

STUDY PROTOCOL

Open Access



Merging citizen science with epidemiology: design of a prospective feasibility study of health events and air pollution in Cologne, Germany

Sara-Marie Soja¹, Robert Wegener², Natalie Kille² and Stefanie Castell^{1,3*} 

Abstract

Background Citizen science as an approach to merge society and science is not a new paradigm. Yet it is not common in public health, epidemiology, or medical sciences. SMARAGD (Sensors for Measuring Aerosols and ReActive Gases to Deduce health effects) assesses air pollution at participants' homes or workplaces in Cologne, Germany, as feasibility study with a citizen science approach. Personal exposure to air pollutants is difficult to study, because the distribution of pollutants is heterogeneous, especially in urban areas. Targeted data collection allows to establish connections between air pollutant concentration and the health of the study population. Air pollution is among the most urgent health risks worldwide. Yet links of individualized pollution levels and respiratory infections remain to be validated, which also applies for the feasibility of the citizen science approach for epidemiological studies.

Methods We co-designed a prospective feasibility study with two groups of volunteers from Cologne, Germany. These citizen scientists and researchers determined that low-cost air-quality sensors (hereafter low-cost sensors) were to be mounted at participants' homes/workplaces to acquire stationary data. The advantage of deploying low-cost sensors is the achievable physical proximity to the participants providing health data. Recruitment started in March 2021 and is currently ongoing (as of 09/22). Sensor units specifically developed for this study using commercially available electronic sensor components will measure particulate matter and trace gases such as ozone, nitrogen oxides, and carbon monoxide. Health data are collected using the eResearch system "Prospective Management and Monitoring-App" (PIA). Due to the ongoing SARS-CoV-2 pandemic, we also focus on COVID-19 as respiratory infection.

Discussion Citizen science offers many benefits for science in general but also for epidemiological studies. It provides scientific information to society, enables scientific thinking in critical discourses, can counter anti-scientific ideologies, and takes into account the interests of society. However, it poses many challenges, as it requires extensive resources from researchers and society and can raise concerns regarding data protection and methodological challenges such as selection bias.

Keywords Citizen science, Acute respiratory infections, COVID-19, Feasibility study, Air pollution, Symptom monitoring, m/eHealth, eResearch, Low-cost sensor, Sensor network

*Correspondence:

Stefanie Castell
stefanie.castell@helmholtz-hzi.de

¹ Department for Epidemiology, Helmholtz Centre for Infection Research, Inhoffenstr. 7, Brunswick, Lower Saxony 38124, Germany

² Forschungszentrum Jülich, Institute for Energy and Climate Research, IEK-8: Troposphere, Wilhelm-Johnen-Straße, Jülich, North Rhine-Westphalia 52428, Germany

³ German Centre for Infection Research (DZIF), Inhoffenstr. 7, Brunswick, Lower Saxony 38124, Germany



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Epidemiology is a term now familiar to most people due to the current coronavirus pandemic. Newspapers and online media report regularly on topics of epidemiology or the work of epidemiologists [1–3]. Citizen science seems to be less familiar [4, 5]. Citizen science is not a science per se but an approach to join citizens and researchers to do science together [6]. Such projects excel at facilitating inclusion of citizens in science, leading to relevant and meaningful results that benefit both science and society [6]. Citizen science has gained popularity in recent years also due to the availability of digital communication devices [7]. Citizen science, however, is not yet well-established in all research areas. So far, it is a rather well-known approach in life sciences or environmental sciences, whereas it is less common in medical research fields, like public health and epidemiology [8–10]. For example, as of September 2022, there are 540 hits for the search term “citizen science” in the SciFinder database with the context of public health, biomedical research, and medicine but 810 for environmental science, environmental monitoring, and biology [11]. On the Web of Science database, there are 234 hits for the search term “citizen science” and “public health”, “biomedical research”, and “medicine”, whereas 436 hits were retrieved for the combination of “citizen science” and “environmental sciences”, “environmental monitoring”, and “biology” [12]. This lack of integration into medical and related research prevents benefits coming into effect, such as empowerment and increased self-efficacy of participants, as well as meaningful prioritization of research areas in the eyes of citizens and society [10, 13]. This absence of use might stem from challenges with data security and ethics [10, 14–16].

The study SMARAGD

Our project “Sensors for Measuring Aerosols and Reactive Gases to Deduce health effects” (SMARAGD) is a pilot study. SMARAGD utilizes collection of a unique combination of targeted air pollution data with time-sensitive infection data provided weekly by participants. Air pollutants are often very heterogeneously distributed with traffic-related emissions varying on the scale of 100 m, but they are hardly measured apart from monitoring stations of the official networks that are often deployed at coarse spatial scales, which fail to capture all relevant heterogeneity [17, 18]. To address this, we measure air pollutants of interest using specific low-cost sensors (Fig. 1) that can be placed at citizens’ place of residence or workplace. The study is based on a cooperation between the Helmholtz Centre for Infection Research (HZI), the Forschungszentrum Jülich (FZJ), and the

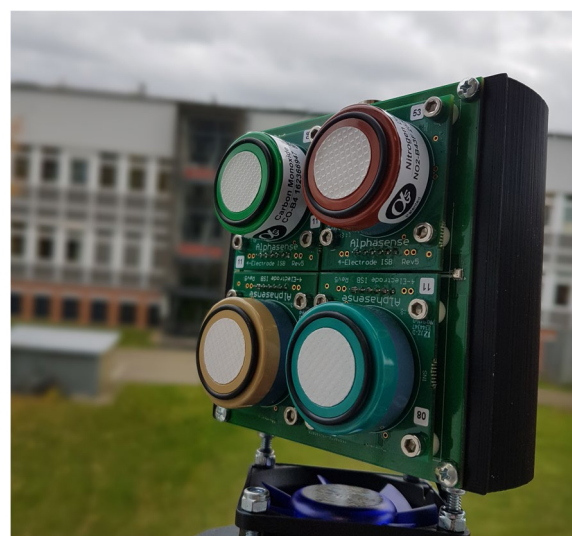


Fig. 1 Low-cost air quality sensors with the weather protection removed. The picture shows the front side with the four gas sensors

Open-Knowledge Lab Cologne (OK Lab) to explore associations between air pollution and respiratory infections in Cologne, Germany. The Helmholtz Centre Munich (HMGU) is also involved as an additional partner to provide background information for the project’s website. As funding of a pilot study is limited, the cohort size of this study has to be restricted. Therefore, there will be limitations in the precision of the estimate of the association between pollution and health. Due to the innovative and unique nature of the project, we also focus on the evaluation of the feasibility of such a study as citizen science project.

