RESEARCH SERIES vol 05

COMMONS LAB

Wilson Center

STAKEHOLDER ANALYSIS

INTERNATIONAL CITIZEN SCIENCE STAKEHOLDER ANALYSIS ON DATA INTEROPERABILITY

Claudia Göbel, Museum für Naturkunde Berlin, Germany

Victoria Y. Martin, Cornell Lab of Ornithology, Ithaca, NY, USA[1]

Mónica Ramírez-Andreotta, Department of Soil, Water, and Environmental Science and Division of Community, Environment & Policy, University of Arizona, Tucson, AZ, USA

STAKEHOLDER ANALYSIS

INTERNATIONAL CITIZEN SCIENCE STAKEHOLDER ANALYSIS ON DATA INTEROPERABILITY

Claudia Göbel, Museum für Naturkunde Berlin, Germany

Victoria Y. Martin, Cornell Lab of Ornithology, Ithaca, NY, USA[1]

Mónica Ramírez-Andreotta, Department of Soil, Water, and Environmental Science and Division of Community, Environment & Policy, University of Arizona, Tucson, AZ, USA

Authors are listed in alphabetical order.
Each author contributed equally to this report.

STAKEHOLDER ANALYSIS:
International Citizen Science Stakeholder Analysis on Data Interoperability.

Commons Lab
Science and Technology Innovation Program
Woodrow Wilson International Center for Scholars
One Woodrow Wilson Plaza
1300 Pennsylvania Avenue, N.W.
Washington, DC 20004-3027

https://www.wilsoncenter.org/program/science-and-technology-innovation-program

Authors: Claudia Göbel, Victoria Y. Martin, Mónica Ramírez-Andreotta
Cover Design: Jeremy Swanston
Layout Design: Kathy Butterfield

Available for download free of charge at

https://www.wilsoncenter.org/research/Commons%20Lab

This report is funded by the Alfred P. Sloan Foundation and published through the Woodrow Wilson International Center for Scholars, in Washington, DC.

The Woodrow Wilson International Center for Scholars is the national, living U.S. memorial honoring President Woodrow Wilson. In providing an essential link between the worlds of ideas and public policy, the Center addresses current and emerging challenges confronting the United States and the world. The Center promotes policy-relevant research and dialogue to increase the understanding and enhance the capabilities and knowledge of leaders, citizens, and institutions worldwide. Created by an act of Congress in 1968, the Center is a nonpartisan institution headquartered in Washington, D.C.; it is supported by both public and private funds.

For more information about the Center's activities and publications, please visit us on the Web at www.wilsoncenter.org.

The Science and Technology Innovation Program (STIP) analyzes the evolving implications of such emerging technologies as synthetic biology, nanotechnology, and geoengineering. STIP's research goes beyond laboratory science to explore new information and communication technologies, sensor networks, prediction markets, and serious games. The program provides critical yet nonpartisan research for the policymaking community and guides officials in the design of new governance frameworks. It gauges crucial public support for science and weighs the overall risks and benefits of technology for society at large.

COMMONS LAB

The Commons Lab of STIP seeks to mobilize public participation and innovation in science, technology and policy.

Commons Lab Staff

Anne Bowser, CoDirector, Commons Lab
Elizabeth Tyson, CoDirector, Commons Lab

Blog: https://www.wilsoncenter.org/blogs/ctrl-forward
Facebook: https://www.facebook.com/WilsonSTIP
Twitter: https://twitter.com/WilsonSTIP

The Commons Lab of the Science and Technology Innovation Program is supported by the Alfred P. Sloan Foundation.

ACKNOWLEDGEMENTS

The Authors are indebted to our interviewees for contributing their precious time, experience, expertise, and ideas to this study, to the reviewers for their critical comments and thoughtful suggestions, and to the Wilson Center staff, Anne Bowser and Elizabeth Tyson for their visionary leadership in establishing this study, and their careful guidance throughout the process. We would especially like to thank: Theresa Crimmins (USA National Phenology Network), Darlene Cavalier (Arizona State University and SciStarter), Ben Gerhardstein (U.S. Agency for Toxic Substances and Disease Registry), Grand Lake Watershed Mercury Study, Steve Kelling (eBird), Connie Walker and Mark Newhouse (Globe at Night), Hagit Keysar (Public Lab Organizer), Dacha Atienza Ariznavarreta (BioBlitzes at NHM Barcelona), Andrea Sieber (Landscape and You-th & BrotZeit), Drs. Mimi Arandjelovic and Hjalmar Kuehl (Chimp&See / MPI-EVA Pan African Project: The Cultured Chimpanzee), Bálint Balázs (Environmental Social Science Research Group), Pieter van Boheemen (Waag Society's Open Wetlab, Public laboratory for biotechnology), project manager from community mapping project on urban air quality in the EU, Douglas Tait (Community Driven Science: Understanding the Impacts of Coal Seam Gas Development in Australia), Jemina Stuart-Smith (Redmap Australia), Ross Goldingay (Lismore Koala monitoring), Rose Herben (EstuaryWatch), Peter Brenton (Atlas of Living Australia), Renae Sayers (Fireballs in the Sky), and Dr. Debra Stokes (Southern Cross University, Australia).

This research was made possible through the generosity of the Alfred P. Sloan Foundation.

Commons Lab

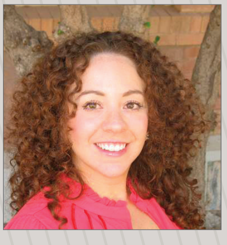
ABOUT THE AUTHORS

CLAUDIA GOBEL

Claudia Göbel is project manager at the European Citizen Science Association and guest researcher at Museum für Naturkunde Berlin. Her background is in science and technology studies. She works on the organization of citizen sciences, links to open science and policy engagement.

VICTORIA Y. MARTIN

Victoria Martin is an environmental social scientist with research experience in a wide range of environmental impact and management issues in Australia and New Zealand. Her recent PhD thesis focused on public engagement in marine citizen science in Australia. Victoria is now a Rose Postdoctoral Research Fellow with the Cornell Lab of Ornithology, USA, where she is continuing her research into citizen science.



MONICA RAMIREZ-ANDREOTTA

Dr. Mónica Ramírez-Andreotta is an assistant professor of Soil, Water and Environmental Science with a join appointment in the College of Public Health and Director of the Integrated Environmental Science and Health Risk Laboratory at the University of Arizona. Her research programs include: developing a fundamental understanding of the fate and transport of contaminants in plant-soil systems, building citizen science programs to increase public participation in environmental health research, creating low cost environmental monitoring tools to improve exposure estimates, and designing data visualizations and risk communication strategies to improve environmental health literacy.

CONTENTS

FOREWORD

I first began thinking about citizen science data standards during the summer of 2013, as a Research Intern with the DataONE Public Participation in Scientific Research (PPSR) Working Group. DataONE was interested in what we called "Project Metadata," perhaps more accurately defined as a set of key terms that could help databases like SciStarter, the Cornell Lab of Ornithology, and CitSci.org share records. It quickly became clear that the value of developing a common, shared vocabulary for talking about citizen science extends far beyond the ability to exchange database records. Standardized metadata documentation promotes the re-use of information by allowing researchers outside of the immediate project team to make decisions about fitness for use. And developing a common data model can allow researchers working on different scales and research domains to exchange information, thus scaling the impact of any single citizen science activity.

The promise of citizen science data interoperability is significant, and recognized by a number of organizations around the globe. By 2015 key initiatives were led by the European Citizen Observatory Web (COBWEB) project, which began developing a common data model through the Open Geospatial Consortium (OGC), and the Atlas of Living Australia(ALA), which began designing an Australian citizen science project database and structured data collection protocol supported by the BioCollect tool. We were excited by these initiatives but also concerned that without comprehensive planning, data standards advanced by a handful of technologists and community leaders would ultimately be imposed on, rather than co-developed by, the diverse global citizen science community.

