditional academic research. They are able, through personal science, to iterate much faster upon their ideas than traditional academic research, which typically operates on time frames cadenced by the application and obtention of grants.

4.4.5 Outcomes and lessons-learnt

Outcomes: The mere fact that Quantified Flu is still being used, and that people are still engaged is in itself a proof of concept of the value of co-creation processes following a bottom-up approach. Even though the requirements are demanding, with users asked to log in every day, some people have been doing it for over a year.

The Quantified Flu data are used mostly by the community of contributors and much less than initially expected by traditional research. Bastian concedes that the vast majority of academics or pharmaceutical companies have little interest in re-using existing data, in particular if this requires working and interacting with the community and potential adjustments to their research protocols. Moreover, they typically need large scale studies, for which community engagement does not lend itself very easily.

In the future, it is hoped that the contributor community will be the first one to take full advantage of the opportunity to perform individualised longitudinal symptom tracking, as they have already demonstrated a strong interest in this.

Did you meet your goals? Quantified Flu achieved very satisfactory participation levels for a crowd-sourced, community-based project, without any dedicated marketing effort. The research team had hoped to scale the project on a larger scale than it did, but with so many things going on at the beginning of the pandemic, it was hard to get people's attention. The free text in the symptoms reports highlights this well, with people reporting potential confounding aspects related to their situation, working from home or sometimes even loosing their job or a close person. On the other hand, by virtue of being smaller-scale and personalised, Quantified Flu offers a unique opportunity to better understand and thus interpret the data. It is even possible to go back to the contributors (through their anonymous identifiers) to ask for further explanations or details about the context of certain entries. Although there are other similar efforts developed on wider scales, such as the smartwatch app launched by Germany's public authority (Corona-Datenspende), which achieved over 1 million users, about 50% of which regularly donating their wearable data, these do not allow to put the symptoms in context.

Quantified Flu was developed as an ad-hoc opportunity that leveraged two individual drivers: the heightened personal awareness of individuals of whether they are getting sick, and people's expanding interest in self-monitoring, with the aim to make the latter more rewarding and valuable. Sustainability and continuation aspects were thus not considered from the onset.

Successes: From a governance perspective, the fact that Quantified Flu was NGO-led meant that it could be developed faster than it would have been possible through research institutions. In particular, the ethics approval process in academia, while entirely valid from an academic research perspective, is not conducive to the involvement of people in the goal-setting and design stages of the projects.

Quantified Flu succeeded in reaching its audience, and was used more broadly than initially envisioned. Although Quantified Flu was initially developed and used for infectious diseases in general, including Covid-19, people use it for symptom tracking more broadly. In particular, Quantified Flu is used by patient communities with other chronic infections, such as hay fever, since the symptoms match fairly well. This is the type of impact the research team was hoping for, showcasing that community-driven processes are a viable approach for designing mobile health tools and interventions that both end up being used and are useful for the people that they target. Now that evidence shows Quantified Flu works for the people interested in using this tool, it would be possible to scale up.

Sustainability: While there are no immediate plans to scale up Quantified Flu, the research team recognises that a more generic symptom tracking tool would be useful. Symptom tracking is a recurrent need in patient communities, and the possibility to ensure that data are standardised between different patient communities allows to pool and aggregate them more easily for analysis. Although there are other commercially-owned citizen-science initiatives allowing this, such as Patients like me, they often inspire skepticism among the patient community, afraid of how their data will be used. Initiatives such as Quantified Flu also open up opportunities for self-assembly of patient communities, and for pooling and aggregating patients data according to their needs, which may not be available in more traditional citizen science projects.

Lessons-learnt: The main lesson from both Quantified Flu and the Covid-19 survey is the importance of engaging the people you target from the start. Bastian and colleagues argue that although it may draw out the development process, especially when many people are involved, it pays off when the outcomes end up being useful and allow to generate innovative insights in the end. Collaborative, personal science offers an unmatched opportunity to steer people's ideas into traditional academic research, and enables people in similar context to support each other.

On a broader level, experience from Bastian's research group with peer-to-peer, bottom-up science shows that it demystifies science. Involving people in the goal setting allows them to realise that science is accessible and that it is possible for anyone to address very simple scientific questions. They also become more critical, in the sense that they understand the process that leads to scientific findings, and the iterations it may take to revise hypotheses and protocols and break down questions into manageable chunks. people are sceptical of the published literature and want to see if they can replicate the findings for themselves.

4.5 Covid Open Survey

4.5.1 About the project

The OpenCovid survey is a short questionnaire created for people who want to help fight the COVID-19 virus from the comfort of their home. After an initial survey, participants are asked daily questions about on they are feeling. A wide variety of questions are asked, some of which can be answered daily, others that appear on a rolling basis, so as to always keep it fresh.

The overarching goal of the Covid Open Survey is to better understand the transmission patterns of the virus and its effect on physical and mental health, so as to guide future decisions surrounding public health.

4.5.2 Context and motivations

Research group and background: Doctor Bastian Greshake Tzovaras, as well as being acting director of research at Open Humans Foundation and instrumental in the development of Quantified Flu (see previous case study), is also a research fellow at the Centre for Research Interdisciplinarity (CRI) at the University of Paris, INSERM. At the CRI, Bastian heads the Peer-Produced Research Lab, working on how to enable the co-creation of citizen science projects. His research focuses on understanding how to peer-produce knowledge and citizen science where people are more involved in the goal setting and design process, and not just data collection and data processing.

Project motivation and goals: The Covid Open Survey project was initiated through the desire of CRI to launch a citizen science initiative around the Covid-19 pandemic issues. The research team was interested in doing something similar to Quantified Flu on a survey-basis, and in ensuring that the data goes back to the community, instead of just being for academic use. This was unusual in the landscape of health citizen science projects, which typically comprises projects that fall under the appellation because they recruit participants online, even though they do not provide any further opportunity for the participants to engage in the scientific or knowledge process.

The Covid Open Survey had multiple research aims from the start, including:

- To collect comprehensive symptom reports with a high temporal resolution to better understand COVID-19 disease progression.
- To measure the impact of the COVID-19 pandemic on mental health.
- To identify effective agents to reduce COVID-19 disease severity
- To characterize factors linked to increased susceptibility to COVID-19
- To understand how a more participatory approach to survey research - that makes participants research partners - changes survey success.

Both the Covid Open Survey and the Quantified Flu project started at the end of March 2020.

Partners: For the project, Bastian's group initially collaborated with partners at the Australian National University, the University of Sydney and at the University of Notre Dame in the USA, who were all interested in developing a citizen science survey where people provide input in the survey design, notably on which questions should be asked. The partnership eventually expanded to become more transdisciplinary and cover more countries, with researchers from Canada, Costa-Rica, India and more hopping on. All partners had already collaborated with at least another member of the team in some form before.

Competition: The research team was aware about similar initiatives that got underway in the same time period, but the Covid Open Survey always stood apart by the fact that it would enable people to give their own input into the survey. Unfortunately, this never materialised, because the initial survey designed by the research team ended up being so large that people dropped out very quickly and did not see the point of adding yet more questions. Direct feedback from a couple of participants confirmed that the initial survey, which took just under half an hour, was just too much.

4.5.3 Design choices and resources

Approach: Covid Open Survey consists of an online questionnaire, followed by regular daily checks of how you are feeling. Although the project brands itself as a five-minute questionnaire people can fill from the comfort of their home, the initial questionnaire ended up longer than expected and took about 20 minutes to complete. The daily checks comprise some fixed questions and a subset that appear on a rolling basis to keep the survey fresh. The idea was to allow people to formulate their own questions for the survey, so that it would answer issues they were interested in.

Survey design: The Covid Open Survey started in the academic realms and although most of the researchers involved had some previous experience with citizen science, it usually involved more traditional in-person or on-site activities, in contrast to the Covid Open Survey which was fully remote. The incubation time for developing the survey took about six months. With few exceptions, the team had no previous experience or expertise in questionnaire design. This meant a high learning curve was needed at the start to think about what questions to ask and how to design the survey. Additionally, the diversity of the team's research interests, covering mental health, genetics as well as pharmaceutical side-effects, meant that the project quickly ballooned into a lot of research questions. The number of academic institutions involved also meant ethics approval was needed in each separate institution, which drew out the process further.