The World Health Organization (WHO) has declared air pollution to be one of the greatest risks for deteriorating health and premature death, especially for low- and middle-income countries, but also for high-income countries [19, 20]. Even though air pollution is decreasing overall in Germany, it still peaks in urban areas [21] and varies greatly depending on individual circumstances [18]. Air pollution affects many systems like the lungs or the cardiovascular system [22, 23]. The extent of the impact of air pollution on respiratory infections is a matter of scientific debate. Research often focusses on indoor air pollution [24–28] or specifically on lower-respiratory-tract infections [26, 29, 30]. Low-cost sensors to measure air pollution are a relatively new tool [31–34], and to our knowledge, no prior study has directly attempted to link their data to the people’s health. Previous studies linking health and air pollution [35, 36] have used data products from the EURAD (EUROpean Air pollution Dispersion) model [37] with a spatial resolution of $1 \times 1 \text{ km}^2$. However, it is known that air pollution can vary on spatial

scales smaller than this [17]. The value of individualized air pollution data, measured where participants spend a large proportion of their time, is not yet recognized in epidemiological research. Due to the ongoing pandemic, COVID-19 came into focus as health outcome as well. The eResearch system PIA (Prospective Monitoring and Management App) is used to measure symptoms and health-related factors [38, 39].

Aims and objectives

Our primary aim is to conduct a feasibility study that describes air pollution data gathered at participants' place of residence or workplace and data on participants' symptoms of acute respiratory infections, including COVID-19. We will explore associations between the amount of air pollution measured locally and respiratory infections. One secondary aim of SMARAGD is to evaluate the feasibility of the citizen science approach in epidemiological studies. For this, we collect data on compliance with app use and satisfaction and evaluate data quality. We presume that such parameters differ from standard epidemiological cohort studies. In a citizen science study, there is a closer relationship of study participants, researchers, and the research topics, and interests and concerns of citizens, especially those affected in areas with potentially high concentrations of air pollutants, are taken into account. Regarding feasibility of our study approach, the performance of the low-cost sensors in the project will be evaluated against standard air pollution measurements.

Thirdly, due to SMARAGD's nature as citizen science project, we also collect reports on exposures like stress, sleep, vaccination, life quality, and adherence to pandemic non-pharmaceutical interventions to be used as control variables in the main analysis but also for explorative analysis regarding effects on symptoms of acute respiratory infections.

Generally, our goal is to enhance knowledge transfer from citizens to researchers and from researchers to citizens making research more accessible and relevant to all stakeholders.

Methods/design

Study design

SMARAGD is a prospective cohort study with a citizen science approach based in Cologne, Germany, similar to a population-based study. One difference to a classical epidemiological observational cohort study is that the participants consist of a convenience sample, i.e., the citizen scientists, rather than a random sample from the population of interest. Moreover, we designed the project with the help of the members of the OK Lab Cologne and the Cologne School of Artificial Intelligence (AI). Both are

groups of people that work as volunteers in their spare time on open-source software and open data [40, 41]. Since there were already air quality monitoring activities in these groups, we established a working relationship. While the overall topic of the study was predetermined with the funding application, the exact scope of the project was intended to be collaboratively decided between citizens and researchers. Therefore, we conducted 14 workshops from November 2019 to November 2021 for and with citizen scientists in which they participated passively and actively, sometimes under guidance. Here, active participation refers to steps such as cognitive pre-testing of questionnaires, whereas passive participation refers to the exchange of information on the project and the discussion of further steps. According to Frederking et al. [9], this corresponds to a level of citizen participation of collaboration, co-production, and co-creation. We did not reach the level of cooperation described by Frederking et al., as the overall topic was determined without citizen scientists.

In addition, HMGU provided content for a study-specific website to deliver high-quality information on air pollutants and their impacts on human health [42, 43]. The topics hosted by the website were decided during the first workshops and are based on questions of the citizen scientists. We planned the field phase with both sensors and the eResearch app PIA for 1 year to cover all four seasons.

Low-cost air pollution sensors

For SMARAGD, we developed multi-sensor instruments for measurements of the air pollutants of interest; both trace gases and particulate matter (PM). We measure the trace gases such as carbon monoxide (CO), nitrogen oxides (NO, NO₂), and ozone (O₃) by electrochemical sensors that are commercially available from Alphasense (part number: CO-B4, NO-B4, NO₂-B43F, O₃-B431, respectively). PM_{2.5} and PM₁₀ (PM with sizes up to 2.5 and 10 µm) are quantified by laser scattering using the NOVA SDS011 sensor. Additionally, we collect data on temperature and relative humidity to correct the low-cost sensor data for cross-sensitivities. We selected these sensors to stay in the proposed budget while achieving a good reliability and robustness for ambient air measurements over the project's duration. Through a collaboration with the Cologne-based company Press Every Key (Cologne, Germany), 100 self-contained weatherproof nodes were produced for SMARAGD. The pure material costs amounted to 392 € per sensor node. We characterized the low-cost sensors against reference instrumentation in a controlled laboratory environment at FZJ. FZJ will distribute the instruments to the participants. The participants will install the sensors outside of their

respective residence or workplace. Once configured, the instruments will automatically measure and send data to a server hosted by FZJ.

PIA — prospective monitoring and management app

PIA is a free and open-source eResearch system developed since 2017 by the Department for Epidemiology at the HZI [38, 39, 44]. It includes a survey tool specifically designed for implementation in cohort studies. PIA displays questionnaires for participants and can be used flexibly via a mobile or a web app. It allows timely data collection and aims to increase response of compliance compared to paper-based studies. Additionally, PIA allows integration and management of samples of bio-specimens [39]. The system is already in use in several epidemiological studies [45].