To address these concerns Claudia Goebel, Elizabeth Tyson, and I conceived of this Stakeholder Analysis at the 2015 meeting of the European Citizen Science Association (ECSA). Through this report we find empirical evidence for the importance of data standards in citizen science, for example learning that some authorities may not use citizen science data because of "*uncertainty about data quality assurance and quality control measures, and a lack of data standardization practices*" (p. 27). Armed with this knowledge, citizen science projects can decide to adopt the standards endorsed by formal authorities to make it easier for their data to be used. Or they may deliberately create their own standards. Some citizen science communities who use bucket sampling to measure air quality, for example, design their own protocols to highlight the absence of existing standards for certain pollutants or point out discrepancies between existing standards for monitoring.[1]

This report uncovered a number of barriers to the adoption of data standards in citizen science. Some projects may doubt the applicability of standards to their research goals, or fear losing relevant information if standards are too general or vague (p. 23). Once articulated, these challenges can be addressed by future initiatives. We offer this report as a roadmap to help coordinate and inform future work on data interoperability, so that any standards produced can be valued or at least understood by a diverse range of citizen science stakeholders around the globe.

ANNE BOWSER
CoDirector,
Commons Lab
Science and
Technology
Innovation Program
Wilson Center

1 Ottinger, Gwen. "Social Movement-Based Citizen Science." Cavalier, Darlene and Kennedy, Eric, Eds., The Rightful Place of Science: Citizen Science, Tempe, AZ: Arizona State University, 2016.

EXECUTIVE SUMMARY

This study examines citizen science projects, with particular emphasis on the stakeholders involved in data and knowledge generation and the use and reuse of information, with the aim of informing data and metadata standardization and interoperability initiatives. It draws on semi-structured interviews with 16 CS projects with different governance models, disciplines, and project aims in the United States (U.S.), Europe, and Australia, along with two citizen science project catalogues and two potential data users in government and academia. In this study, interoperability is defined as the ability of different information technology systems and software applications to communicate, exchange data, and use the information that has been exchanged. It rests on the development of standards and seeks to give parties outside the immediate project team, i.e. external and potential stakeholders, access to data and knowledge.

Achieving greater interoperability would enable data generated by different citizen science projects to more easily be reused by diverse parties (e.g., volunteers, researchers, and decision-makers); by combining data sets of different scales (e.g., local, regional, national, and global scales); and, by combining different types of data (e.g., to help answer transdisciplinary research questions). For example, water quality information collected by one local citizen science project could be combined with diverse types of environmental monitoring or biodiversity data from a regional network and data collected by a public health program. This would enable scientists to advance research on a range of topics, like by using plant phenology to predict allergy seasons, understand how lead concentrations in drinking water relate to lead levels in children's blood, or combine data on avian presence, distribution, or health air quality data.

Main findings: The projects reported a broad spectrum of stakeholders who engage in project decision-making, provide support, and/or use project data. Six stakeholder groups were identified in the interviews: (1) Civil society organizations, informal groups and community members; (2) Academic and research organizations; (3) Government agencies and departments; (4) Participants; (5) Formal learning institutions such as schools; and, (6) Businesses or industry. Barriers to the involvement of a wide range of stakeholders included: a lack of awareness by these stakeholders about the project, or vice versa; difficulty in accessing or knowing how to access potential stakeholders; and, time and resource constraints for doing so.

Perspectives on data standardization and interoperability efforts were heterogeneous across projects. Benefits were seen by many interviewees, although overall the understanding of what interoperability means was limited. Many projects either used research domain specific, regulatory, or community standards while making their data accessible via domain specific databases, or shared online tools. Some projects also built infrastructure for data interoperability. For projects dealing with qualitative data or focusing more on non-data related activities like knowledge sharing or devising experimental designs, it was largely unclear what data interoperability could mean to them and how they would be impacted. Main concerns raised regarding the promotion of interoperability in citizen science included limited applicability due to the natural science bias of standards; costs of adopting standards; doubts about adaptability to local circumstances; resistance to curtailing stakeholder participation and passing of burden; fear of losing relevant information and decreasing data quality; and, difficulties in agreeing on common metadata terms.

Conclusions: Interoperability is only slowly becoming a topic of concern in the citizen science community. The heterogeneity of data sharing practices and adoption of diverse types of standards represent major challenges for interoperability initiatives. A broad concept of interoperability is needed to work across disciplines and project types. Advancing interoperability both facilitates and rests on the involvement of internal, external and, if possible, potential stakeholders.

Recommendations: (1) Citizen science project managers should invest time at the outset of projects to identify stakeholders who could use, and potentially reuse data and knowledge. Existing data and metadata standards should be used for data management whenever possible. (2) Interoperability initiatives should be transparent, open to all types of citizen science projects, involve internal and external stakeholders, and consult potential stakeholders. Shared citizen science data and metadata standards should be adaptable to the needs of citizen science projects, leverage existing community standards, and be open to review and extension. Effort should be dedicated to building a concept of interoperability that goes beyond data, the natural sciences and information technologies. Future interoperability initiatives should embrace plurality in a comprehensive way. (3) Citizen Science associations and other networks should offer capacity building on interoperability and facilitate the adoption of data and metadata standards.

Photo Credit: Gardenroots Citizen Science Project

Citizen Science at Work

A community gathering for data sharing through the Gardenroots project. A university student discusses site-specific soil and plant quality data with a community garden group. Under Gardenroots, community members collected soil and culturally relevant crop samples from their community garden. Ensuring that projects can co-design protocols on the community level while still generating data that can be re-used by external researchers is a key challenge of interoperability projects.

INTRODUCTION

1

Mobilizing data for research and making data count as evidence for decision-making are among the major driving forces of citizen science and other forms of participatory research.[2, 3] Data and information access, licenses, ownership, and quality are central topics for citizen science endeavors. Considering the diversity of emerging projects along with the specificities of how they are implemented locally, it is no wonder that interoperability - or, the ability of different information technology systems and software applications to communicate, exchange data, and use the information that has been exchanged - is key to making citizen science research transformative and more widely accessible to science, policy, and society at large (see Figure 1 for a definition of interoperability and other important concepts).

In the past years, various activities have emerged with the aim of fostering interoperability, both by standardizing the information used to describe citizen science projects, and by developing joint data and metadata standards and protocols for data collection and sharing to support exchange and reuse. In 2012, U.S. researchers affiliated with the DataONE Public Participation in Scientific Research (PPSR) working group began working on PPSR CORE,[4] initially designed as a metadata standard for describing key facets of citizen science projects to help existing project repositories exchange records. The following year, researchers in Europe started developing SWE4CitizenScience, a common data model that once implemented and deployed will support data interoperability between almost all types of crowdsourcing projects. Actors from around the globe are involved in these projects, including working groups in the U.S.- based Citizen Science Association (CSA), the European Citizen Science Association (ECSA), the Open Geospatial Consortium (OGC), Biodiversity Information Standards (TDWG), The Committee on Data for Science and Technology (CODATA), a COST Action on citizen science, platforms like CitSci.org, SciStarter and the Atlas of Living Australia, the Joint Research Center of the European Commission, the Global Biodiversity Information Facility, and the Woodrow Wilson International Center for Scholars.

Citizen science programs each have different technical requirements, project goals, interests, and stakeholders. Adopting a common and shared vocabulary for describing citizen science projects and citizen science data (hereafter: "data standards") can allow information collected by, for example, a local air quality monitoring project to be understood and re-used by researchers and communities working on other related topics, such as public health issues, at national and even global scales. However, any standardization effort also poses challenges, including technical hurdles in working with diverse and heterogeneous communities of practice as well as inclusiveness questions regarding transparency and openness for co-creation and co-design in the process of defining both common and project-specific principles and tools.