As a result, the survey could only be launched in August 2020, by which time most of the team's initial excitement had worn off. The Covid Open Survey was a side project for everyone involved, with no dedicated funding.

Resources: The total time spent on the project is difficult to estimate, as the work was a side project for everyone and happened in bursts, according to the needs - be it agreement on survey questions, ethics approval process, or other. No dedicated funding was available for the project, but the Open Clinical survey tool provided in-kind support by making their software platform much cheaper for running the survey, the rest being complemented through discretionary funds from CRI.

4.5.4 Outcomes and lessons-learnt

The Covid Open Survey did not lead to any tangible outcomes and fell short of its initial goals. The main success was that the project did manage to be launched at all. It also engendered broad international collaborations, inclusive of different perspectives. But in hindsight, Bastian reckons that the ad-hoc partnership formation was probably one of the downfalls of the project. Starting smaller, through a more community-driven process might have helped the project take off faster and meet its audience. The error was in the design process, where the more traditional academic research interests took over the participatory process. Although the project cornerstone was the community involvement in the survey design, this was not given any space in the initial project development. The inclusiveness of the project, with the trans-disciplinarity and diversity of perspectives represented, was also its silver lining: every researcher had its own questions it wanted to see answered, resulting in an overly long initial survey and an overly drawn-out design process. Together, these factors led to reduced volunteer engagement and withered the research's team drive, given this was a side project.

4.6 CovidWatcher

4.6.1 About the project

Covid Watcher is a research, advocacy and policy platform developed by Columbia University in collaboration with New York community partners. It aims to support New York City officials, community organisations and hospitals to identify the needs and target resources to deal with the impacts of the Covid-19 pandemic. Ultimately, it hopes to inform the development of more effective policies that ensure the health, safety and economic vitality of the city as it reopens.

Participants can respond to the CovidWatcher survey on the project website as well as via the CovidWatcher app- available on iOS and Android. The project comprises a baseline survey, and four distinct surveys concerning Medical needs (medical, mental), Resource needs (access; employment, safety, education, information); Covid-19 symptoms and testing (symptoms, exposure, testing; smell); personal behaviour (social distancing, seeking info). More recently, a survey for small business owners was also added.

4.6.2 Context and motivations

As the covid pandemic ripped through New York City, challenging medical resources as well as all aspects of society, the local scientific community quickly came together to aid in understanding and problem solving the novel coronavirus pandemic.

Research group and background:

CovidWatcher was initiated by Noémie Elhadad, Associate Professor in the department of biomedical informatics at Columbia University (New York), together with a handful of colleagues within her faculty with a strong background in computational biology across different disciplines (biomedical informatics, systems biology, and infectious diseases). Noémie's own research investigates ways in which electronic health records and patient-generated data (e.g., online health community discussions, mobile health data) can help improve healthcare and the health of patients. It is thus fairly natural that the team envisioned developing a data-collection app and website for tracking the impacts of the pandemic in New York City.

In mid-March 2020, at the beginning of the lockdown in New York City, the researchers created a slack team and recruited students to help out. The drive was the strong disconnect between the data provided by the public health agencies, coming from hospitals, and the reality on the ground. The surge in positive cases was such that very quickly the wards were overwhelmed, and there was no testing going on. The original idea was to develop a crowd-sourcing platform that would survey New Yorkers about their exposure to the virus, their symptoms and how they felt, thereby helping to fill important data gaps.

Within a few days, the team realised that possibly this was not where they could best contribute. New York City was impacted in a way never witnessed before: within a few days, there was no more food in grocery stores. People had a lot of instrumental questions around the pandemic. The researchers thus swiftly pivoted from a symptom tracking-type questionnaire to more material questions, asking people whether they had access to food, transportation, or other basic needs. They also surveyed people about their mental health.

This shift fundamentally changed the way the researchers went about engaging participants. They needed a more representative sample from the population, including from underserved communities. More importantly, they needed to make the survey instrumental and put this information to good use. Within two weeks, they earnestly contacted several public authorities in New York (e.g. the department of health, of education, New York State). When it became obvious that both the State and the city were overwhelmed and did not have the capacity to partake in new initiatives, the team decided to serve the communities directly, by working through community-based organisations. Esther Fuchs, professor of international and public affairs at Columbia joined to reinforce the team to lead the community engagement process and ensure it was inclusive. Gradually, the aims of the project crystallised around how to channel this information to guide resources (food, ambulances, etc.) where they were most needed, and to enable city officials and medical centers better coordinate their response.

Landscape: Early on in the pandemic, Noémie Elhadad recalls that enthusiasm was high among researchers around the world. Her group got in touch with numerous other research teams, but in the end, funding and publication concerns often seemed to prevail, such that people would go and launch their own study. This illustrates the difficulty in establishing new collaborations from scratch, if there was no prior working relationship.

These discussions led the group to focus on New York City, and to expand its focus beyond health issues, in order to differentiate itself from other prominent projects tracking the symptoms and spread of Covid-19, such as a Harvard-led project, or the “Stop covid NYC”, a web-based app launched by Mount Sinai Health System, New York’s largest academic medical system. CovidWatcher demarcates itself from other coronavirus apps and trackers by including surveys on concerns about education, access to resources, transportation, and mental health. A number of add-ons were also developed, to make use of the sensors on the participant, and track their spatial locations, or physiological data from their wearables, such as heart rate and body temperature.

Partners: All core team members already knew each other and had collaborated before. This was necessary, given the long work hours involved. After CovidWatcher pivoted to more material issues, a researcher from urban policy was brought in to support the core team. She fit in seamlessly, and brought her experience and existing links with New York City community organisations, allowing CovidWatcher to connect to the city's high-risk and underserved populations, and to ensure that the data were reliable and inclusive.

The team developed strong community partnerships with New York advocacy organizations, but also universities, government agencies and elected officials that helped disseminate the CovidWatcher survey in their constituent communities. These partnerships were essential to fill critical gaps in knowledge as new issues arose in the city. Community Partners include: NewYork-Presbyterian, ColumbiaDoctors, Clinical Directors Network, Older Adults Technology Services, Human Services Council, Harlem- 125th Street Business Improvement District*, The Pride Agenda, Jewish Community Relations Council of New York, Common Cause of New York, UJA Federation of New York, Academy of Political Science, Citizens' Committee for Children of New York, Community Service Society, Restoration, FPWA, Association for a Better New York, Gale Brewer, Mark Levine, United Neighborhood Houses, Citizens Union, Civic Hall, Association to Benefit Children, NYC Hospitality Alliance, Tech NYC, NYC Planning, The Manhattan Chamber of Commerce, and Partnership for New York City.

The team also received some support from industry volunteers, including Facebook, Google, Elektra labs, Redscout.

Project motivation and launch:

The initial trigger for the project was the speed at which everything that was taken for granted seemed to fall apart in New York. It seemed that maybe technology and citizens could help to problem-solve.

The decision to focus on New York City took shape quickly: New York was the first city in the United States to be widely affected by the Covid-19, and although the team approached health departments in other cities, these were unconcerned, and considered it a New York issue. New York was where the need was at that time, and it would allow the project to be more actionable at the community level.

By mid-April, the project had fully redesigned itself, and an App both for androids and iPhone was launched, as well as a web-based version. The unprecedented speed at which the app was developed was possible through the collaborative energy that sprang from all sides. The project very quickly mustered a large team of volunteers from the scientific community, including both faculty members and students. People at Google and Facebook volunteered their time to help out. The community-based organisations were also galvanised by the sense of urgency, and the project established collaborations with over 15 (as detailed above) of them in that short timeframe.