Study population

Adults living in Cologne who are willing to use the eResearch system PIA on either a smartphone, tablet, or computer, who are able to speak and read German and are willing to operate an air pollution sensor or live in a household/work with someone who is willing to do so, are eligible to participate. We do not exclude vulnerable groups such as pregnant, breastfeeding women, or elderly people, as no intervention is part of this study. All participants complete an informed consent form before receiving any questionnaires. Initially, we recruit 100 participants who will receive a sensor. The number of participants is limited due to financial limitations for sensors and nasal swabs. These first 100 participants are referred to henceforth as the “sensor group”. As many requests were already made to the researchers for more spots to participate in the study even without sensors, an additional cohort is implemented for family, household members, or colleagues (in case the sensor will be placed at the workplace). Inclusion criteria remain the same.

Recruitment

Recruitment started in March 2021 with the members of the OK Lab Cologne as a pilot run to enable changes if required. Following the 2-week trial, the recruitment of citizens started via a snowball system.

Questionnaires

We discussed questionnaires with the members of the OK Lab Cologne and the Cologne School of AI during several workshops, and 18 questionnaires were pre-tested. As a result of these workshops, we modified 12 questionnaires. The remaining six questionnaires were pre-validated standardized questionnaires that could not be changed. Participants receive a weekly health questionnaire to report symptoms of respiratory infections,

including COVID-19. If someone reports an infection, they automatically receive a detailed symptom questionnaire. It is also possible to report an acute infection with a so-called spontaneous reporting form. We will use parts of the further questionnaires to adjust the statistical analyses for confounding variables and to assess risk factors commonly associated with respiratory infections, such as sociodemographic variables, preexisting medical conditions, stress, workload, exercise, quality of life, vaccinations, and travel. Due to the current pandemic, we also included specific questionnaires about the preventive measures against COVID-19, such as hygienic measures and contacts. For assessment of the usability of PIA, we provide a technology readiness questionnaire 30 days after the individual starting time [46] as well as an usability questionnaire 90 days after the start [47]. All questionnaires remain accessible in-app after they have been completed, in case the participants may want to use them for their documentation purposes.

Nasal swabs

Each participant in the sensor group is given the opportunity to collect a nasal swab at home. Participants can send the swab by using a provided pre-paid box to the Institute for Virology of the Hannover Medical School (MHH) at no cost. Each swab is tested by MHH for viruses such as influenza A and B, parainfluenza 1 to 4, rhinovirus, human metapneumovirus, adenovirus, respiratory syncytial virus (RSV), four endemic human coronaviruses (229E, NL63, HKU1, OC43), and SARS-CoV-2 via multiplex PCR. The results of those tests are provided in-app for the participants individually in their own personal accounts. Any infections with influenza or SARS-CoV-2 have to be reported to the health department of Cologne according to the §6 of the Law of Preventing Infections (IfSG) [48].

Statistical analysis

Due to the small sample size as a pilot and feasibility study, we will focus statistics on descriptive methods. Hence, we did not perform a power calculation. Where feasible, we will pool data from SMARAGD with data from similar studies we conducted to gain greater statistical relevance and comparison between different cohorts. Regarding feasibility, we define success as follows: (1) 75% of participants fill in weekly questionnaires on health status at least every other week on average, (2) at least 50% all other questionnaires are submitted; (3) missing items within submitted questionnaires are less than 10%, (4) 75% of nasal swabs are sent for laboratory analysis, and (5) 75% of sensors collect data during 75% of time of deployment.

Discussion

To our knowledge, we set up the first study that links air pollution with regularly, digitally reported health data that uses low-cost sensors which provide scientifically acceptable data quality.

SMARAGD and the benefits of citizen science for epidemiology

SMARAGD allows researchers and citizens to work collaboratively on a specific project. As a result, the researchers learn how scientific approaches can be made more accessible and integrated into participants' daily life. In our case, the citizens requested additional questionnaires regarding sleep patterns and workload. We also integrated a combination of weekly and monthly as opposed to daily questionnaires on health status, which were originally proposed by the epidemiologists. Hence, we added questionnaires on topics of interest to the participants or their preferred structure respectively that would otherwise have been ignored. After deciding which questionnaires and topics to use, we tested questionnaires via on-site and online cognitive pretests [49]. During these tests, many requests for changes to the questionnaires were expressed by the citizens, which we not only implemented within SMARAGD but will also take into consideration in future studies. However, for some pre-validated questionnaires [46, 47, 50–52], no changes were possible as otherwise the possibility of a comparison with other studies would no longer be possible and the validation would no longer apply. This led to critical discussions between researchers and citizens about the use of those questionnaires, and concerns were raised regarding comprehensibility of the questionnaires themselves. If citizens deem phrases unintelligible, the benefits of validated tools may need to be reconsidered. Working with citizens in the project also allows addition of the citizens' (sometimes highly specialized) expertise which should be taken into account. In SMARAGD, we used this specific expertise of our citizen partners for the development of the low-cost air pollution sensors. In addition, participants with a certain area of expertise can aid recruitment among acquaintances via a snowball system. The use of citizen science enables scientists and research institutes to focus on the issues that the public cares about. The desire to determine particulate matter levels led to the creation of initiatives, where citizens operate particulate matter sensors and feed the data into a database [53]. Other civil society movements on climate change such as "Fridays for Future" are fighting for the right to a healthy future and for politicians to acknowledge the need for a livable world in the decades to come [54]. Scientists see the same need, creating visibility for

urgent needs such as climate change [55]. These movements are examples of society wanting its views and needs recognized. In Germany, politicians have reacted by planning to introduce citizens' assemblies. They will meet on specific issues and develop recommendations for action to be considered by the federal government in Germany [56]. With the SARS-Cov-2 pandemic, people also started to recognize knowledge on health topics as urgent and valuable. Society learned about the meaning of concepts such as incidence, the reproduction number (R_0), different kinds of vaccination, and ways to protect oneself based on scientific results. Scientific television formats [57] or podcasts about the coronavirus pandemic with renowned epidemiologists and virologist [58] became popular. All these phenomena demonstrate that citizens want to engage in topics that concern them. Citizen science can also support and leverage such needs by providing access to high-quality scientific information or by fostering the ability to think scientifically when engaging in a critical discourse. It also creates a connection between nature, environment, and the citizen scientists' own health and well-being [10]. The project SMARAGD takes these demands for participation into account and offers inclusion in areas like engineering (sensors), digital matters and information technology (PIA), environment (air pollution), and health (respiratory infections) in one project. Both citizen groups who are part of the project existed before and worked with local pollution data. SMARAGD is just one project among others on which they have worked. Therefore, the lessons learned from the project will not only benefit the researchers for future projects but also may benefit the citizens for their future engagement in other projects.