Figure 1: Key Concepts

Key Concepts

This study seeks to explore the range of stakeholders of citizen science projects to better engage them in using data standards to promote interoperability in citizen science.

In the context of this study, a **stakeholder** is an individual or organization that contributes to realizing a citizen science project, has a vested interest in a citizen science project, and/or benefits from the research activities and data produced.

Data standards include common vocabularies, information formats, and protocols that may be used to exchange information about citizen science projects, datasets, and data.

Interoperability is the ability of different information technology systems and software applications to communicate, exchange data, and use the information that has been exchanged.

Adopting a common and shared vocabulary for describing citizen science projects and citizen science data ("data standards") can allow information collected by, for example, a local air quality monitoring project to be understood and re-used by researchers and communities working on other related topics, such as public health issues, at national and even global scales. However, any standardization effort also poses challenges, including technical hurdles in working with diverse and heterogeneous communities of practice as well as inclusiveness questions regarding transparency and openness for co-creation and co-design in the process of defining both common and project-specific principles and tools.

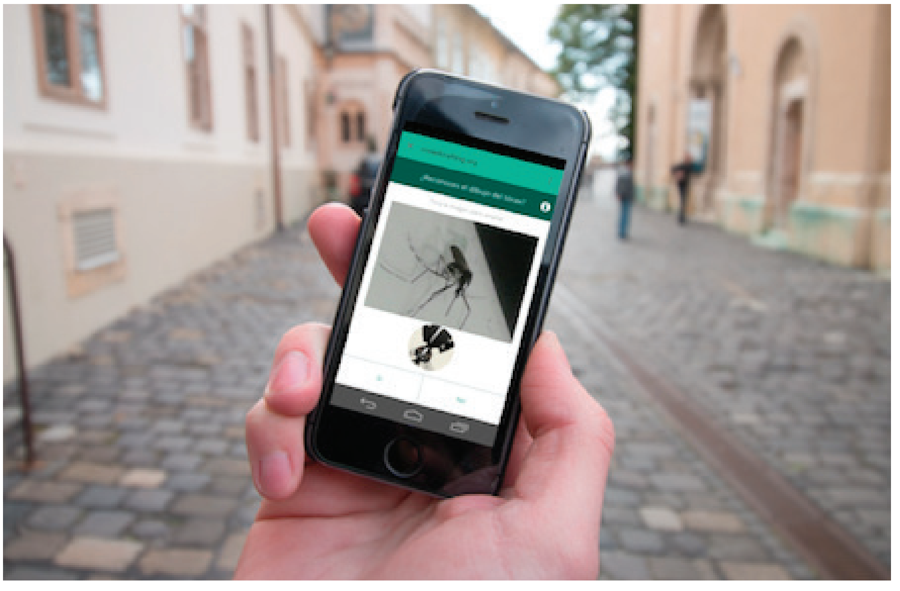


Photo Credit: Mosquito Alert Spain

Global Citizen Science

The Global Mosquito Alert is a consortium of citizen science projects committed to developing a common protocol and shared platform for mosquito monitoring. This is one example of a citizen science data interoperability initiative that will benefit from involving diverse stakeholders around the globe.

AIMS AND METHODS

2

In the context of ongoing work on data standardization in citizen science, this exploratory study aims to (1) understand which stakeholders or parties may be impacted by these efforts, and (2) identify potential consequences of standardization and interoperability. The focus is on individuals, groups, and organizations involved in citizen science activities, as well as those affected by or potentially benefiting from the data and knowledge generated in such projects. On this basis the study seeks to inform ongoing and future standardization and interoperability efforts as well as to facilitate broader participation by global citizen science communities.

To accomplish this goal, specific objectives are to:

1. Survey a purposive sample of citizen science projects representing a full spectrum of governance models and participatory approaches, as well as projects from different scientific disciplines.[5]
2. Identify the various types of stakeholders involved in citizen science efforts, and explore additional stakeholders who could potentially benefit from (re) using the data and knowledge created.
3. Gather citizen science project leaders' perspectives on:
   a. Data generation, management, ownership, accessibility, and sharing.
   b. Data standardization and interoperability efforts.

A stakeholder analysis is a methodology originating in the social sciences used to identify parties that are or will be affected by, and thus "have a stake in" a political program or particular issue, such as an oil spill or a plan to construct a factory in a given location. While it is a common approach in business and natural resource management, it has become popular as a preliminary study for designing and implementing broad stakeholder engagement mechanisms in recent years.[6] In the context of this study, a stakeholder is understood as an individual or organization that contributes to realizing a citizen science project, has a vested interest in a citizen science project, and/or benefits from the research activities and data produced. While a broad spectrum of contributions, e.g. money, time, equipment, etc., is considered, special attention is paid to the generation and use of data, knowledge and similar outputs.

SAMPLING PROCEDURE

Given the complexities and variability within and between citizen science projects, and the exploratory nature of this study, a qualitative study design was selected. While this precludes a broad study of a large number of projects, the method instead allows for exploration of the issues in greater depth than can be achieved through quantitative research. Potential projects for the study were identified through emails to regional citizen science networks, online searches, and researchers' familiarity with projects through their work. The project leaders were contacted with a request for an interview, all of whom agreed.

Based on a purposive sampling strategy,[7] a total of 16 projects[8] were selected to represent different categories of citizen science governance models across different scientific disciplines and regions (Table 1). The five governance models[9] are: contractual, contributory, collaborative, co-created and collegial. These models represent different roles and responsibilities members of the public may play in citizen science research. This well-recognized typology enables the study to cover a wide range of activities and stakeholders. Regarding scientific disciplines, projects were selected to ensure representation of various fields from natural sciences, social sciences and humanities in order to include different methodologies and research outputs. The selection of scientific disciplines for the sampling framework was also informed by previous research.[10] Three regions (Europe, USA and Australia) were chosen to provide an international perspective in the global citizen science landscape. However, due to the limited number of interviews, no comparisons are offered between geographic regions.

The majority of projects selected during the initial round of sampling focused on the collection of data through direct observation. The initial sample was then extended to support a broader perspective on interoperability. In addition to data collection through direct observation, other projects surveyed use qualitative methodologies or are concerned with collecting physical samples, use sensors to collect data, analyze data online, or conduct lab experiments. In addition to initial interviews with citizen science projects, four additional interviews were conducted to explore the perspective of key stakeholders identified during the first round of interviews. These additional interviews were conducted with representatives from research infrastructures including a national citizen science database (Atlas of Living Australia) and an international citizen science project repository (SciStarter), both of which are involved in ongoing interoperability initiatives, as well as an environmental scientist and a government employee (both as potential data users).

Interview Protocol and Data Analysis

The stakeholder analysis interview protocol was guided by a review of relevant literature and discussions with research domain experts, including an exploratory stakeholder mapping exercise at a citizen science data workshop held by the Joint Research Centre of the European Commission.[12] The interview questions were developed by the

Table 1. Sampling framework*

	Discipline			
Citizen Science/ PPSR Governance Model	**Biology**	**Environmental & Earth Sciences**	**Social Sciences & Humanities**	**Astronomy and Planetary Sciences**
Contractual		CSG impacts (AU)		
Contributory	Koala monitoring (AU) USA National Phenology Network - Nature's Note-book (U.S.) eBird (U.S.) BioBlitz Barcelona (EU)	EstuaryWatch (AU)		Fireballs in the Sky (AU)
Collaborative	Redmap (AU) Chimp&See (EU)			Globe at Night (U.S.)
Co-created	Public laboratory for biotechnology (EU)	Grandlake Watershed Mercury Study (US) Community mapping of urban air quality project (Anonymous, EU)	Community supported agriculture study Hungary (EU) Science and Youth / BrotZeit (EU)	
Collegial		"Civic View from Above" - Public Lab's (U.S.) DIY aerial photography toolkit in spaces of political conflict (project takes place in Israel/Palestine)		
Other interviews	**CS Databases and Project Repositories:** • SciStarter (U.S.) • Atlas of Living Australia (AU) **Potential External Stakeholders (not currently using CS data):** • Environmental scientist (AU) • Health agency (U.S.)			