4.6.3 Design choices and resources

Approach: The rationale for CovidWatcher is to inform governments, businesses, philanthropy, community and advocacy groups about what is and is not working for people and communities, so that they can act more strategically and make appropriate decisions about healthcare, jobs and social services. It consists in a panel of surveys specifically designed to fill in the gaps in knowledge on Covid-19 is impacting New Yorkers. Surveys are anonymous, offered in multiple languages, and available on the web and via the Covid-19 app-. CovidWatcher relies on community partners to ensure it reaches all income levels, communities, ages, racial and ethnic backgrounds.

Why citizen science? The decision to develop a citizen science project derived from the altruistic feeling that people deserved a platform to let everyone know what the issues were in their community. This was unusual for New York City which is not characterised by tight community spirit. The idea was that in return for the information they provide, citizens would gain insights about the impacts of the pandemic in the city.

The intention was to feed the results back to the community, through real-time, interactive visualisations. Infectious disease experts and public health officials could then use the visualizations to monitor trends and plan accordingly. However, this required substantial investment and the team did not manage to secure the funding to finance this venture.

Survey design: The project team had prior experience in designing surveys, which meant it was able to rapidly hone in on the major issues in a given topic. The survey was designed in a modular fashion, with different modules targeting different issues and audiences, allowing the surveys to remain relatively short. Most of the questions stemmed from the team and were based on standard instruments, but a sizable amount was also designed ad-hoc, in response to the emerging needs, or to issues that emerged from the community. The modules related to family and education for instance initially stemmed from a request of Columbia University, to try and understand whether students would come back on campus.

Noémie Elhadad recollects how incredible it was to witness that the situation was changing at such a pace that CovidWatch questions regularly became meaningless. Access to food was a pressing issue at the start of the pandemic for example, but within two months it had dissipated. The same happened with access to surgical masks. One of the project's initial aims was to identify hotspots of infection in the city, so this could be used to send ambulances in these areas. But within a couple of weeks, the situation had worsened to the point that it was clear this would be a worthless effort. Frequently, the project team would come up with a whole set of questions that would have to be discarded before even being used, because they had become irrelevant. New issues arose all the time, and it was hard to predict what they would be. For instance, the team thought about launching a survey investigating people's access to testing facilities. But until June-July, there were no testing facilities, so the question could not be asked. In some cases, the project was able to pick up early signs of emerging issues, such as with the spike in anti-Asian discrimination in New York city that got picked up through the Safety survey.

The research team kept adapting the survey, to respond to new and emerging concerns, but maintained the previous survey questions. Contributors were free to choose which sections to answer, and it was very telling to see which modules people were interested in answering at different times. A common comment from contributors was that the number of questions was overwhelming.

App development: In addition, the team developed additional functionalities in the Covid Watcher App. It was able to develop these quickly and efficiently by building on its previous experience in developing similar functionalities in the past. The functionalities were developed in a very modular way, with informed consent and appropriate measures to protect user privacy. They included three passive sensing options: heart rate and body temperature monitoring for wearable watch users, or how much people get in and out of their home, by including a geofence allowing to determine whether a person was within 200 meters of its home location or had travelled outside the 200 m perimeter.

Data validation: The Covid Watcher team considered data quality issues carefully in the development of the surveys. A lot of standard instruments traditionally used for surveys were found to be inappropriate in the Covid-19 context. In considering how to develop new surveys that would ask the relevant questions at appropriate frequency, the researchers opted for a pragmatic approach. They developed questions and adapted them based on the feedback received from the community-based organisations. Noémie Elhadad reflects that several academic researchers across the country faced similar issues, which led to some fruitful exchanges and progress in the field.

Data accessibility: Although the project applied an open data policy, it exercised it with caution due to the sensitivity of the data and risk for unintended consequences. All personal data were collected with the informed consent of the participants and anonymised. The de-identified data were open to community-based organisations, as well as for research, provided the researchers were able to justify a good use-case. Aggregated forms of the data can readily be shared upon request.

Resources: All work was performed on a volunteer basis. By the time the project started, regular work had slowed down dramatically, which meant that the core team was able to spend time on the project on the side. About half of Noémie Elhadad's lab members worked on the project full time, while the rest of her lab members had the occasional involvement. Discretionary funding from the department financed the non-trivial costs for the development of the website, app. and database. Additionally, Columbia University was much more willing to help go through the small administrative hurdles of research projects, including ethics and protocol reviews, which helped the project get going quickly.

4.6.4 Participation and engagement

Participants and engagement: The project managed to secure about 5,000 active CovidWatcher users. It took some adaptation for the team to find the best strategy to secure engagement. The large majority of participants preferred the website survey, and was not interested in the app, even though it offered nifty functionalities in addition to the survey. It was also possible to secure longer engagement of the app users by sending them notifications to remind them to participate. But it was hard to convince people to use the app in the first place. Recognising that the populations the project was trying to reach were not app-savvy, the team found that the most effective way to ensure engagement was to communicate through the community organisations, and to regularly remind people by emails.

The main communication interface for the project was through the project email, to which the project team did its best to respond in an accurate and timely manner. The team also participated in numerous meetings with the community in an effort to secure engagement and share key project outcomes. It made calls to action through Columbia university, town hall meetings, as well as through specific meetings with community organisations, where the constituents could give their feedback. Although this approach proved effective, it required a lot of time on the part of the team, and only reached the people that were already interested in the first place.

What did the participants gain? Deliberate efforts were made to communicate project outcomes and make the data available to support local decision-making. The data collected were mapped geographically to help community partners identify viral hotspots in New-York city. Although dynamic visualisations of further project results were initially planned, these did not materialise due to lack of funding. But Covid Watcher data were made available in aggregate form to community organisations, city officials and healthcare institutions. Detailed reports were also shared with the community organisations so that they could share the results with their constituents.

Outreach: All outreach efforts were carried out by the project team, including reaching out to community organisations and city officials. Columbia university's public relations service did not support the project launch, through dissemination or advertisement activities, as the University required some preliminary outcomes before considering official backing. However, the Covid Watcher app received official backing from Columbia University, a process which normally takes a long time, which helped establish its legitimacy.

4.6.5 Outcomes and lessons-learnt

Outcomes: Covid Watcher did not fully succeed in holding its promise to support decision-making and mitigate the socio-economic impacts of Covid-19 spread in New York city. City officials already had established channels to collect data and were not willing to change their ways of working. Community organisations were successfully engaged, but for several, lack of funding limited their ability to use and analyse the data. At the time, the Covid Watcher team itself did not have the capacity to perform tailored data analyses. The project survey directed at small business owners was perhaps the most successful, because the businesses had a strong need to be heard and the capacity to make it happen. The survey analysed the long-term impact of the pandemic on small-businesses, with questions on their needs as New-York was reopening. The research is still ongoing, but the results have already been quite useful, and transferred to officials in New York City and in business development initiatives.

From an academic perspective, the project did not lead to any research publications, although there would be scope for reporting on the community-based set up of the project.

Lessons learnt: Despite the mixed outcomes, the team was able to learn from this experience and to carve the less successful pieces out to develop new projects. For instance, although the app did not end up being used as much as anticipated, the development work did not go to waste: it became the basis for the development of a handful of new apps, including the in-house contact tracing app developed by Columbia University, and an NIH funded project. The greatest disappointment for the team was probably its failure to secure funding to feed back the results to the community. CovidWatcher shows that it can take some time for the project to find the niche where it will deliver added value. It took a gradual process, involving discussions with other academics and city officials, to redirect the project to the local communities and to the material needs they faced in the pandemic. Although this was not a straight path, Noémie Elhadad confesses that she is not sure there was another way for them to do this. Lessons learnt in one population do not necessarily fully translate in another context. Nevertheless, drawing on her accomplishments from previous citizen science projects and from the small business survey, Noémie postulates that projects tend to be more successful when the communities they target have a strong need to be heard.

On a more personal side, Noémie Elhadad acknowledges that the project completely changed the way she sees her research. It broadened her view of her research questions, to ensure they are more pertinent and more relevant to the end goal. It reinforced her motivation to involve and work with the community, something she had already started doing with previous projects, building on advocacy groups, community organisations or social networks such as reddit. More generally, the CovidWatcher experience spurred her research to be more multi-disciplinary, and identify collaborators able to bring a complementary approach. She admits that this makes the research more complex, but is confident that it allows for more meaningful outcomes. It also changed the way she mentors her students: each lab member is encouraged to have a side-project where they volunteer to contribute to the community.