Citizen science as an approach can also be used to counteract anti-scientific worldviews by engaging citizens in science and showing them how scientists come to their conclusions. This provides a possibility of changing opinions and values and of providing people with access to influence policies and politics [10, 13]. This is especially important in current times, where misinformation is prevalent. Experiencing and contributing to science may be a way to show that scientists usually aim to work for the benefit of the public. Whether this is something that SMARAGD can also convey will likely be assessable in project evaluations.

For epidemiology, it is described that those affected by something should be involved in its respective research [59]. German Medical Ethicist Alena Buyx remarked: "Every participant is a PI" [60]. This lays the foundation for citizen science in epidemiology. Nevertheless, only very few projects in the field of health and medicine have adopted the approach, e.g., in May 2022, only 30 projects

out of 230 on the EU-citizen science platform are tagged as health and medicine projects [61].

Limitations of citizen science in epidemiology

Citizen science also comes with challenges. Motivating citizens to participate in projects in their spare time, just to gain knowledge and understanding, may be difficult. Generally, citizens engage in science due to personal interests or because they want to contribute to something meaningful [62–65]. However, these reasons may change over time, which is in contrast to the timeline of many epidemiological projects that take years from conceptualization to publication and therefore require long-term commitment. Relevant participation in projects, i.e., from start to the end, may often take longer than anticipated at the beginning and necessitate perseverance from people who do this as a leisure time activity. With SMARAGD, we intend to keep motivation high through meetings between citizens and researchers, even if these have, to date, mostly been possible only online due to the coronavirus pandemic. Whether these meetings serve the purpose of maintaining sufficient continuous engagement will only become clear at the end of the project.

From a scientist's perspective, one of the benefits of citizen science is to gain different opinions from multiple backgrounds, e.g., regarding intelligibility of questionnaires. However, this can be challenging because time and effort must often be invested in communication between researchers and citizens [10, 13]. In addition, citizens may be interested in other topics than researchers. For example, for SMARAGD, one of the longer discussions was which questionnaires to include or exclude. As researchers from the field of infectious disease epidemiology, we focused our attention strongly on the then-emerging COVID-19 pandemic. The scientific team therefore wanted to include questionnaires on contacts, hygienic measures, and the like. However, the citizen group was more interested in allergies, which were perceived to have a greater affect in their daily life over the long term. Compromises pertaining to the scope of questionnaires had to be found to suit both sides. For SMARAGD, questionnaires about hygienic measures were implemented less often than intended because of the citizens' lack of specific interest while including basic questions on allergies, which were not originally conceived in the study scope as defined by the scientists.

Concerns regarding data security may make it difficult to recruit study participants in Germany [66]. This might also be applied to citizen science approaches. As a countermeasure, we specifically offered discussions regarding data security resulting in a traditional epidemiological approach with questionnaire and medical data being pseudonymized and data primarily accessible only

to researchers who are bound by data privacy regulations and practices.

Another factor that could lead to a general problem in citizen science is the fact that citizens invest their spare time in the project, whereas researchers are usually paid, even though both parties may spend an equal amount of resources [15, 16]. So far, this has not been an issue between citizens and researchers in SMARAGD. However, in the first workshop, we jointly decided that the major part of communication and study organization should remain with the researchers, rather than be implemented as shared tasks. Therefore, the majority of the overall work rests with the researchers. This asymmetry is reflected in the fact that our citizen partners are not among the authors of this paper.

Probably the most well-known challenge of citizen science is that specific groups of people are more likely to engage in specific projects. Therefore, it can be argued that projects using a citizen science approach do not represent the general population and therefore might be biased [10, 67]. In the case of SMARAGD, we expect a disproportional number of young- to middle-aged men, with relatively good health and higher education, as this was the composition of the preexisting group (OK Lab Cologne) supporting the project. However, this does not always have to be considered a challenge or even a problem; especially for causal research aims, the composition of the population does not necessarily have an impact [68] which applies to the exploration of the relation between air pollutants and health indicators in SMARAGD. Whereas, e.g., for surveys that aim at describing characteristics of a certain population, representativeness is essential [68]. In SMARAGD, we assess feasibility of a specific study type in a convenience sample, i.e., citizen scientists, without deducing a general rule for all of epidemiology. Another objective of SMARAGD is to compare measurements of low-cost sensors to those from high-quality reference instrumentation; therefore, the selection of participants needs to be described but does not necessarily affect parts of the research done in SMARAGD. However, an awareness of any bias is undoubtedly always required, but this applies to all population-based epidemiological research, not only citizen science. On the contrary, citizen science may even offer the chance to include groups into science that are otherwise possibly left out [10, 67].

One challenge that is still fraught with great uncertainty in citizen science projects is data protection [69]. If citizens engage deeply in scientific projects, they might encounter the collected raw data. In Germany, there are regulations in place considering data protection like the General Data Protection Regulations (GDPR) that provide the respective framework [70]. In particular, medical

data, as stated in §9 of GDPR are considered highly sensitive and are protected accordingly [70]. The GDPR or guidelines like the “Good Epidemiological Practice” [59, 70] do not consider the involvement of nonscientific personnel who have not signed any agreement with data protection regulations. The “Leitfaden für rechtliche Fragestellungen in Citizen-Science-Projekten [Guide to Legal Issues in Citizen Science Projects]” provides some information on the GDPR in the context of citizen science [69] but is considered as an overview rather than a specific and concrete guideline. For example, if participants get to know each other in workshops, they might recognize each other’s provided data sets, even if answer categories are grouped providing less detailed information on individuals while at the same time losing information that might be of scientific interest. In a small project like SMARAGD, recognition can easily occur, especially because many of the participants have known each other for a long time.

One specific data protection issue is the availability of the sensor location data to all participants during the data collection period as the place of residence or workplace is

one of the main aspects in the project. Therefore, we had to find a suitable way to collect this information and provide them to participants without violating the GDPR. Therefore, we superimposed a hexagon pattern over the city of Cologne, which is structured in a way that usually more than 30 people live in one hexagon (Fig. 2). If fewer than the expected 30 individuals live in one hexagon, data is represented in an adjacent hexagon. This allows for the study of spatial patterns but ensures that housing locations of participants remain confidential. Each sensor is sorted into the respective area by researchers who have access to address data but not to medical data and who have signed a data privacy notice. Participants will be able to see information regarding air pollutants of the sensors online but will not be able to identify the address of a specific sensor, only its approximate location within the city.