**Projects included in the study according to Shirk et al.'s[11] PPSR framework, scientific disciplines and regions (Australia =AU, Europe = EU, United States of America = U.S.).*

authors and Wilson Center staff to cover several themes, including for each project: a description of the project; stakeholders; data generation and management practices; degree of interoperability with other projects; and the project's perspective on data standardization in citizen science (a full interview protocol may be found in the Online Appendix). Each interview lasted between 45 and 90 minutes. Responses were recorded in a research journal, with initial results categorized by each author individually using thematic analysis.[13] The three authors then compared their research notes and early analysis to iteratively define the categories of stakeholders presented below, as well as highlight recurring themes and major tensions. In addition to interview questions and research journals, a selection of exemplary and broad stakeholder groups from the literature review[14] was used to prompt further reflection by the authors. The stakeholders mentioned by interviewees were then grouped and sample categories adapted.

> While some barriers to data openness and accessibility should be addressed, others highlight the fact that (partial) non-openness is functional, rather than dysfunctional, for the operation of citizen science projects and thus should not simply be treated as "barriers to access."

Interview Participants

Europe
Barcelona, Spain
Klagenfurt, Australia
Leipzig, Germany
Budapest, Hungary
Amsterdam, Netherlands
Brussels, Belgium

United States
Tucson, Arizona (2)
Philadelphia, PA
San Francisco, CA
Boston, MA
Ithaca, NY
New Orleans, LA

Australia:
Lismore, Australia (3)
Hobart, Tasmania, Australia
Melbourne, Australia
Canberra, Australia
Perth, Australia

Interview participants in the United States, Europe, and Australia. We invite future researchers to expand the scope of inquiry to new geographic areas.

RESULTS

3

STAKEHOLDERS OF CITIZEN SCIENCE PROJECTS

The citizen science projects in this study reported a broad spectrum of stakeholders who engaged in project decision-making, provided support (in the form of time, funding, expertise, equipment etc.), and/or were users of the project data. Six stakeholder groups were identified in the interviews. These groups were: (1) *Civil society organizations, informal groups and community members*; (2) *Academic and research organizations*; (3) *Government agencies and departments*; (4) *Participants; (*5) *Formal learning institutions such as schools*; and, (6) *Businesses or industry.*

Nature of stakeholder involvement

These stakeholder groups exhibited varying levels of involvement, according to how frequently they were mentioned and their types of contributions. The spectrum of involvement in projects ranged from a **high level of involvement** (which describes stakeholders who participate in many different aspects of a project, performing roles such as project leads, decision-makers, supporters, and/or users of the data or knowledge) to a **low level of involvement** (which describes stakeholders who participate in only one or two roles in the project).

Stakeholder groups such as *Civil society organizations, informal groups and community members*, and *Academic and research organizations* reported the highest level of involvement in the citizen science projects in this study. Group (1) refers to formally organized (incorporated) and less formally or informally organized stakeholders from civil society, including activist and advocacy groups, as well as individual community members. *Civil society organizations*, *informal groups* and *community members* participated in different ways for all governance models. This group was found to support, collect and/or use the data or knowledge in all other governance models, and lead or make decisions in collaborative and co-created projects. In all governance models, members of communities, or community groups also collected data, provided indirect support such as access to land and used project data or knowledge. The study also observed that the community members/groups who participated in project decision-making for collaborative, co-created and collegial projects also led many aspects of project design and implementation. For example, in one collegial project, community members decided what tools to use, received training online or through a local organizer, and then created their own projects, which is the inverse of a contributory or contractual project. *Academic and other research institutions* were also highly involved in citizen science and formed the largest group of lead organizations for the projects, and were present in projects of all governance models. As a stakeholder group, academic and research organizations were also instrumental in decision-making and support, and made use of the project data or knowledge.

Government agencies and departments were mentioned as important stakeholders in all governance models, with the goal of influencing policy outcomes. Compared to other stakeholders, the presence of this third group in the citizen science projects was moderate. Government involvement in citizen science was strongest in contributory projects, where it was found to fund, lead projects, make decisions, and/or support and use the data/knowledge. Government support, decision-making and/or data or knowledge use was also present in collaborative and co-created projects, yet appeared in the contractual and collegial projects only as a potential data user. *Participants* represented the volunteers[15] who contributed to the projects across all governance models. While these participants were present in all projects, the number of roles they played in the projects was variable, and usually fewer than the group leading the project.

The final two stakeholder groups were more removed from the core functions of the citizen science projects. *Formal educational institutions* were not involved in the contractual project. However, these institutions supported contributory and co-created projects (especially for data collection) and used citizen science data or knowledge from contributory, collaborative, co-created, and collegial projects, primarily for educational purposes. Two projects included schools in decision-making, one of which was a co-created project, and the other a collaborative project. *Industry and business* involvement was absent in the large majority projects. However, they were more deeply involved in some of the co-created projects, where they assisted decision-making, provided support, and used the data and knowledge. These were also identified in the contractual project (as a data user) and contributory projects (for support and data use), yet were absent from the collaborative projects interviewed.

Potential for engaging new stakeholders

Potential stakeholders who were not currently involved in projects or using citizen science data were mentioned by all governance levels except the contractual project. The list of stakeholders for possible future collaborations included *Local, state, or federal government agencies or departments*; *Business or industry*; *Civil society organizations, informal groups and community members*; and, members of other *Academic and Research Orga-*

nizations. The projects indicated they would like these stakeholders to provide support and/or use the data and knowledge generated by the projects. Barriers to involvement included a lack of stakeholders' awareness of the projects (and in some cases, a lack of the project team's awareness of potential stakeholders), difficulty in accessing or knowing how to access potential stakeholders, and time and resource constraints for doing so, as discussed in depth below.

Internal and external stakeholders

Stakeholders could be distinguished as being **internal** and **external** to a given citizen science project based on whether interviewees considered stakeholders as part of the immediate project team or as independent contributors or users of results. Consequently, SciStarter and the Atlas of Living Australia (**actual** stakeholders), the US health agency and the Australian scientist (**potential** stakeholders), need to be considered external stakeholders, since they were neither

Table 2: Matrix of stakeholder groups

Examples of stakeholder types identified by interviewees	**Internal** Stakeholders form part of the immediate project team and research process(es)	**External** Stakeholders are not considered part of the immediate project team and research process(es), yet have a vested interest in the project
Actual Stakeholders that are currently or have been involved with a CS project	• Partners/collaborators setting the research question in a collegial project. • University researchers leading and analyzing data from their contributory CS project. • Government department selecting location for CS activity. • NGO hosting a community lab. • Neighbor committee or community members disseminating project findings. • High-school students doing CS as part of their coursework.	• SciStarter using CS project metadata. • Atlas of Living Australia making CS data accessible. • Volunteers in a contributory project analyzing data. • Local business informing customers on CS activity. • Schools using CS activities in formal education.
Potential Stakeholders that have not been involved with a CS project, yet could be in the future	• Federal and state agencies funding project and using CS data. • Special interest group participating in knowledge generation.	• US health agency using CS data; Australian scientist and "other researchers" using CS data.

mentioned as part of a citizen science project by any of the other interviewees nor do they operate their own projects. All six stakeholder groups listed contain internal and external stakeholders (Table 2). The distribution of internal and external stakeholders varied according to the governance model. For instance, in co-created and collegial projects, the project team was usually much larger than in other governance models, and included various stakeholders in making decisions. These stakeholder groups were seen as external parties in contributory projects, among them project participants and local authorities. As elaborated below, this distinction is important for understanding data sharing practices and promoting interoperability.

