Wilson Center

Authors
Anne Bowser
Peter Brenton
Rob Stevenson
Greg Newman
Sven Schade
Lucy Bastin
Alison Parker
Jessie Oliver

Citizen Science Association Data & Metadata Working Group:
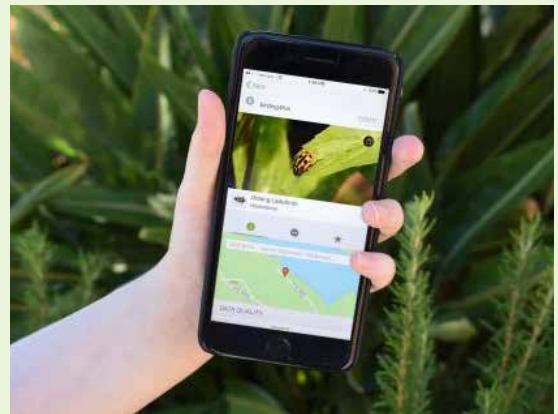
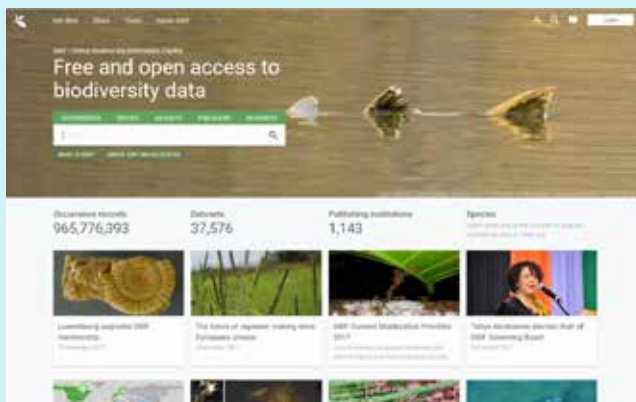## Report from CSA 2017 and Future Outlook

December 2017

The Importance of Data Standards

Data standards help local information, such as one biodiversity observation collected in one national park, to scale through aggregation and re-use.

*Standardized or interoperable data can be used by **local** communities to inform **local** management practices.*

*By sharing records with a data repository such as GBIF– support research on **global** problems like climate change.*

Authors

Anne Bowser- The Wilson Center

Peter Brenton- Atlas of Living Australia

Rob Stevenson- University of Massachusetts-Boston

Greg Newman- Colorado State University

Sven Schade- European Commission, Joint Research Centre

Lucy Bastin- European Commission, Joint Research Centre

Alison Parker- ORISE Fellow hosted at the U.S. Environmental Protection Agency

Jessie Oliver- Queensland University of Technology

Citizen Science Association Data & Metadata Working Group:
Report from CSA 2017 and Future Outlook

In 2016, the U.S. National Park Service (NPS) celebrated its centennial anniversary with over 100 Bioblitzes hosted in parks around the country. Because these events engaged local communities in documenting their natural environment, each park was given the freedom to decide how their celebration should unfold, for example by specifying which species to document or when to collect information. Still, all parks across the United States used the iNaturalist[1] mobile application for data collection. Because iNaturalist collects and stores biodiversity data in line with the Darwin Core standard, information collected in each local park can be verified by a community of experts and shared with GBIF,[2] a global database of biodiversity observations.

This case study illustrates how — through interoperable data standards — information considered important by local communities can "scale" to be used in national or global research and policymaking. *The goal of the Citizen Science Association (CSA) Data and Metadata Working Group (WG) is to make this vision a reality by promoting interoperability not just in one research domain like biodiversity, but across each and every research domain where citizen science is taking root and growing.*

Introduction

With the growth of citizen science comes the challenge of coordinating people, projects, and data. These challenges also present an opportunity. Through the use of data and metadata standards and other mechanisms to promote interoperability, data can support multiple research questions, allowing citizen science to help address ever-grander issues and problems on local, regional, national, and global scales.

In 2015, the U.S. Citizen Science Association (CSA)[3] founded a Data and Metadata WG to promote collaboration in citizen science through the development and/or improvement of international standards for citizen science data and metadata. The first formal meeting of the Data and Metadata WG was held at the CitSci2017 Conference, on May 17th, 2017, in Minneapolis, Minnesota. The goals of this meeting were to:

- Convene new and existing working group members.

- Refine the governance structure and leadership team.

- Review the history and mission of the working group.