The fact that participants see the data set might also conflict with the individual’s right not to know about conditions or results of any medical measure. Participants can indicate in the consent form that they do not want to be informed about the result of the nasal swab laboratory

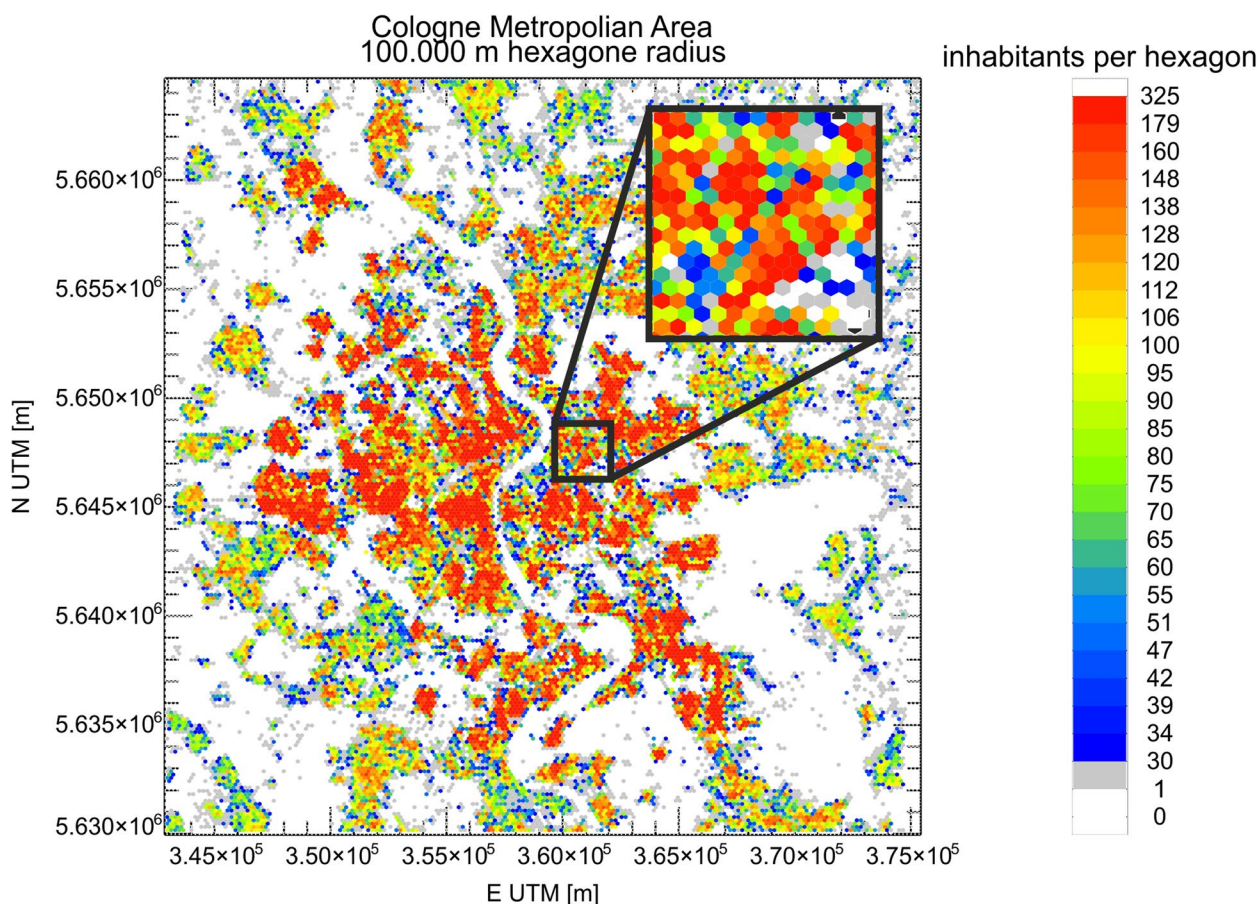


Fig. 2 Cologne Metropolitan Area with 100-m hexagon radius showing the number of inhabitants per hexagon. Data from <https://dataforgood.fb.com/docs/high-resolution-population-density-maps-demographic-estimates-documentation/#datasets> downloaded on 2021/04/18

analysis. However, they might identify themselves and their result in the data set. Therefore, we decided that any combination of identifiable answers that occurs less than five times will be recategorized in a less detailed manner so that no identification can take place. In general, citizen science approaches are easier to apply for noncritical topics. In SMARAGD, we chose not to include, e.g., questions on sexual contacts, although these are possibly associated with the risk of infections. The new “Citizen Science Weißbuch (Citizen Science White Paper)” in Germany also recognizes the great uncertainty regarding data protection in citizen science; however, it too does not offer a ready-to-use solution regarding data privacy [71]. Still, for the provision of the data set to the citizen, all these data protection aspects must be considered.

Conclusion

Citizen science can make research more meaningful by including thoughts and opinions of those affected. Bringing researchers and citizens together allows knowledge transfer from both ends so that everyone benefits from science. It also allows civil engagement which many people demand. Therefore, citizen science can be a way to take science onto a new level. Even though citizen science brings many benefits, it also bears its challenges. Long-term motivation of citizens to participate and communication barriers between researchers and citizens are among the more general limitations. Other challenges may arise, particularly in epidemiological studies, such as data protection issues due to working with medical data or selection bias. Real-world projects such as SMARAGD are needed to assess the specific benefits and develop how to overcome the challenges.