DATA SHARING, ACCESSIBILITY, USE & REUSE

Projects collect and share heterogeneous types of data and knowledge. These ranged from quantitative data (e.g. on air quality, light pollution, biodiversity), to qualitative information (e.g. on perceptions of air pollution and adaptation strategies or consumption behavior, aerial mapping of land, human rights and political conflict), to knowledge (e.g. on lab experiments, local cultural traditions or social transformation). Other relevant outputs generated by projects included data on project participants, metadata on citizen science projects, maps, programming code, designs, objects, lab experiments, institutions, education, and skills. In all cases more than one output was generated. This suggests that discussions on interoperability should not be limited in focus to projects seeking to gather quantitative natural science data. Even for projects in which quantitative data generation is the central aim, learning along with co-creation methods and other tangible outputs, such as programming code or physical objects, represent valuable resources for potential reuse.

Across projects and geographies, sharing data, knowledge and procedures was generally perceived as important. The value of data and knowledge to science, the environment, volunteers, collaborators, and society at large was underlined. Specific reasons for sharing included enabling comparisons of citizen science data to official records; making the use of data possible for everyone; improving data quality; facilitating cross-border and long-term studies; fulfilling a societal responsibility as recipient of public funding; increasing accessibility to low cost monitoring tools; enabling joint action; and, raising awareness about important issues.

Sharing and accessibility of data and knowledge

While all projects used various channels to share their raw data, results, and/or procedures, the level of accessibility varied. Some projects allowed unhindered access to raw data sets, for example as files could be downloaded directly from project websites or were shared through discipline-based repositories and archives. Others only granted restricted access to raw data, for instance, sharing them upon request or through online interfaces that support analysis and interpretation, sometimes following embargoes. Instead of or in addition to direct access to raw data, some projects shared processed data, study results, or procedures through publication in academic journals (not necessarily open access), public reports, online graphics, repositories, or social media. In one case, due to potential conflicts and local constraints, the data was only shared and used by the community that generated the data, in hard copy format.

The reasons interviewees offered for not making project data more accessible included:

- Technical difficulty and lack of technical capacity in the project team.
- Sensitivity of subjective knowledge and personal data at the center of the project.
- Privacy considerations that might hamper participants' willingness to contribute.
- Difficulty finding journals that publish datasets.
- Embargo periods ensuring privileged access to the research group for conducting. analyses and publishing results before others can do so.
- Perceived lack of interest in qualitative raw data.
- Concern about impacts on high-value or threatened species.
- Potential for political conflict.
- Perceived lack of interest beyond actual internal stakeholders.
- Language barriers.

While some barriers to data openness and accessibility should be addressed, others highlight the fact that (partial) non-openness is functional, rather than dysfunctional, for the operation of citizen science projects and thus should not simply be treated as "barriers to access." The reasons for limiting accessibility mentioned above are rooted in the complex interplay of disciplinary traditions, methodological questions, project aims, and the design and implementation of the project, all of which merit further investigation.

Across all projects, results were usually shared among internal stakeholders, and rarely with external stakeholders. A comparison of project governance models illustrates three aspects of how different conceptions of internal and external stakeholders manifest in nuanced sharing practices. First, *raw data sharing* was generally practiced within the project team, and less so with external parties. Co-created projects usually gave participants access to raw data (as well as to procedures and results) while only some contributory projects did so. Contributory projects, in turn, were more likely to share data with external stakeholders for analytical or environmental management purposes. Second, for contributory researcher-driven projects, *sharing raw data* was usually understood as making project data *available to the research community*, while *sharing processed data or results*, especially in the form of graphics, maps, or report cards was understood as a form of making their data *accessible to the public and/or contributing participants*. Third, the collegial project provided yet another perspective of how *sharing knowledge* - in this case tools and methodologies - is linked to the social structure of citizen science activities. The community from which the project stems was built around the use of open source tools. The tools were openly available for everybody to choose which to apply to their own projects. In turn, tool users were encouraged to share their experiences within the community to improve the methodologies, adding rigor and expanding the applicability of the tools. Thus, accessibility, use and reuse of knowledge and methodologies are ingrained in the functioning of the projects and larger community.

Use and reuse of data and knowledge

Examining how stakeholders are involved in project governance is also relevant for better understanding the use of citizen science data as well as of knowledge and procedures. In most cases, data use by external stakeholders was limited, non-existent or unknown. While some projects said there were no barriers to their data, several believed that potential users were unaware of the data or the project itself. Several co-created projects stated they had already involved all relevant stakeholders as project partners (e.g. by co-designing study questions or choosing activity locations), whom they considered having an interest in using the data, such as local or national authorities. Although these co-created projects give various stakeholders access to results, future (re)use by potential new internal or external stakeholders is rather neglected in these cases. Some projects that emphasized sharing knowledge and procedures reported their methodologies had often been used and requested by external parties seeking to use the project or replicate co-creation methodologies and experimental designs they developed. This highlights the importance of sharing project results beyond simply data.

To understand the motivations of potential external stakeholders for using citizen science data, interviews were conducted with an employee of the federal health agency in the U.S. and an environmental scientist in Australia. Neither interviewee was using citizen science data, citing the main reasons as uncertainty about data quality assurance and quality control measures, and a lack of data standardization practices for environmental sampling. The Australian scientist also mentioned a lack of awareness of citizen science data sets amongst scientists generally and, more specifically, her lack of knowledge about the location of data sets, ownership and publication or access rights. Given these concerns the U.S. health agency employee said they prefer to stick to conventional data collected following U.S. Environmental Protection Agency (U.S. EPA) standard methods that have been published in scientific literature. The Australian scientist said the adoption of standards in citizen science would help alleviate some concerns about the influence of externalities, the procedures used, and allow for multi-region comparisons, all of which is important for the type of research this scientist undertakes.

DATA STANDARDIZATION AND INTEROPERABILITY

In the interviews, interoperability was defined as "the ability of different information technology systems and software applications to communicate, exchange data, and use the information that has been exchanged," which rests on the adoption of data and metadata standards. Further questions explored what interoperability meant for projects, and what practices, if any, they were undertaking to ensure project data or other information could be easily shared.

What is currently being done by projects and rationale

Interoperability as defined here was not considered by most interviewees and if so, only infrequently in the beginning of projects. At the same time, exemplary work on making project data interoperable was conducted in some of the surveyed projects. Interviewees from eBird (U.S.), BioBlitz Barcelona (Europe), and Atlas of Living Australia said they use established scientific, industry and/or regulatory data standards, like Darwin Core. For example, eBird data, with a location and date, can go to a national data repository like the U.S. Geological Survey (USGS) or National Atmospheric and Space Administration (NASA), pull the data and *relate* their data to the location and information they have. Globe at Night, eBird, and USA National Phenology Network are already merging topic-based datasets and have developed or are developing interoperability infrastructures to do this. This approach is effective because these projects collect the same data that can fall under a similar field, using the same terminology. Limitations of this approach include when data of diverse types or from diverse sources need to be combined, as highlighted when an interviewee representing Globe at Night discussed the National Park Service implementing a system to measure light using an entirely different standard of measurement, which creates difficulties when attempting to work with the National Park Service or combine data.

Overall understanding and perceptions of interoperability

In general, interoperability was poorly understood by most interviewees. Nevertheless, across projects—regardless of governance model, geography, or area of study—combining data sets and using joint standards for certain types of data and/or application cases was generally perceived as valuable. Interestingly, most projects had not considered what interoperability means specifically within the citizen science context. They reported using existing standards for publishing data in disciplinary databases, government standards, or sharing documentation through online content platforms such as Wikipedia, YouTube as methods they used to make data interoperable with others. Others simply stated they use scientific methods of data collection and reporting that are accepted within their disciplinary field. For projects that dealt with qualitative data, or are more focused on knowledge sharing, data analysis, or devising lab experiments, it was unclear what data standardization could mean for their fields of activity and how they could and should relate to the ongoing interoperability initiatives. Consequently, although interviewees were aware of the value of sharing, knowledge on how to make project data and results interoperable was often lacking.