4.7 Outbreaks Near Me

4.7.1 About the project

Outbreaks Near Me is the next generation of the Flu Near You platform that was created in 2012 and maintained by the Boston Children's Hospital team (USA). Outbreaks Near Me expands the Flu Near You platform to all respiratory diseases, to supply public health officials and academics with real-time information that could help stop the current pandemic, or prevent the next one.

Created by epidemiologists and software developers at Harvard, Boston Children's Hospital and a group of volunteers from across the technology industry, Outbreaks Near Me uses crowdsourced data to visualize maps to help citizens and public health agencies identify current and potential hotspots for the recent pandemic coronavirus, COVID-19, and the annual influenza. Outbreaks Near Me relies on voluntary participation from the general public, to report if they or their family members are sick or healthy. Those who are feeling sick are asked to describe their symptoms, recent travel history, and healthcare information.

4.7.2 Context and motivations

Research group and background:

Autumn Getz is pursuing her MSc in Epidemiology at Boston University School of Public Health, interested in emerging infectious diseases, and the mechanisms of human interaction with the disease. She joined HealthMap in the fall of 2019 and is the current project manager for Outbreaks near me, responsible for coordinating stakeholders, users, and the development team.

"Outbreaks Near Me" was born out of the successful "Flu Near You" project, that tracks and maps the distribution of influenza-like diseases in the United States and Canada. Flu near You was created in 2011, through collaboration between HealthMap at Boston Children's Hospital, American Public Health Association (APHA) and Skoll Global Threats Fund, with the aim to help prevent the next influenza pandemic by providing an early-detection system. It relies on voluntary participation from the general public through a website or mobile app. Registered participants are asked to take a few seconds each week to report on the influenza-related symptoms. Flu Near You displays the aggregated weekly results in interactive maps, and provides an interface to compare its data with data from the Centers for Disease Control and Prevention (CDC) sentinel influenza network. Users are also able to view the number of people reporting and the percentage with influenza symptoms statewide; they can also locate their nearest vaccination center and get basic information about vaccines. The long-term goal of Flu Near You is to provide early warnings of influenza occurrence that provide real-time information to the general public, but also public health officials and healthcare services, and may help prevent the next pandemic. Flu near You, which currently counts nearly half a million participants in the United States has been highly effective at providing an understanding of the spread of influenza-related diseases, providing relevant and complementary signal for influenza surveillance when compared with national surveillance data, including generating heat maps that identify emerging pockets of illness.

As soon as December 2019, the Flu Near You team within the Computational Epidemiology lab at Boston Children's Hospital started thinking about adapting the Flu Near You tool to monitor Covid-19. This early alertness was possible since some of the researchers in the team are also involved in Health Map, an online tracking tool for emerging infectious diseases across the world. Health Map picked up evidence of clusters of respiratory illness around Wuhan at the end of December 2019. Initially, the team planned to adapt the Flu Near You system, to also allow for Covid-19 symptoms tracking. But since Flu Near You was built on older technology, the changes were not as easy as expected. Accordingly, the team decided to create an independent sister platform dedicated to Covid-19 tracking, Covid Near You. This would also allow to bring new volunteers to self-reporting. Covid Near You was launched in March 2020, and is very similar in its structure to Flu Near You. Through a user-friendly interface, participants are asked to provide daily updates on Covid-19 related symptoms, or report feeling healthy. Covid Near You then maps this information to provide real-time, interactive, zoomable heat maps of the spread of Covid-19.

In December 2020, the Flu Near You and Covid Near You platforms were merged into a single disease tracking platform that uses citizen science and participatory surveillance to detect respiratory illnesses. The new platform was rebranded “Outbreaks Near Me”. The process started in the summer of 2020, as the team increasingly recognised the value of merging its two platforms, which had some overlap since a few Covid-19 related questions had been added to the Flu Near You platform and their data were integrated. Moreover, a platform with a broader scope would be more adaptable, and allow to track any potential future respiratory epidemic. The team had in any case been contemplating an update of the Flu Near You platform for some time, so it seized the opportunity when it received interest in developing a single platform for respiratory diseases from potential funders.

Initial motivation:

Project motivation: The gap in Covid-19 testing at the beginning of the pandemic was one of the key motivations behind the launch of the Covid Near You platform. The idea was that until both healthy and symptomatic people can be tested at scale, citizen science can help fill the gaps in monitoring the spread of Covid-19. Discussions between Professor John Brownstein, co-creator of Flu Near You and Chief innovation officer at the Boston Children Hospital and Prem Ramaswami, head of product at Alphabet’s Sidewalk

Labs, who had collaborated before, led to the idea of co-opting the Flu near you technology for better Covid-19 monitoring and to the creation of the Covid Near You platform.

The development of Outbreaks near me was in continuity with the existing work of the Flu Near You team. The cornerstone for the effectiveness of these projects is the continued engagement of participants, which allows to have longitudinal data for surveillance. The team thus built on the elements that made the success of Flu near you, by designing very short and focused surveys. "It is a matter of finding the right people and not requiring too much of them", affirms Autumn Getz, "then people will continue to report even after the pandemic". This approach differentiates Outbreaks near me from many other self-reported symptom tracking surveys, which tend to be very extensive and exclusively focused on Covid-19.

Landscape and partners: The overall spirit has been collaborative, and the team already had strong established research network and collaborations through its past projects. Outbreaks Near Me team is in particular collaborating with a few other groups. It has a partnership with Momentive (previously SurveyMonkey), which hosts on its website a longer version of the Outbreaks Near Me survey, that includes questions about attitudes, feelings and beliefs. The Outbreaks Near Me team is also exchanging with other research groups about the instruments, to gauge the questions they are asking and how, as well as the type of responses they get. Outbreaks near me team also collaborated with the University of Maryland and Carnegie Mellon University for their Covid-19 Trends and Impact Survey, that covers aspects related to symptoms, as well as mental health, economic conditions and demographics of respondents, and has been promoted in partnership with Facebook.

In the past year and half, Outbreaks near me also strengthened its partnerships with the network of participatory surveillance systems tracking flu in other countries, as they worked closely together on how to adapt the flu tracking systems to Covid-19. Outbreaks Near me is in regular contact with colleagues in some of these networks (e.g. FluTrackers and Flunet in the Eu, Australia and New Zealand), and compares data and outcomes with them.

Autumn Getz sees a good potential for broader collaborations in the future. Several groups may not have the resources, or funding to continue to support their survey in the longer term, and after the pandemic, the Outbreaks Near Me platform may be able to host their work and ensure its legacy.

4.7.3 Design choices and resources

Why citizen science? Outbreaks Near Me defines itself as an online crowdsourced, or participatory surveillance system more than as a citizen science project per se. However, Flu Near You is partnered with the citizen science association in the United States.

Arguably, the project may become citizen science for the registered users, who really engage by reporting weekly, responding to users surveys, and monitor the project's outputs to see what comes out of the data. As compared to more traditional disease surveillance systems, the participatory systems have the benefits of speed, sensitivity and scalability of users. They are particularly useful for respiratory diseases such as flu, for which many people do not bother going to the doctor, and who thus go undetected by conventional surveillance strategies. The participatory systems actively engage the general public in reporting and in return, provide them with timely information about the disease and its spread. In exchange for participating, users can visualise the Flu or Covid-19 activity on maps and get links to trusted sources of information.

Another key benefit of participatory surveys is that they allow us to get additional insights from the community. The symptom reporting is only the first layer of necessary data. Additional questions can be asked to the volunteers to provide more in-depth understanding of the pandemic and understand how factors such as mask-wearing, social distancing or access to healthcare influence the spread of the disease, and may vary according to the social or demographic context. The whole idea behind Outbreaks Near Me is to empower people and make them actors in public health.