Abbreviations

AI	Artificial intelligence
EURAD	European Air Pollution Dispersion
FZJ	Forschungszentrum Jülich
GDPR	General Data Protection Regulation
HMGU	Helmholtz Centre Munich
HZI	Helmholtz Centre for Infection Research
MHH	Hannover Medical School
OK Lab	Open Knowledge Lab
PI	Principal investigator
PIA	Prospective Management and Monitoring App
SMARAGD	Sensors for Measuring Aerosols and ReActive Gases to Deduce health effects

Acknowledgements

We greatly acknowledge the work of all members of the OK Lab Cologne and the Cologne School of AI for their contributions, support, and help with the design of the project and especially Marcel Belledin for the organization. We also thank Tim Becker and Matthias Krauß from Press Every Key for their great contribution to the development of our sensors. We thank Elmar Frieze (FZJ) for providing EURAD data and Michael Decker and Helga London for their support on the IT infrastructure, data visualization, and the project website. Many thanks also for the support of Dr. Carolina Klett-Tammen, Christiane Steguweit, Daniel Schulze, Irina Janzen, Jana-Kristin Heise, Frederic Leone, and

Tina Barohn of the Helmholtz Center for Infection Research for supporting the study and the manuscript. We also thank Gertrud Hammel for their contribution to the content of the website. We would also like to thank all participants for taking part voluntarily in the study.

Authors' contributions

SMS wrote the first draft of the paper with contributions from all co-authors. SC conceptualized PIA as eResearch system and together with RW designed the study. RW and NK supported the development of the instruments and performed their characterization. RW, NK, and SC edited the manuscript. The authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. The SMARAGD project is one of the Helmholtz Association's projects funded by the Impulse and Networking Fund with a focus on citizen science (reference number: CS-0010).

Availability of data and materials

Not applicable.

Declarations

Ethics approval and consent to participate

The study was approved by the ethics committee of Hanover Medical School (no. 9587_BO_S_2021), and the data protection concept was advised by the data protection officers of the involved institutions and submitted to the Federal Commissioner for Data Protection and Freedom of Information (“Bundesbeauftragter für den Datenschutz und die Informationsfreiheit,” BfDI). Participants have given written informed consent before participating in the project.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 26 September 2022 Accepted: 17 January 2023

Published online: 22 February 2023

References

1. ZEIT Online. 281 Suchergebnisse für >>Epidemiologie<<. 2022. Available from: <https://www.zeit.de/suche/index?q=Epidemiologie>. Cited 2022 May 25.
2. Bild. Bild.de Suche: Ergebnisse 1 - 10 von insgesamt 116 Treffer(n) Ihrer Suche nach "Epidemiologie"; 2022. Available from: <https://www.bild.de/suche.bild.html?query=Epidemiologie>. Cited 2022 May 25.
3. Tagesschau. Suche in tagesschau.de: Suchbegriff: Epidemiologie. Ihre Suche nach "Epidemiologie" ergab 76 Treffer; 2022. Available from: https://www.tagesschau.de/suche2.html?query=Epidemiologie&sort_by=date. Cited 2022 May 25.
4. Tagesschau. Suche in tagesschau.de: Suchergebnisse für 'Citizen Science': 1 Treffer; 2022. Available from: www.tagesschau.de/suche2.html?query=Citizen+Science&sort_by=date. Cited 27.10.2022.
5. Tagesschau. Suche in tagesschau.de: Suchergebnisse für 'Bürgerwissenschaften': 0 Treffer; 2022. Available from: www.tagesschau.de/suche2.html?query=b%C3%BCrgerwissenschaften&sort_by=date. Cited 27.10.2022.
6. ECSA (European Citizen Science Association). Ten Principles of Citizen Science. Berlin; 2015. Available from: https://ecsa.citizen-science.net/wp-content/uploads/2020/02/ecsa_ten_principles_of_cs_german.pdf. Cited 2022 May 25.
7. Rüfenacht S, Woods T, Agnello G, Gold M, Hummer P, Land-Zandstra A, Sieber A, et al. Communication and dissemination in citizen science. In: Vohland K, Land-Zandstra A, Ceccaroni L, Lemmens R, Perelló J, Ponti M, et al., editors. The science of citizen science. Cham: Springer