When comparing citizen science governance models, interviewees differed in their knowledge about and prioritization of data standardization and interoperability. In some cases, interoperability was well understood. Various large-scale ecological contributory projects were using Darwin Core, an established biodiversity informatics data standard. Contributory projects were also working on the interoperability measures described above. In addition, Hagit Keysar stated that her collegial projects benefit greatly from the Public Lab website where staff, organizers and users post best practices and standards on how to use the technologies, and that this is important for scaling up projects and making the projects and methods meaningful to other stakeholders. It was stressed that the Public Lab methods are informed by the users who contribute online or offline and that this iterative process is about creating and facilitating ways to conduct collaborative work.

On the other hand, interoperability beyond the core project team is not a priority for some citizen science projects. One co-created project stated it was more important to ensure that local community members received the data in a format they could use than adhere to strict data standards. Although the interviewee recognized the benefits of combining data sets through standardization, it is not a primary concern given their specific goals. However, a different co-created project stated that interoperability was not a concern, yet as the interview proceeded, the interviewee realized how much it would have helped their own work if the local and state government agencies and nonprofit organizations had standardized their data.

Concerns regarding interoperability

The interviews revealed six main concerns about interoperability in citizen science:

Photo Credit: Horseshoe Crab Citizen Science Project Cedar Key.

Florida Sea Grant

It's spring and that means horseshoe crabs are nesting in high numbers along many of Florida's beaches. A new citizen science program in the Nature Coast is helping state managers collect important data about horseshoe crab populations. This program is a collaboration between Florida Sea Grant, UF IFAS Nature Coast Biological Station, FWC, and UF's Biology Department. Sea Grant Agents like @scbarry are training citizens to count nesting horseshoe crabs and apply tags to the crabs as part of a long-term population study and you can help! Keep your eye out for tagged horseshoe crabs the next time you are walking the beach or near the water.

Florida Sea Grant GIS Workshop May 2017

- **Limited applicability due to natural science bias.** Several interviewees were critical of the fact that social science standards remain largely absent from discussions, highlighting that the interface of any general citizen science data standard would need to be accessible to a wide range of projects and research methodologies. One project interviewee argued for a more comprehensive conception of interoperability pointing out that the one put forward in the stakeholder analysis was focused on the natural sciences in general and specifically on observational data, and not applicable to all forms of knowledge. Further, for the projects working with qualitative data it was unclear what such a standard could look like. Another issue highlighted by interviewees included the lack of clarity on how to treat data gathered on participants including demographic information and participant evaluations.

- **Costs of adopting standardization protocols and modifications to existing infrastructure.** Projects were willing to adopt standards in principle as long as they could be incorporated in a way that did not fundamentally change how the project currently operates, or change the usefulness of the data already collected. There was a great deal of uncertainty regarding what data standards for citizen science might look like, and how they might dictate necessary changes to already established projects and their information and technology (IT) systems. Interviewees said for new standards to be adopted, they would need to be easy to incorporate into existing data collection processes, and not render existing data useless. Interviewees also worried how efforts to standardize may create significant work for project teams, specifically for their computer information and technology specialists. The costs involved in implementing such changes may present a great challenge for a considerable portion of the citizen science community since many of the interviewees said they were operating on relatively small budgets (and in some cases, in-kind support) for IT services. These budgetary constraints will likely limit the uptake of interoperability measures.

- **Doubts about adaptability to local circumstances.** A co-created project expressed concern on how to develop standards and interoperability with locally-based projects. While they recognized there would not be an off-the-shelf version that suited all projects, they suggested a template could be provided. However, even if the template was provided, this interviewee said they would have to work with their community to build the infrastructure and refine it, concluding that doing this is not easy at a local scale.

- **Resistance to curtailing participation and passing of burden.** Some interviewees expressed concern about decreasing participation in projects arising from a requirement for participants to set up and/or follow a set of standards. Interviewees said they would consider adopting a general standard for citizen science data when certain preconditions were met, including not limiting participation by requiring participants to conduct significantly more work, or by other barriers to participation (such as reduced privacy).

- **Fear of losing of relevant information and decreasing data quality.** One interviewee with extensive experience in data standardization stressed that unintended consequences can arise from standardization; for example, the data may become too general, richness is lost, and metadata may become broader. This interviewee said every project is collecting data differently based on the area of research or discipline. They emphasized that all researchers need to agree on the metadata fields that need to be included across projects, and only data that is managed and curated with the same quality and rigor as their own projects can be integrated. This interviewee thought it was critical for projects to maintain high standards of data management and quality and "*to be able to control for and accurately describe all the bias in the dataset – this is more important than standardization.*" However, a second interviewee suggested Monarch Watch as a successful example where existing project databases were merged to create a more robust dataset and perform rigorous analyses.

- Difficulties in agreeing on concrete metadata terms. Where scientific research is highly specialized, there was some question about the value of adopting broad-reaching standards (e.g. set of properties, elements, fields, columns, or attributes), which may not be appropriate for certain projects. One interviewee stated that the main metadata fields needed to allow for "data relatedness"[16] are: who collected the data, where the data was collected, when the data was collected, and how the data was collected. Currently, most projects do not collect the *how* information.

DISCUSSION

4

From the exploration of stakeholders, data accessibility and use practices in citizen science projects, several topics have emerged that are relevant to standardization and interoperability discussions.

INTEROPERABILITY IS ONLY SLOWLY BECOMING A TOPIC OF CONCERN IN THE CITIZEN SCIENCE COMMUNITY

This became evident from the lack of clarity about what data accessibility means and what activities are related to interoperability, as well as the considerable variation in the understanding of who owns the data generated in citizen science projects (ranging from the community to only the project team). Clearly, there is a need for communication and training material on making citizen science data, methods and knowledge open and interoperable as well as sharing of good practice. While there is no overarching body responsible for providing this service to projects, this is perhaps an area that could be more comprehensively addressed by citizen science associations across the globe. As the momentum of citizen science practice continues to grow, these associations are well-placed to encourage the community to think about and implement interoperability mechanisms, including data and metadata standards.

The fact that interoperability measures are currently only implemented by the contributory and collaborative projects surveyed in this study suggests a vast potential for reaching projects with other governance models. In addition, projects from the social sciences and humanities using qualitative methodologies were highly uncertain about the applicability and impact of standards to their field. This highlights challenges for standardization across scientific disciplines. For qualitative data, there needs to be more exploratory work done on what standardization could mean, why it would be useful, and what it would look like. More research is also needed to understand differences of knowledge production across scientific disciplines. In addition, current initiatives on interoperability were largely unknown to the interviewees, as illustrated by the fact that associations like

ECSA or CSA (who are supporting those initiatives) were not mentioned as stakeholders by interviewees. These initiatives need to become more transparent and inclusive to engage more projects and subcommunities in citizen science. In particular, it would be useful to specify ways in which interested projects could engage in interoperability initiatives and reach out to a broader community. This issue may be best addressed through the support of citizen science associations themselves, as further discussed below.

HETEROGENEITY OF DATA SHARING PRACTICES AND ADOPTION OF STANDARDS FOR INTEROPERABILITY

Considering the breadth of citizen science project designs and research methodologies, the clearest benefits of interoperability appear to persist among projects collecting similar types of data and working with the same media or species. For biodiversity monitoring, a body of standards (such as Darwin Core) improves interoperability and promotes connection to environmental health efforts, yet differences in the ways in which diverse types of data are collected, stored and shared means broad-scale adoption of interoperability will not be simple. Yet, this is an issue that requires urgent attention, particularly since there has been limited uptake within CS practice, because one of the great benefits of interoperability is the potential to contribute to avoiding environmental disasters. Incidents like the Flint, Michigan (U.S.) lead catastrophe might be avoided in the future if water quality data collection efforts were standardized, not just with citizen science, but also between federal agencies, such as the EPA, USGS, and NASA, which often have their own standards and procedures for collecting, storing and sharing data.