Survey design: the researchers transferred the general questions from Flu Near You into Covid Near You and then Outbreaks Near Me. They also performed some user interviews to identify the motivations of participants and how much participants are willing to contribute. A couple of questions were added to answer questions about healthcare utilisation of interest to one of the funders (whether people were seeking care when they were sick, and what type of care). The partnership with SurveyMonkey allowed Outbreaks Near Me to develop a specific survey to get greater depth of information about people's testing status, political affiliation and attitudes and beliefs toward issues like mask-wearing, social distancing or travel to assess how they affect Covid-19 transmission.

The team was careful to strike a balance between the baseline survey questions to follow through time and the specific questions to answer more specific research interests, always sensitive to the fact that if the survey was too long or complex, people would not want to report weekly. Some of the questions were thus more standardised, validated instruments, other questions, in particular those related to Covid-19 were developed ad hoc and adapted as needed based on feedback from the users or from other symptom-tracking projects. The Outbreaks Near Me app features a standard set of questions, but the website includes some additional questions about vaccination and testing, and the user can answer in free text to some of these. The team also uses its SurveyMonkey survey to ask a broader range of questions, or to test the specific phrasing of a question.

Open data: the data are available upon request to research organisations, upon providing an explanation of how they will be used and signature of a data transfer agreement. General public users cannot download the data themselves, but have access to the visualisations. Outbreaks Near Me plans to extend these visualisations to allow users to interact with the data in aggregated form.

The reporting is quick, easy and anonymous. Anyone over 13 years of age can participate.

What do the participants get back? The team strives to provide value back to its users by making the collected data understandable and accessible. This is where the partnership with Tableau comes in. Off the main page, Outbreaks near me users can look at the reports which takes them to a dashboard in Tableau allowing to visualise self-reported Covid-19 symptom data, as well as additional information through aggregated metrics according to political affiliation, education level, attitudes about mask-wearing, healthcare access, and more. The aim is to provide a flexible way for the users to query the data, so that they can look at them in the way they are more interested in, at the scale that matters to them, and maybe discover important emerging patterns.

Resources: The Covid Near You project started based on the Flu Near You team's own resources. They were able to get backdated funding which helped them get started. Following on, they secured additional funding, mostly from organisations with whom they had worked with in the past, including federal contracts and grants from nonprofit and philanthropic organisations.

The Covid Near You platform was developed at lightning speed. The group already had a lot of experience in developing public health tools, and could conceive the platform fast and before the confinement started, allowing for more effective collaborations. Thanks to the contributions of over 20 volunteers from the Tech companies like Google, Apple, Amazon and others, the Covid Near You site was up and running within a week.

The development of the Outbreaks Near Me platform was made possible with collaboration from SurveyMonkey, Delphi Group, Symptom Study and the American Public Health Association and thanks to the generous contributions of several donors, including SurveyMonkey, AtScale, AWS, BlazeMeter, Cloudflare, Datadog, MongoDB, TechSoup. Although the site is already operational, the team still plans on developing several additional features.

Altogether, for Covid Near You and then Outbreaks Near Me, it is estimated that about 20 people were involved on a part-time basis, notably for the development aspects, but also for designing and marketing. About seven to eight people are working full time on the projects. It is hoped that within the next year, the Flu Near You platform will have fully migrated on the Outbreaks Near Me platform.

Sustainability: The project was built with sustainability in mind, since the team already has the Flu Near You system running for over ten years. One of the reasons to combine the Covid Near You and Flu Near You platforms was to technologically upgrade them so that they have the capacity and are readily adaptable. In the event of any future outbreak, it will be easy to add new symptoms or syndromes.

The team works according to the funding cycle, ensuring that the system stays operational when funding is low, and investing in improvements and additional features when financial resources are higher. The team is already exploring future funding opportunities, already noticing that public interest in Covid-19 is slowly fading away, and anticipating it will continue. It is in particular looking for ways to keep research and public health interests up, since the regular user base is key.

4.7.4 Participation and engagement

Participation: Outbreaks near me reached up to six million reports at one point in time, but the number of regular users ranges between 20 to 40 thousand every week. In comparison, Flu Near You reaches on average 12,000 contributions per week during flu season, and less off season.

Engagement: High engagement levels and consistent participation of users are essential to accurately track disease over time and estimate burden. The Flu Near You model is built on securing high engagement from a group of registered users, through a simple easy to use interface. Similar principles were applied for the Covid Near You and then Outbreaks Near Me systems. Additionally, Outbreaks Near Me engagement levels are bolstered through regular text reminders to registered users, issued twice a week. Outbreaks Near Me does not have an online user community, but participants can reach the team through its email address. An annual user survey is planned for Outbreaks Near me, similar to the one sent out in Flu Near You, where the team also asks feedback and provides free text for comments.

Participants: the participants are from the United States, Canada and Mexico and seem to reflect similar biases as those observed in the Flu Near You platform. Participation tends to be skewed towards older people that are educated, scientifically literate and health conscious. It is often people that work in the healthcare sector, or who care about health. Outbreaks Near Me wants to continue to diversify its user base, so that it brings more data, that are more representative of the entire population. The partnership with SurveyMonkey , which counts nearly two million daily users on its online polling platform, is already enabling Outbreaks Near Me to broaden its user base and data inputs. In the future, the team hopes to focus on more targeted recruitment towards younger users, rural populations, people with no higher education, as well as increasing racial and ethnic representation.

What did the users gain? Outbreaks Near me platform provides links from its Frequently Asked Questions page to external sources of trusted information on Flu and Covid. The team is planning to send a quarterly newsletter out to participants; in Flu Near You, emails are sent out periodically

Outreach: Outbreaks Near Me team engaged in a lot of media outreach, involving a combination of traditional and social media activities. In particular, John Brownstein of the Boston's Children Hospital, co-founder of Healthmap and Flu Near You, was able to leverage his wide-ranging network and media connections. The Outbreaks Near Me team also communicated to the current Flu Near Me users and encouraged them to participate.

4.7.5 Outcomes and lessons-learnt

Outcomes: The primary outcome of Outbreaks Near Me is assessing the burden of influenza-like and covid-like illnesses. One of the key benefits of participatory systems like Outbreaks Near Me is that they allow earlier and more sensitive detection than conventional systems. The reporting is real-time, whereas it can take weeks for the official surveillance system to identify an outbreak, since it needs to compile and analyse the data from the weekly reports sent by healthcare providers across the country. Participatory systems also allow to capture cases that are not being picked up by traditional surveillance systems, which only pick up the share of patients that go to the doctor.

Additionally, Outbreaks Near Me was able to assess healthcare utilization among the participants reporting symptoms, and estimate the amount of people who might be sick and not seeking care in the population. The project also enables the assessment of vaccination rates for Flu and Covid.

The team has been collaborating with the US Centre for Disease Control and Prevention historically, and the government organisation uses the platform's participatory surveillance data to complement its surveillance system. The joint survey with Momentive (SurveyMonkey) led to policy recommendations and a scientific publication on mask-wearing and the transmission of Covid-19. Some of the data about how people were feeling after receiving the vaccine were also useful to draw recommendations about vaccination

Did you meet your goals? A major satisfaction for the project team was to see that the team had the flexibility, historical knowledge and data to react and provide useful insights in real time during an emergent threat like covid19. The Outbreaks Near Me team did not set out very ambitious goals at the start. The focus was on getting the project operational quickly, so that it could help fill the information gaps. Getting the site up and running early in the pandemic was thus a major success in itself. The team was also very pleased with the levels of engagement achieved on the Covid Near You platform, which exceeded its expectations. It is still a little early to say for the Outbreaks Near Me platform.

The Outbreaks Near Me data have been used throughout the pandemic and help both improve understanding of the disease and inform policy. However, there is still a lot of potential for the Covid-19 tracking data to get used more, reckons Autumn Getz. While the participatory data was very useful to health departments early on in the pandemic, now that many more case counts projects are available, interest for these data is waning. The team plans to proactively reach out to its contacts in different health departments to encourage continued uptake of its data as a complementary source of information.