- International Publishing; 2021. p. 482. Available from: https://pure.iiasa.ac.at/id/eprint/16993/1/2021_Book_TheScienceOfCitizenScience.pdf. Cited 27 Oct 2022.
8. Audubon. Audubon; 2022. Available from: www.audubon.org/about. Cited 2023 Feb 20.
 9. Frederking A, Richter S, Schumann K. Citizen Science auf dem Weg in den Wissenschaftsaltag. Berlin: VDI; 2016. (iit perspektive; vol 26). Available from: www.iit-berlin.de/publikation/citizen-science-auf-dem-weg-in-den-wissenschaftsaltag/. Cited 2022 May 25.
 10. den Broeder L, Devilee J, van Oers H, Schuit AJ, Wagemakers A. Citizen science for public health. *Health Promot Int*. 2018;33(3):505–14.
 11. CAS SciFinder. References search for “citizen science.” 2022. Available from: <https://scifinder-n.cas.org/search/reference/628e35a18ea60d2323d05708/1>. Cited 2022 May 25.
 12. Web of Science, References search for “citizen science”; 2022. Available from: <https://www.webofscience.com/wos/woscc/summary/e95fa35b-aca5-4c2d-ad8b-77ac1b8e1dd8-3bc21882/relevance/1>. Cited 2022 May 31.
 13. King AC, Winter SJ, Sheats JL, Rosas LG, Buman MP, Salvo D, et al. Leveraging citizen science and information technology for population physical activity promotion. *Transl J Am Coll Sports Med*. 2016;1(4):30–44.
 14. Novak J, Becker M, Grey F, Mondardini R. Citizen engagement and collective intelligence for participatory digital social innovation. In: Bonn A, Vogel J, Makuch Z, Bowser A, Haklay M, Hecker S, editors. *Citizen science: innovation in open science, society and policy*. London (UK): UCL Press; 2018. p. 124–45. Available from: <https://muse.jhu.edu/book/81913/>.
 15. Riesch H, Potter C. Citizen science as seen by scientists: methodological, epistemological and ethical dimensions. *Public Underst Sci*. 2014;23(1):107–20.
 16. Tauginiene L, Hummer P, Albert A, Cigarini A, Vohland K, et al. Ethical challenges and dynamic informed consent. In: Vohland K, Land-Zandstra A, Ceccaroni L, Lemmens R, Perello J, Ponti M, et al., editors. *The Science of Citizen Science*. Cham (CH): Springer Cham; 2021. p. 397–416. Available from: <https://link.springer.com/book/10.1007/978-3-030-58278-4#book-header>. Cited 27 Oct 2022.
 17. Bonn B, von Schneidmesser E, Andrich D, Quedenau J, Gerwig H, Lüdecke A, et al. BAERLIN2014 – the influence of land surface types on and the horizontal heterogeneity of air pollutant levels in Berlin. *Atmos Chem Phys*. 2016;16(12):7785–811.
 18. Wegener R, Klemp D, Dubus R, Rohrer F, Wahner A. Mobile measurements reveal high NO₂ and ammonia concentrations in German inner cities. In: 4th ICACGP Quadrennial Symposium and 15th IGAC Conference. 2018.
 19. World Health Organization. Air quality database: update 2016. 2016. Available from: www.who.int/data/gho/data/themes/air-pollution/who-air-quality-database/2016#cms. Cited 2022 May 25.
 20. WHO Regional Office for Europe. Evolution of WHO air quality guidelines: past, present and future. Copenhagen (DK): WHO Regional Office for Europe; 2017. Available from: https://www.euro.who.int/__data/assets/pdf_file/0019/331660/Evolution-air-quality.pdf. Cited 27 Oct 2022.
 21. Schultz MG, Klemp D, Wahner A. Luftqualität. In: Brasseur G, Jacob D, Schuck-Zöller S, editors. *Klimawandel in Deutschland*. Berlin, Heidelberg (DEU): Springer Spektrum; 2017. p. 127–36. Available from: https://link.springer.com/chapter/10.1007/978-3-662-50397-3_13#. Cited 27 Oct 2022.
 22. Gurjar BR, Jain A, Sharma A, Agarwal A, Gupta P, Nagpure AS, et al. Human health risks in megacities due to air pollution. *Atmos Environ*. 2010;44(36):4606–13.
 23. Peters A, Hoffmann B, Brunekreef B, Knäzi N, Kutlar Joss M, Probst-Hensch N, et al. Die Rolle der Luftschadstoffe für die Gesundheit. *Schweiz Ärzteztg*. 2019;100(23–24):796–9.
 24. Kafando B, Windinpsidi Savadogo P, Millogo T, Sana A, Kouanda S, Sondo B. Pollution de l'air intérieur et prévalence des infections respiratoires aiguës chez les enfants à Ouagadougou. *Sante Publique*. 2018;30(4):575–86.
 25. Misra P, Srivastava R, Krishnan A, Sreenivaas V, Pandav CS. Indoor air pollution-related acute lower respiratory infections and low birthweight: a systematic review. *J Trop Pediatr*. 2012;58(6):457–66.
 26. Perez-Padilla R, Schilmann A, Riojas-Rodriguez H. Respiratory health effects of indoor air pollution. *Int J Tuberc Lung Dis*. 2010;14(9):1079–86.
 27. Raju S, Siddharthan T, McCormack MC. Indoor air pollution and respiratory health. *Clin Chest Med*. 2020;41(4):825–43.
 28. Bardana EJ. Indoor pollution and its impact on respiratory health. *Ann Allergy Asthma Immunol*. 2001;87(6):33–40.
 29. Mehta S, Shin H, Burnett R, North T, Cohen AJ. Ambient particulate air pollution and acute lower respiratory infections: a systematic review and implications for estimating the global burden of disease. *Air Qual Atmos Health*. 2013;6(1):69–83.
 30. Jary H, Simpson H, Havens D, Manda G, Pope D, Bruce N, et al. Household air pollution and acute lower respiratory infections in adults: a systematic review. *PLoS ONE*. 2016;11(12):e0167656.
 31. Mead MI, Popoola O, Stewart GB, Landshoff P, Calleja M, Hayes M, et al. The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmos Environ*. 2013;70:186–203.
 32. Considine EM, Reid CE, Ogletree MR, Dye T. Improving accuracy of air pollution exposure measurements: statistical correction of a municipal low-cost airborne particulate matter sensor network. *Environ Pollut*. 2021;268(Pt B):115833.
 33. Cross ES, Williams LR, Lewis DK, Magoon GR, Onasch TB, Kaminsky ML, et al. Use of electrochemical sensors for measurement of air pollution: correcting interference response and validating measurements. *Atmos Meas Tech*. 2017;10(9):3575–88.
 34. Kim J, Shusterman AA, Lieschke KJ, Newman C, Cohen RC. The Berkeley atmospheric CO₂ observation network: field calibration and evaluation of low-cost air quality sensors. *Atmos Meas Tech*. 2018;11(4):1937–46.
 35. Lucht S, Hennig F, Moebus S, Führer-Sakel D, Herder C, Jöckel K-H, et al. Air pollution and diabetes-related biomarkers in non-diabetic adults: a pathway to impaired glucose metabolism? *Environ Int*. 2019;124:370–92.
 36. Rodins V, Lucht S, Ohlwein S, Hennig F, Sopka V, Erbel R, et al. Long-term exposure to ambient source-specific particulate matter and its components and incidence of cardiovascular events - the Heinz Nixdorf Recall study. *Environ Int*. 2020;142:105854.
 37. Elbern H, Strunk A, Schmidt H, Talagrand O. Emission rate and chemical state estimation by 4-dimensional variational inversion. *Atmos Chem Phys*. 2007;7(14):3749–69.
 38. hzi-braunschweig/pia-system. 10.5281/zenodo.6389565. Available from: <https://zenodo.org/record/7189235#Y1pRQ7ZCRPY>. Cited 2022 Oct 27.
 39. HZI (Helmholtz Centre for Infection Research). PIA Die App für Gesundheitsforschung; 2022. Available from: www.info-pia.de. Cited 2022 May 25.
 40. Code for Germany. OK Lab Köln; 2022. Available from: <https://codefor.de/koeln/>. Cited 2022 May 25.
 41. GitHub. School of AI. 2022. Available from: <https://github.com/orgs/Cologne-School-of-AI/projects>. Cited 2022 May 25.
 42. Lecker V. SMARAGD: Luftschadstoffe. 2022. Available from: <https://www.fz-juelich.de/de/iek/iek-8/projekte/smaragd/luftschadstoffe>. Cited 2022 Jun 1.
 43. Lecker V. SMARAGD: Gesundheitsauswirkungen. 2022. Available from: <https://www.fz-juelich.de/de/iek/iek-8/projekte/smaragd/gesundheitsauswirkungen>. Cited 2022 Jun 1.
 44. Heise J-K, Dey R, Emmerich M, Kemmling Y, Sistig S, Krause G, et al. Putting digital epidemiology into practice: PIA- prospective monitoring and management application. *Inform Med Unlocked*. 2022;30:100931.
 45. HZI (Helmholtz Centre for Infection Research). PIA Projekte; 2022. Available from: <https://info-pia.de/projekte/>. Cited 2022 May 25.
 46. Neyer FJ, Felber J, Gebhardt C. Entwicklung und Validierung einer Kurzskaala zur Erfassung von Technikbereitschaft. *Diagnostica*. 2012;58(2):87–99.
 47. Brooke J. SUS: a “quick and dirty” usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. *Usability evaluation in industry*. CRC Press; 1996. p. 207–12.
 48. Gesetz zur Verhütung und Bekämpfung von Infektionskrankheiten beim Menschen: Infektionsschutzgesetz - IfSG; 2000. Available from: www.gesetze-im-internet.de/ifsg/_6.html. Cited 2022 May 25.
 49. Porst R. Fragebogen: Ein Arbeitsbuch. 4th ed. Wiesbaden: Springer VS; 2014.
 50. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res*. 1989;28(2):193–213.