A BROADER CONCEPT OF INTEROPERABILITY IS NEEDED TO WORK ACROSS DISCIPLINES AND PROJECT TYPES

Several interviewees criticized the concept of interoperability employed in this study for being too narrow in its focus on information technologies, natural sciences and quantitative data, which are typical characteristics of mainstream CS projects, and often contributory or collaborative governance models. Such a perspective misses the other various concrete outputs generated in CS projects, such as tools, experimental design, and programming code, as well as the impacts of some citizen science projects, especially learning and transformative action. The later are more common in projects that enable greater involvement of the participants in decision-making and project design, for example co-created or collegial projects. Plurality—including through regional, disciplinary, terminological, and methodological considerations—is necessary and enriching for science and consequently it should not be reduced. This means that standards for citizen science data need to enable, not infringe upon different forms of plurality. In order to account for further developments in the field and to allow for various practitioner communities to contribute, standards should continue to have formal review and revision processes in place and in their architecture, as well as

To understand the motivations of potential external stakeholders for using citizen science data, interviews were conducted with an employee of the federal health agency in the US and an environmental scientist in Australia. Neither interviewee was using citizen science data, citing the main reasons as uncertainty about data quality assurance and quality control measures, and a lack of data standardization practices for environmental sampling.

linkages between different sets and types of standards. In addition to existing standards in academic research and industry, standards and workflows originating within the citizen science community itself, such as the approach to open source tools and information embraced by Public Lab, need to be taken into account and built upon. Accessibility to tools and methodologies that are being standardized by the users themselves is critical for facilitating community or civic science activities. This is essential for groups that do not have the existing infrastructure, whether tools or programs, to advocate or investigate community-identified issues. Bottom-up initiatives and project types should be recognized by other stakeholders and strongly considered in interoperability efforts. A more inclusive concept of interoperability is needed that is suitable for various forms of knowledge including tribal ecological knowledge and subjective and experiential knowledge. Interoperability should also allow for inter-generational exchange.

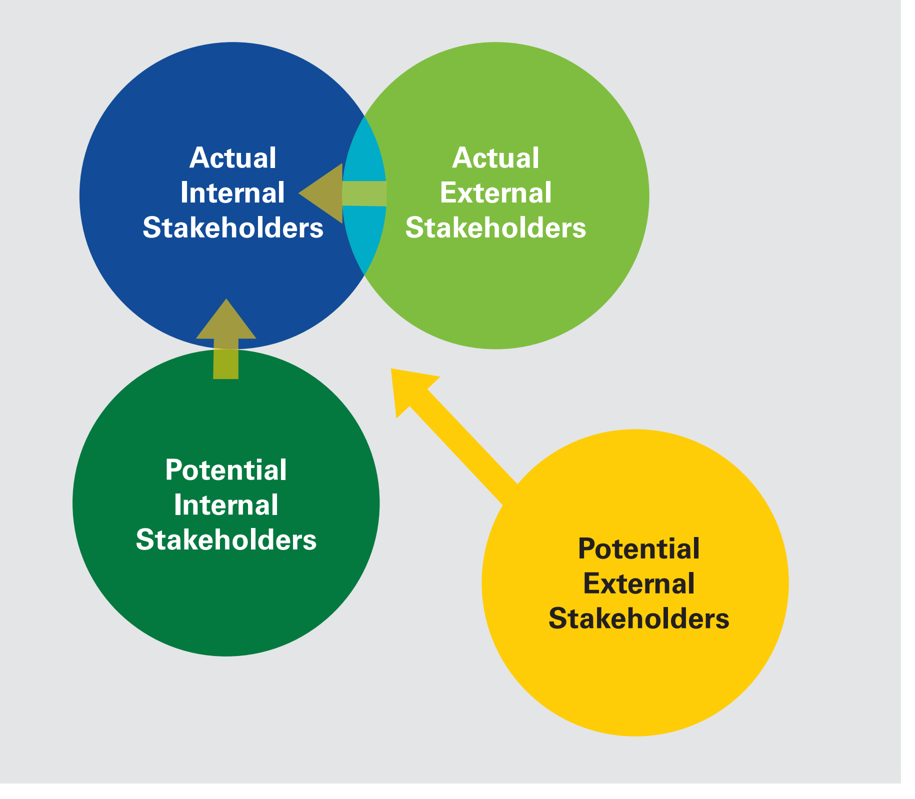
ADVANCING INTEROPERABILITY BOTH FACILITATES AND RESTS ON INVOLVEMENT OF STAKEHOLDERS

The study highlighted the various types of stakeholders, both internal and external to CS projects,[17] which have interests in the production, management and/or use of the project data. Considerable restrictions on the use and sharing of data amongst project stakeholders were detected,[18] which contrasts with an expectation of some that CS data should be open and easily accessible. The implementation of interoperability measures will be essential in resolving this mismatch between expectations of some in the CS community and the reality of CS practice.

With a few exceptions, many of the interviewees in this study think about interoperability as an issue of data usability rather than an issue related to the engagement of internal and external stakeholders. At the same time, potential stakeholders—such as the US health agency representative and the Australian scientist described earlier—report difficulty in accessing the data, knowing that it exists in the first place, and whether data meets quality assurance/quality control standards. Better interoperability can help mitigate these concerns. While for some projects, data sharing and access beyond the core project team will not be important or practical, data interoperability nevertheless remains a major challenge for the progression and impact of citizen science as a diverse input into knowledge production and decision-making. Interoperability also has the potential to make external stakeholders' data and information needs more easily understood and accessible to CS projects, and therefore needs to be considered for both directions of exchange.

It is clear that the implementation of interoperability measures has the potential to extend sharing and reuse of data and knowledge beyond internal project stakeholders, where it is currently concentrated, towards more external and potential stakeholders (Figure 2). In turn, identification of, and consultation with, stakeholders at the planning stages of CS projects will help support interoperability. The same is true for additional work on citizen science data standards, which would benefit from the involvement of both more diverse projects and external stakeholders. Greater and more diverse stakeholder involvement would enable broader and more effective use of citizen science data, methodologies and knowledge in the future.

Figure 2. Potential of interoperability measures for extending sharing of results with stakeholders of CS projects. Most sharing of data and other outputs currently happens among internal stakeholders of CS projects and less with external ones (blue circle). Implementing interoperability measures (arrows) has the potential to facilitate more sharing with external stakeholder and re-use by potential stakeholders.



LIMITATIONS OF THE STUDY

In this exploratory study, a total of sixteen interviewees representing sixteen citizen science projects were recruited through a purposive sampling technique. This relatively small sample may not represent the full landscape of citizen science projects and perspectives. Further, the two additional interviewees—one government employee, and one scientist—are not sufficiently representative of potential external stakeholders. Due to the scope of this research, international differences in facilitators or constraint of interoperability related to government and research systems were not considered. By limiting the projects to the U.S., Australia and Europe, this study excludes vast parts of the globe with long-established participatory research traditions as well as growing citizen science activities. Language, other ways of knowing, and cultural models such as traditional ecological knowledge, were not considered when selecting projects. These are opportunities for future research.

OVERALL RECOMMENDATIONS TO THE FIELD AND,FUTURE RESEARCH

5

Based upon the work described here, recommendations are proposed to guide future citizen science interoperability efforts. These recommendations have been crafted and tailored for citizen science projects, overarching interoperability initiatives, and citizen science associations and networks.

Citizen Science Projects: Regardless of governance model, it is important to identify actual and potential stakeholders (including partners to support the project and/or benefit from using the outputs), and have data and knowledge management and sharing protocols in place to facilitate information use and reuse. It is also critical for projects to maintain high standards for data management and data quality, and to control for and accurately describe biases in datasets. Projects should consider seeking out information on existing data standards, along with similar projects for their experiences, and get support where needed.