Next steps: The team is already planning ahead and intends to use Outbreaks Near Me to closely monitor when flu and other respiratory diseases will come again, since they have been depleted during covid. They also intend to assess more metrics, notably in relation to healthcare utilisation, testing and vaccination. Future plans also involve extending the collaboration with the US Centre for Disease Control and Prevention to ensure the results of the platform are more generalisable to the general population, with their support for outreach efforts and the survey extension.

Lessons-learnt: Looking back, if we had been able to build up Flu Near You to collect more Covid-19 data, in addition to the Covid Near You platform, says Autumn Getz, we might have been able to get more representative or sensitive data, and leveraged the more established user base of Flu Near You. Additionally, building in ways to retain Outbreaks Near Me participants from the start would have been helpful, since as the fear of Covid-19 goes down, the platform risks loosing some unique profiles.

4.8 Summary of results

Projects producing citizen science data have increased significantly over the last decade, creating opportunities for capacity-building, public engagement and crowdsourcing of big data, in particular in times of crisis. Not only can this information act as an early warning, which may help to save lives and livelihoods, it also has the potential to generate shared understandings of hazardous phenomena, improve communication and help communities at risk take actions to build their resilience during, after, and in preparation for future hazardous events (Hicks et al., 2018) There is a pressing need to accelerate the use of citizen science data for public good (Balestrini et al., 2021) by understanding the enabling conditions that can improve their emergence.

This case study analysis explores the measures that can help increase the likelihood and relevance of citizen science initiatives to respond to unforeseen challenges, using the Covid-19 pandemic as a case example. Recognised success factors of citizen science initiatives include having clear project goals, stable funding, sustained participant engagement and an appropriate technological infrastructure (Balestrini et al. 2021; Newman et al., 2012), conditions that can be challenging to meet when responding to an emergency. For instance, reaching out to a different community or scaling up all require considerable efforts. Securing funding and technological infrastructure typically operates on fairly long time scales, not commensurate with urgency.

In this analysis, we consider the conception of projects and how citizen participation was considered, as well as how they were designed and the outcomes they achieved. The interviewed project leads also shared their own lessons-learnt from the experience. The overall aim is to enrich the discussion about how citizen science projects can be conducted and what it takes to make them more scalable, spreadable and sustainable (Balestrini et al., 2021).

4.8.1 Crowd-sourced and mobile technology initiatives significantly informed the response to Covid-19

All case studies were crowd-sourcing initiatives, reflecting the rise of crowd-sourced citizen science in health sciences and epidemiology. This rise has been spurred by mobile technologies that give access to big data, rapidly and with good accuracy (Katapally, 2020). In particular, mobile technology offers unprecedented opportunities for supporting data collection relying on images, movement patterns and basic physiological variables. Smartphones can also be used to ask people about their behaviours and feelings, and to link them with specific services or locations. The billions of smartphones in use worldwide, allow getting real-time insights directly from the people on the ground and to deploy and adapt projects relatively rapidly. It is thus no surprise that the Covid-19 pandemic spurred researchers to fast-track citizen science efforts that use mobile technologies to gather information about the disease, significantly informing the response to Covid-19 worldwide (Ravindran, 2021).

One of the biggest challenges in the current COVID-19 pandemic has been the difficulty to effectively track the virus, predict public health outcomes and recommend quick public health interventions. Through web or mobile apps, citizen-driven efforts can effectively fill those gaps, whether volunteers regularly log details about their location (CovidWatcher), their symptoms (Quantified Flu, Outbreaks Near Me), their well-being (Covid Open Survey), or the material resources available in their community (CovidWatcher, Safcast Covid-19 map). When the information is supplied by those directly affected, it makes a difference and provides an added value. Data from Outbreaks Near Me are complementing the official CDC Covid-19 tracking data to help identify early signs of spread and provided early evidence of the benefits of wearing facial masks. CovidWatcher and Safecast Covid-19 map provide credible sources of information that bring to light concerns of communities and give them a voice, empowering people and allowing them to free themselves from the dependence on government-based information. Quantified Flu also empowered people to establish how they might be able to sense the early symptoms of infection. Crowdsourced initiatives have also been a tremendous support in the identification of drugs or treatments against Covid-19, as highlighted by the tremendous participation in volunteer-computing initiatives to support molecular simulations in Covid Phym and OpenPandemics, which sped up the research process.

The case studies also underscored the sense of purpose citizen science can help deliver. The global pandemic that affects everyone has led to growing waves of anxiety, fueled by risk on the one hand, but also the overall uncertainty, restrictions of freedoms and overall poor governmental communications. The urgency and scale of the matter of concern made for high commitment. Often with time on their hands, people were eager to contribute to finding solutions, helping people in their community of family, understanding their own symptoms and risks. Collectively, the case studies underscore the unique potential citizen science offers to tap in the generosity of people to cooperate and contribute to a greater social good, and empower them, providing them with options to generate independent data.

4.8.2 Recommendations for addressing future emerging challenges through citizen science

Building on the good practices identified in the case studies, we highlight some key elements that were shared by all case studies and seem to enable responsiveness and impact delivery during a crisis.

- **Preparedness:** citizen science can be incredibly powerful, and scientists need to be aware of its potential and ready to tap into it. But this cannot get started in a day, a week, or a month. All case studies were able to launch their projects and secure engagement very rapidly (within less than six months of the onset of the pandemic). Their own work or their network acted as an early-warning signal, making them aware early on of the seriousness of the issue. They also all had previous experience in developing citizen science initiatives.
    - **Cooperate**: The Covid-19 situation created a unique and unprecedented commitment to a shared goal for researchers. Scientists in all fields suddenly wanted to work together and form teams with exceptional complementary expertise. But the collaborations that really worked were the ones that built on existing relationships, infrastructures, or ideas. In the case of Covid Watcher, some prior knowledge was needed. "We could not have gone to identify the communities from scratch, or to gain the expertise about all the survey instruments from scratch, and even less to develop the apps from scratch", says Noémie Elhadad of CovidWatcher. Responding to an emerging threat requires generating and using knowledge from across disciplines and sectors. When building your project, cooperate with projects and\or people that are already involved in the subject of research. It takes a lot of effort and funding to create and run such a project. It may be better to cooperate with people who do citizen science, than to try and do it all yourself.
    - **Modular infrastructure:** From a technical point of view, a nimble, adjustable infrastructure is needed, that can be grabbed and deployed at scale. OpenPandemics, Covid Phym, Outbreaks Near Me and Safecast Covid 19 map all built from an existing volunteer community and big data infrastructure. CovidWatcher, Quantified Flu and Covid Open Survey were also able to build on previous experiences in developing citizen science projects. Overall, open-source systems and flexible technologies can be rapidly and cheaply deployed, and are easier to scale up and tweak. When users contribute through an app or website, researchers can easily add or modify survey questions, to match the interest or the target audience. Together, these examples highlight the importance to invest in the citizen science infrastructure early-on, before the next emergency will arrive.
    - **Plan for success:** Think big. When you have a project with a high potential, it is worth investing in it from the start and plan for success. Redesigning the system a posteriori required significant resources. It is thus advisable to plan ahead to design a system that can be expandable and has the potential to scale. This may also require **thinking about the funding model**. The Flu Near You platform could not be extended to include Covid-19 symptoms

because its technology was too obsolete. This led to the creation of the Covid Near Me Platform, then fully redesigned and replaced by Outbreaks Near Me, when funding was available to develop a more expandable platform. This is also the experience from Safecast with its radiation and air quality monitoring maps. The network did not anticipate the amount of data it would get. As a result, the current database and map are already the third generation, and the system had to be fully redesigned in 2013 to allow it to accommodate incoming data and become global. To a different extent, OpenPandemics updated its platform and technology in the course of the project, which was made possible thanks to the financial support from the World Community Grid. **S**hort-term financing model creates uncertainty and can hinder project development of adaptability. CovidWatcher was unable to secure the financing to create the visualisations that would enable it to make the project more meaningful and useful for the volunteers and general users. Covid Open Survey, which was a side-project for all involved, with no dedicated funding ended up falling short of its objectives.