51. Fuchs R, Klaperski S, Gerber M, Seelig H. Messung der Bewegungs- und Sportaktivität mit dem BSA-Fragebogen. *Zeitschrift für Gesundheitspsychologie*. 2015;23(2):60–76.
52. Lincke HJ, Vomstein M, Lindner A, Nolle I, Häberle N, Haug A, Nübling M. COPSOQ III in Germany: validation of a standard instrument to measure psychosocial factors at work. *J Occup Med Toxicol*. 2021;16(1):50. <https://doi.org/10.1186/s12995-021-00331-1>.
53. Sensor.community; 2022. Available from: <https://sensor.community/de/>. Cited 2022 Nov 1.
54. FridaysforFuture. FridaysforFuture. 2022. Available from: <https://fridaysforfuture.org/>. Cited 2022 Mar 16.
55. Capstick S, Thierry A, Cox E, Berglund O, Westlake S, Steinberger JK. Civil disobedience by scientists helps press for urgent climate action. *Nat Clim Chang*. 2022;12(9):773–4.
56. Mehr Demokratie e.V. Vom Bundestag beauftragte Bürgerräte kommen noch in diesem Jahr; 2022. Available from: www.mehr-demokratie.de/news/voll/vom-bundestag-beauftragte-buergerraete-kommen-noch-in-diesem-jahr-1. Cited 2022 Jun 1.
57. ZDF. Mithink X Die Show; 2022. Available from: www.zdf.de/show/mithink-x-die-show. Cited 2022 May 26.
58. Norddeutscher Rundfunk. Das Coronavirus-update von NDR Info. 2022. Available from: www.ndr.de/nachrichten/info/podcast4684.html. Cited 2022 May 26.
59. Hoffmann W, Latza U, Baumeister SE, Brünger M, Buttman-Schweiger N, Hardt J, et al. Guidelines and recommendations for ensuring good epidemiological practice (GEP): a guideline developed by the German Society for Epidemiology. *Eur J Epidemiol*. 2019;34(3):301–17.
60. Buyx A, Del Savio L, Prainsack B, Völzke H. Every participant is a PI. Citizen science and participatory governance in population studies. *Int J Epidemiol*. 2017;46(2):377–84.
61. EU-citizen.science. Project search. 2022. Available from: <https://eu-citizen.science/projects?keywords=&orderby=&country=&status=&hasTag=&difficultyLevel=&topic=Health+%26+Medicine&participationTask>. Cited 2022 May 26.
62. Dickinson JL, Shirk J, Bonter D, Bonney R, Crain RL, Martin J, et al. The current state of citizen science as a tool for ecological research and public engagement. *Front Ecol Environ*. 2012;10(6):291–7.
63. Greenwood JJD. Citizens, science and bird conservation. *J Ornithol*. 2007;148(S1):77–124.
64. Pocock MJ, Chapman DS, Sheppard LJ, Roy He. Choosing and using citizen science: a guide to when and how to use citizen science to monitor biodiversity and the environment. Wallingford (UK): Centre for Ecology & Hydrology; 2014. Available from: https://www.ceh.ac.uk/sites/default/files/sepa_choosingandusingcitizenscience_interactive_4web_final_amended-blue1.pdf. Cited 27 Oct 2022.
65. Societize Consortium. Green paper on citizen science. In: Citizen science for Europe: towards a better society of empowered citizens and enhanced research. Brüssel: European Commission Belgium; 2013. p. 5–51.
66. Beullens K, Loosveldt G, Vandenplas C, Stoop I. Response rates in the European social survey: increasing, decreasing, or a matter of fieldwork efforts? In: Survey methods: insights from the field. 2018. <https://doi.org/10.13094/SMIF-2018-00003>.
67. Heinisch B, Oswald K, Weißpflug M, Shuttleworth S, Belknap G, et al. Citizen humanities. In: Vohland K, Land-Zandstra A, Ceccaroni L, Lemmens R, Perelló J, Ponti M, et al., editors. The science of citizen science. Cham: Springer International Publishing; 2021. p. 97–118.
68. Rothman KJ, Gallacher JEJ, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol*. 2013;42(4):1012–4.
69. Freyberg L, Euler E, Schneider UK, Gromek G. Leitfaden für rechtliche Fragestellungen in Citizen-Science-Projekten. Berlin (DE): Museum für Naturkunde Berlin; 2020. Available from: www.buergerschaffewissen.de/sites/default/files/grid/2021/01/19/Citizen_Science_rechtlicher_Leitfaden_19.01.2021_WEB.pdf. Cited 27 Oct 2022.
70. Datenschutz-Grundverordnung: DSGVO; 2016. Available from: <https://dsgvo-gesetz.de/>. Cited 2022 Jun 1.
71. Bonn A, Brink W, Hecker S, Herrmann TM, Liedtke C, Premke-Kraus M, et al. Weißbuch citizen science Strategie 2030 für Deutschland. 2021. Available from: <https://www.econstor.eu/bitstream/10419/247794/1/1777622972.pdf>. Cited 01 Nov 2022.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