Interoperability Initiatives: Current standardization and interoperability initiatives should become more transparent and open to internal and external stakeholders of CS projects. Trust and dialogue must be built with internal and external stakeholders and their experiences and necessities need to be taken into account. Potential stakeholders at various levels (local, national, international), especially regulatory agencies and government, should be included into interoperability initiatives to express their needs and contribute expertise. When developing common vocabulary and interoperability protocols, it is imperative to build in mechanisms for review and extension for new issues and communities in the future to make standards facilitate, not curtail, the innovative potential of CS as a developing approach. Standards should also be flexible and adaptable to various fields of CS research, e.g. implemented as a modular framework like the Creative Commons (CC) licenses, which allow users to select the best choice from a range of options. Existing community standards should be leveraged where possible. Work needs to progress towards a concept of interoperability that goes beyond data, natural sciences and information technologies and embraces plurality by being open and useful for various forms of knowledge, qualitative methodologies, regional specificities, and across different languages and generations.

Citizen Science Associations and Networks: More capacity building is needed for citizen science projects to inform others in the citizen science community about interoperability issues and the benefits of greater interoperability in citizen science. There are numerous ways in which associations and other networks can facilitate greater adoption of data standards and additional work on interoperability in the future. These methods include hosting workshops and presentations, facilitating access to data experts, and providing recommendations and guidance for best practices. In addition, the associations may be able to assist projects in identifying other similar projects that are already collecting standardized, interoperable data, and facilitate inter-project learning. Project identification systems, such as SciStarter and the Atlas of Living Australia's Project Finder, may also be able to assist in matching projects according to data requirements. Finally, associations could provide support for projects to conduct their own stakeholder analyses, for example by providing advice such as toolkits on key aspects of project design. This will help citizen science projects research diverse stakeholders' needs, and ultimately maximize their impact.

However, the provision of information on its own will not be enough to bring about the considerable change in citizen science practice that is required for broad-scale adoption of data standards to promote interoperability. Practices, values, and norms will need to be more deeply embedded in the culture and expectations of citizen science communities. This will require conversations about the importance of these measures, and the impact they can have on wider issues, such as leveraging information for knowledge generation and decision-making, across all sectors and types of citizen science activities, in addition to conversations held within specialized data and metadata working groups.

FUTURE RESEARCH AND PRACTICAL NEEDS

Future research is needed on the role of different stakeholders in citizen science, both in relation to data standardization efforts and in relation to other research and development activities. Other important external stakeholders may include citizen science data repositories, project databases and practitioner associations. Although this study included two citizen science databases/project repositories, neither of these citizen science project types were mentioned as stakeholders by the other interviewees implementing projects. As these actors lead citizen science data interoperability initiatives, we recommend their inclusion in future interoperability stakeholder analyses. To further complement and successfully expand this existing study, more research into the perspectives of a broader set of stakeholders is warranted. The distinction between internal and external stakeholders is offered as a heuristic for linking project governance models with data sharing and interoperability practices, and should be explored further for empirical richness and conceptual depth.

Many of the citizen science projects contributing to this research shared significant practical limitations in terms of access to resources such as funding and technological expertise that would be required to advance the use of data standards.[19] Resources provided by the associations, as suggested above, can help mitigate knowledge gaps. However, funding or technological expertise provided through another channel, such as hackathon-style events to remediate exiting data or database records, is also needed. It is crucial that these resources support the full range of citizen science projects, including small, growing, and grassroots initiatives, in addition to the larger and more established projects that enjoy privileged access to national and international level funding schemes.

There was a great deal of uncertainty regarding what data standards for citizen science might look like, and how they might dictate necessary changes to already established projects and their information and technology (IT) systems. Interviewees said for new standards to be adopted, they would need to be easy to incorporate into existing data collection processes, and not render existing data useless.

ENDNOTES

1 V. Martin completed this study at Southern Cross University, Lismore, Australia.

2 Cooper, Caren, Janis Dickinson, Tina Phillips, and Rick Bonney. “Citizen science as a tool for conservation in residential ecosystems.” *Ecology and Society* 12, no. 2 (2007).

3 Kullenberg, Christopher, and Dick Kasperowski. “What is citizen science?–A scientometric meta-analysis.” *PloS one* 11, no. 1 (2016): e0147152.

4 Citizenscience.org. “PPSR_ CORE Metadata Standard”. October 9, 2015 <http://citizenscience.org/2015/10/09/ppsr_core-metadata-standard>.

5 Shirk, Jennifer, Heidi Ballard, Candie Wilderman, Tina Phillips, Andrea Wiggins, Rebecca Jordan, Ellen McCallie et al. “Public participation in scientific research: a framework for deliberate design.” *Ecology and Society* 17, no. 2 (2012).

6 Reed, Mark S., Anil Graves, Norman Dandy, Helena Posthumus, Klaus Hubacek, Joe Morris, Christina Prell, Claire H. Quinn, and Lindsay C. Stringer. “Who's in and why? A typology of stakeholder analysis methods for natural resource management.” *Journal of environmental management* 90, no. 5 (2009): 1933-1949.

7 Miles, Matthew, Huberman, Michael and Johnny Saldana, *Qualitative Data Analysis,* Thousand Oaks, CA: SAGE Publishing, 1994.

8 One interviewee reported on experiences from two projects of same governance model and discipline. They are treated as one project in the analysis.

9 Shirk, Jennifer, Heidi Ballard, Candie Wilderman, Tina Phillips, Andrea Wiggins, Rebecca Jordan, Ellen McCallie et al. “Public participation in scientific research: a framework for deliberate design.” *Ecology and Society* 17, no. 2 (2012).

10 Follett, Ria, and Vladimir Strezov. “An analysis of citizen science based research: usage and publication patterns.” PloS one 10, no. 11 (2015): e0143687.

11 Shirk, Jennifer, Heidi Ballard, Candie Wilderman, Tina Phillips, Andrea Wiggins, Rebecca Jordan, Ellen McCallie et al. “Public participation in scientific research: a framework for deliberate

design." *Ecology and Society* 17, no. 2 (2012).

12 Joint Research Centre (JRC) of the European Commission. "Citizen Science: data and service infrastructure meeting". January 26-27, 2016 <https://ec.europa.eu/jrc/en/event/workshop/citizen-science-workshop>.

13 Bryman, Alan. *Social Research Methods (4th Edition).* Oxford University Press, 2012.

14 For instance see Durham, Emma, Helen Baker, Matt Smith, Elizabeth Moore, and Vicky Morgan. "The BiodivERsA stakeholder engagement handbook." *BiodivERsA, Paris* (2014): 108.

15 In the case of the collegial project, participants were specifically called collaborators rather than volunteers.

16 See "What is currently being done by projects and rationale" section above.

17 See also Tiago, Patricia. "Social Context of Citizen Science Projects". In: Ceccaroni, Luigi and Jaume Piera (editors). *Analyzing the Role of Citizen Science in Modern Research.* IGI Global (2016): 168-191.

18 Groom, Quentin, Lauren Weatherdon, and Ilse R. Geijzendorffer. "Is citizen science an open science in the case of biodiversity observations?." *Journal of Applied Ecology* 54, no. 2 (2017): 612-617.

19 See also Wiggins, Andrea. "Free as in puppies: Compensating for ICT constraints in citizen science." *Proceedings of CSCW 2013.* ACM Press (2013): 1469-1480.

Wilson Center

One Woodrow Wilson Plaza
1300 Pennsylvania Avenue, N.W.
Washington, DC, USA 20004-3027
202-691-4000
www.wilsoncenter.org

COMMONS LAB

The Commons Lab of STIP seeks to mobilize public participation and innovation in science, technology and policy.

http://bit.ly/CommonsLab

The Commons Lab is supported by the Alfred P. Sloan Foundation.