- **Simple, open, realistic design:** almost all case studies had a clear straightforward goal. In some cases, that goal developed and adapted over time (CovidWatcher). When the goal was too broad (as in Covid Open Survey), this led to difficulty in maintaining sustained engagement, both from the researchers and participants side.
    - **Keep it simple**. This will help make the project more accessible and enlist more users (e.g. Outbreaks Near Me), rather than daunting them off through complex or overly long surveys (e.g. Covid Open Survey).
    - **Be realistic and manage expectations**: **Agree on what your project goal is and frame your project accordingly**. In particular, ensure that the project team is well-aligned on key tradeoffs, such as sustainability vs. speed. Covid Near You (then Outbreaks Near Me) initially decided to develop something fast, and later considered sustainability issues, which worked well for the project given its history. **Evaluate what you expect from the project, and manage expectations.** Most case studies acknowledge that the project was a positive experience for all of the research team. Despite the long hours and sometimes haphazard path CovidWatcher took, team members felt rewarded to contribute to society, and it proved a very worthwhile training activity for students. In OpenPandemics, the positive and sometimes moving testimonials of the volunteer community spurred the team along the way. It is also important to be transparent to volunteers about what they will contribute to, and to deliver on these expectations.
    - **Design openness into the system from the beginning**. Few projects really have open, easily accessible data. Most of the case studies will make the

    data available upon request and appropriate justification of its use. Usually, only aggregated results are directly accessible, but the raw data is not downloadable. It is often the case that projects start, thinking that they will implement data openness afterwards. But it is almost impossible to implement data openness unless you do it transparently from the very beginning, acknowledges Azby Brown, because to make the data easily downloadable and shareable, you need to get the buy-in and agreement from the entire community.

- **Keep your contributors happy**: People are enthusiastic about contributing, but eager to know that their efforts are not in vain. Sustained engagement is key to the success of citizen science projects, and in particular for health sciences, where it allows to collect longitudinal data. All case studies, at the exception of Covid Open Survey and Quantified Flu, expanded significant efforts in outreach and in activities to engage the community. Expectations cannot always be met and citizens may feel deceived.
    - **Make sure the project is open to everyone:** People have very different motivations for participating, some people participate because they want to make something, not because they care about the bigger issue. Others are strongly politically motivated, while some are directly concerned about their local environment and need to have good data, others yet just want to be involved in citizen science. Leaving no one behind also means ensuring that your project is inclusive and open to the most vulnerable in society. CovidWatcher successfully repositioned itself to engage different underserved communities in New York City. Outbreaks Near Me is now considering additional ways to engage with the segments of society not represented in its survey.
    - **Understand and listen to people's concerns**. Today, with the availability of open-source tools, it is relatively easy to make sensors or crowd-source data. The harder part is building a community, and this can only be based on trust. It requires discussions, spending time with the community and listening to people's concerns. Much of Safecast's Covid-19 map activity is a response from listening to people about their concerns: people want to know what makes things safe, where to get tested, if they can get data from their own neighborhood on the map. Safecast tries to respond to these concerns, and when people see that their interest and activity is reflected in the identity of the group, it is a tremendous motivator. CovidWatcher listens to the communities and addresses the issues that concern them. In the bottom-up citizen science project Quantified Flu, the community is the one coming up with the research question and the project design, the researchers are simply facilitating and enabling the process.

- **Give back to your participants to secure trust and engagement:** Researchers should be aware that citizen science (even in its most low-level participation shape, such as in distributed computing), should be as bi-directional as possible. You need to be ready to feed the community with the right information and data to maintain enthusiasm. In turn, it is fundamental to be very clear and honest when providing feedback. Therefore, project goals and outcomes should be set explicitly at the outset of a project, to ensure informed participation, trust and motivation (Eleta et al., 2019). For instance, the participation in OpenPandemics increased each time the team made an announcement. But one should be cautious about over-promising; "Once you say that you are getting ready to do something, you create an expectation that you cannot disappoint" claims Stefano Forli. Regular feedbacks to the volunteer community, transparency about progress and hurdles, but also options to see and interact with the data (e.g. Outbreaks Near Me, Quantified Flu) are essential in building trust.
- **Partner with NGOs.** Building a community and sense of trust is essential, but it is very time consuming. This can be challenging for the research team, that needs to manage its resources optimally in order to advance with the research, but also provide regular feedbacks to the volunteer community. Partnerships with non-scientific organisations, whether community organisations (CovidWatcher), non-governmental organisations or distributed computing platforms (e.g. Ibercivis for Covid Phym, World Community Grid for OpenPandemics, Open Humans for Quantified Flu) help nudge the researchers and encourage social accountability. It can also help with outreach, enabling projects to reach wider, more diverse audiences.

4.8.3 Some challenges

One of the greatest challenges with the rise of crowd-sourced health data, is to protect citizen privacy and anonymity, in particular when tracking people's individual symptoms, movements or behaviours. Anonymity was a guiding principle across case studies, guaranteed through strong encryption processes. Volunteers were typically required to sign an informed consent before participation. In Quantified Flu, citizens were also provided the option to drop out and delete their data, as well as full control on how their data are shared.

The case studies highlight the strong relationship between matters of concern and engagement (Balestrini et al., 2021). However, few of the case studies considered the sustainability aspect from the outset. The challenge for all these projects is to sustain the engagement once the risk fades away or eventually disappears. Outbreaks Near Me is already noticing a slight drop in participation for instance. Balestrini and colleagues

(2021) argue that sustainability should be associated with the maintenance of the social bonds that can make the community more resilient in the face of future problems. This is what both Outbreaks Near Me and OpenPandemics are trying to achieve, by developing the tools and infrastructure for future pandemics. In turn, some projects may be destined to be short-lived, having fulfilled their mission during the pandemic (e.g. CovidWatchers).

The pandemic also directly challenges the relationship between science and people. It has made it increasingly evident that scientific advances and public health operate on much longer timescales than citizens expectations of responses. This discordance creates a high risk of loosing public trust in science. On the one hand, emerging disasters such as the Covid-19 pandemic affect the public understanding of science; people have been fed endless amounts of information on epidemiology. On the other hand, mis-information and the inherent uncertainties associated with scientific data, can lastingly affect the trust and public perception of scientific knowledge. Experience from the Fukushima meltdown and other natural disasters shows that public trust has not yet been recovered. It remains to be seen how that relationship will evolve post-Covid.

4.8.4 Conclusion

The report shows there is considerable opportunity to improve the outcomes of citizen science in response to emerging challenges. Such unforeseen events can have a transformative power of their own, as was witnessed with the Covid-19 pandemic, where entire populations were suddenly locked down, modifying their behaviour, including travel restrictions, social distancing, pausing work or studies, often accompanied with growing sentiment of anxiety, but also time on their hands. Participation in many citizen science programs skyrocketed in this period, demonstrating the unique potential citizen science offers to tap in the generosity of people to contribute to a greater good. It is now a great time for citizen science. The Covid-19 pandemic has laid bare some of the gaps in capacity of governments and traditional science. The prevalence and scope of mobile technologies is providing citizens options for generating independent data that can empower them. With the increasing power of citizens to affect change, citizen science is now increasingly recognised and promoted across the board, earning a place in national science policies, complementing the efforts of governmental organisations (Katapally, 2020). So, if you pick something that meaningful for you and the people around you, advises Azby Brown, lead researcher at Safecast, you will find the people around the world who are also interested in this and you can connect with.

References

Balestrini, M., Kotsev, A., Ponti, M. et al. Collaboration matters: capacity building, up-scaling, spreading, and sustainability in citizen-generated data projects. Humanit Soc Sci Commun 8, 169 (2021). https://doi.org/10.1057/s41599-021-00851-5

Bonney R, et al. 2009. Citizen science: a developing tool for expanding science knowledge and scientific literacy. Bioscience. 59(11), 977-84. https://doi.org/10.1525/bio.2009.59.11.9

Créquit P, Mansouri G, Benchoufi M, Vivot A, Ravaud P. 2018. Mapping of Crowdsourcing in Health: Systematic Review. J Med Internet Res. 20(5):e187. URL: http://www.jmir.org/2018/5/e187/; doi:10.2196/jmir.9330; PMID:29764795

Den Broeder L, Devilee J, Van Oers H, Schuit AJ, Wagemakers A. 2016. Citizen Science for public health. Health Promot Int. 33(3), 505-14.

Eleta I, Galdon Clavell G, Righi V, Balestrini M. 2019. The promise of participation and decision-making power in citizen science. Citiz Sci Theory Pr 4(1):8. doi: https://doi.org/10.5334/cstp.171

English PB, Richardson MJ, Garzon-Galvis C. 2018. From Crowdsourcing to Extreme Citizen Science: Participatory Research for Environmental Health. Annu Rev Public Health. 39, 335-50. PubMed https://doi.org/10.1146/annurev-publhealth-040617-013702

Haklay M. 2013. Citizen Science and Volunteered Geographic Information – overview and typology of participation. In: Sui, DZ, Elwood, S, and MF Goodchild (eds.). Crowdsourcing Geographic Knowledge. Berlin: Springer, 105-122.

Hicks A, Barclay J, Chilvers J et al. 2019. Global mapping of citizen science projects for disaster risk reduction. Front Earth Sci 7:226. https://doi.org/10.3389/feart.2019.00226

Katapally TR 2020. A Global Digital Citizen Science Policy to Tackle Pandemics Like COVID-19 J Med Internet Res, 22(5):e19357; doi: 10.2196/19357

Newman G, Wiggins A, Crall A, Graham E, Newman S, Crowston K. 2012. The future of citizen science: emerging technologies and shifting paradigms. Front Ecol Environ 10(6):298–304

Ravindran, S. 2021. Nature India, Smartphones help test and track a pandemic. doi:10.1038/nindia.2021.70 Published online 13 May 2021

Roy, H., Pocock, M. Preston, C. Roy, D., Savage, J., Tweddle, J. et al. 2012. Understanding Citizen Science & Environmental Monitoring. Final Report on behalf of UK-EOF.

Segal, E., Zhang, F., Lin, X., King, G., Shalem, O., Shilo, S., Allen, W.E., Alquaddoomi, F., Altae-Tran, H., Anders, S. and Balicer, R., 2020. Building an international consortium for tracking coronavirus health status. Nature medicine, 26(8), pp.1161-1165.

Shirk et al. 2012 Shirk, J. L., H. L. Ballard, C. C. Wilderman, et al. 2012. Public participation in scientific research: A framework for deliberate design. Ecology and Society 17(2). doi: 10.5751/ES-04705-170229.

Socientize Consortium, 2013. Green paper on citizen science. Citizen Science for Europe. Towards a better society of empowered citizens and enhanced research. Brussels.

Swan M. Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem, 2012. J Med Internet Res;14(2):e46 doi: 10.2196/jmir.

Turbé, A., Barba, J., Pelacho, M., Mugdal, S., Robinson, L.D., Serrano-Sanz, F., Sanz, F., Tsinaraki, C., Rubio, J.M. and Schade, S., 2019. Understanding the citizen science landscape for European environmental policy: An assessment and recommendations. Citizen Science: Theory and Practice, 4(1).

Wiggins A, Wilbanks J. 2019. The Rise of Citizen Science in Health and Biomedical Research. Am J Bioeth. 19(8), 3-14. PubMed https://doi.org/10.1080/15265161.2019.1619859

Appendix 1 – Case studies interview template and schedule

Interview template

Covid-19 citizen science projects - Questionnaire structure

Main question	Response
1. General information	
1.1 Project name	
1.2 Contact name	
1.3 Role in the project	
1.4 Date, time, interviewer	
1.5 Consent	
2. Design of the project	
2.1 Motivation for designing the project	
2.2 Did you have previous experience in designing participatory projects?	
2.3 How did you develop the partnership?	
2.4 What were the resources needed (staff, funding)	
2.5 How was your initiative received by colleagues	
3. Aims of the project	
3.1 How did you decide on the research questions?	
3.2 How did you design the survey (if relevant)	

3.3 How did you envisage the participation?	
3.4 Did you put in place some specific quality assurance procedures?	
4. Participant engagement	
4.1 How did you make your project known?	
4.2 Did you use specific means to motivate participation?	
4.3 Number of participants	
4.4 Participant profile	
4.5 Did you actively disseminate the project results?	
5. Outcomes	
5.1 General description	
5.2 How useful do you think the data produced was/is?	
5.3 Did you meet your goals?	
5.4 What do you think were the main successes?	
5.5 Is there anything you would do differently today?	
5.6 Do you intend to continue?	
5.7 Any recommendations/lessons-learnt for future projects?	
6. More details	
Any further comments or questions?	

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

Interview schedule

Project	Contacts	Interview date
COVID-PHYM	Javier Martínez de Salazar Javier Ramos Díaz Francesco Sanz	28/06/2021
OpenPandemics	Stafano Forli	28/06/2021
CovidWatcher	Noemie Elhadad	30/06/2021
Quantified Flu	Bastian Greshake Tzovaras	29/06/2021
Covid Open Survey	Bastian Greshake Tzovaras	29/06/2021
Safecast	Azby Brown	14/07/2021
Outbreaks Near Me	Autumn Getz	14/07/2021

Standard reporting format

- About the project (short project summary)
- Context and motivations
    - Research group and background
    - Project motivations
    - Landscape (partnerships, competition)
- Design choices and resources
    - Approach (why citizen science)
    - -Principles (/Survey design)
    - - Data aspects (accessibility, protection, validation)
    - - Resources
- Participation and engagement
    - Participants
    - Engagement and outreach
    - What did the participants gain
- Outcomes and lessons-learnt
    - Outcomes
    - Did you meet your goals
    - Successes and challenges
    - Sustainability

  - Lessons-learnt

Informed consent form

All interviewees were asked to sign the informed consent form below prior to the interviews.

Citizen science in response to emerging challenges
Letter to the research participants

Greetings,
Subject: Conducting research about citizen science in response to emerging challenges

Research Summary:
This study is part of the CS Track consortium for broadening knowledge about Citizen Science and the impact its activities can have. The study aims to investigate the responses of citizen to emerging challenges, and particularly to the COVID-19 pandemic. You have been invited to participate in this study because you have been involved in a citizen science project related to COVID-19 research. We value your commitment to this field and hope to learn from your experiences.
This research, including all section implemented in its framework, was approved by the "Ethics Committee" of the Inter-College Research Authority at the Mofet institute (approval 872522).
Participation in the research is voluntary. You have the right to withdraw your participation at any stage without being affected in any way. Your information will be stored securely, and your identity will be kept strictly confidential. Study findings and data may be published, but you will not be individually identifiable in these publications. We undertake to maintain your dignity and your rights throughout all stages of the research.
As part of your participation in the study, you will be asked to take part in an interview discussing your involvement in citizen science projects related to COVID-19 research. The duration of the meeting will be approximately 1 hour and will be conducted via zoom, at a time of your choice. Interview questions will include general information about your project its design and aims in addition to questions on participation patterns and project outcomes.
The research will be performed by Dr. Anne Turbe and Ms. Tslil Farchi and supervised by Dr. Yaela Golumbic from the Mofet Institute and on behalf of CS Track project. You can learn more about CS Track here- https://cstrack.eu and about the research program by contacting the research team by email - research-group@cstrack.eu

If you agree to participate in the research, please sign the consent form below and send it in reply to this email.
Sincerely,
Dr. Yaela Golumbic

yaela

Citizen science in response to emerging challenges
Consent form for participation in a research

To: Dr. Yaela Golumbic
Subject: Expressing consent for research participation

I agree to participate in a research on citizen science in response to emerging challenges which will be performed by Dr. Anne Turbe, Ms. Tslil Farchi at the Mofet Institute on behalf of CS Track project.
I have received information about the purpose of the research and the framework in which it will be carried out. I am aware that my participation in the research is voluntary, and that I can withdraw at any stage.
I give consent to the audio recording of the zoom interview.

____________	______________________	____________________
Date	Name	Signature