- Establish an agenda for advancing the working group's mission through projects and task forces.

This consensus document reports on the outcomes of CitSci2017 and subsequent discussions and advances the shared agenda of working group members.

History and mission of the CSA Data and Metadata Working Group

During the fall of 2013, organizations supporting citizen science convened in Albuquerque, NM and Philadelphia, PA to discuss how to best share information about the growing number of citizen science, or public participation in scientific research (PPSR), projects. These early meetings were supported by the DataONE PPSR Working Group and SciStarter. Attendees represented databases of citizen science projects compiled by SciStarter, CitSci. org, the Cornell Lab of Ornithology, and the Wilson Center.

These organizations reached a landmark agreement - to share basic information across databases that catalog citizen science projects. The collaborative success for citizen science and PPSR led to the development of an initial PPSR-Core project metadata protocol (Appendix A). Sharing information across databases maximizes everyone's efforts, getting more "bang for the buck" out of information about citizen science projects everywhere. It leads to current, up-to-date, synchronized project lists and promotes a broader understanding of the global practice of citizen science.

In July 2015, these partners convened at the Wilson Center in Washington, DC along with members of the growing Federal Community of Practice for Crowdsourcing and Citizen Science (CCS). As one result of the meeting, consensus was reached around seven "core" or required fields to support data sharing between SciStarter, CitSci.org, the Wilson Center's Federal Catalog, and other databases. The July 2015 meeting also produced a list of optional and in-progress fields that still needed refinement.

Also in July 2015, members of CSA, the European Citizen Science Association (ECSA), and the Australian Citizen Science Association (ACSA) met to discuss citizen science data standardization and interoperability at the first ACSA conference in Canberra, Australia. Members of these associations agreed that any effort to develop and promote data and metadata standards for citizen science should be global in scope. Meeting attendees also recognized that while the initial PPSR-Core data sharing protocol was developed to share information about citizen science projects, a mature version of PPSR-Core should also help organizations share citizen science datasets and data.

Members of CSA, ECSA, and ACSA jointly proposed a CSA Data and Metadata WG with these goals in mind. In November 2015, CSA formally approved the Data and Metadata WG Charter. Also around this time a memorandum[4] issued by Dr. John Holdren, Director of the US White House Office of Science and Technology Policy (OSTP), instructed executive branch agencies to "develop the metadata requirements" for a federal catalog of crowdsourcing and citizen science projects. This memorandum helped engage the Federal Community of Practice for Crowdsourcing and Citizen Science (CCS) in developing project metadata for an expanded version of the catalog hosted by the Wilson Center.

> Sharing information across databases maximizes everyone's efforts, getting more "bang for the buck" out of information about citizen science projects everywhere. It leads to current, up-to-date, synchronized project lists and promotes a broader understanding of the global practice of citizen science.

● ● ● ● ● ●

In 2016, the European Commission Joint Research Centre (JRC) invited 20 international participants, including members of the three associations, to Ispra, Italy for a workshop to discuss data and service infrastructures for citizen science. Participants were asked to: (1) discuss the relationships between existing databases; (2) identify major requirements for interoperability between citizen science project databases, including a new database to be hosted by the European Commission; (3) draft a reference model for analyzing and sharing citizen science tools and data; and, (4) define a high-level roadmap with checkpoints for synchronizing ongoing activities related to standardization. This meeting led to an initial list of key projects for the working group.

As stated in our charter, the **purpose** of the Citizen Science Association Data and Metadata WG is to support, advance, and facilitate data interoperability among and between citizen science projects and other data repositories; and, to promote collaboration in citizen science via the development and/or improvement of international standards for data and metadata. The **goals** of the CSA Data & Metadata WG are to track and coordinate global efforts towards data and metadata interoperability and to ensure that all relevant stakeholders are included in these discussions.

Governance structure and leadership team

As with other CSA working groups, the CSA Data and Metadata WG is led by two co-chairs. Governance over the working group is shared between the co-chairs along with liaisons to partner Citizen Science Associations, including the European Citizen Science Association (ECSA) and the Australian Citizen Science Association (ACSA), and with liaisons to Standards Development Organizations (SDOs) and similar groups. Together, the co-chairs and liaisons comprise the CSA Data and Metadata WG leadership team.

The work of the CSA Data and Metadata WG is supported and implemented by the leadership team and members at large. The leadership team, in consultation with members at large, is responsible for identifying big-picture **projects**. The leadership team also pursues funding opportunities and coordinates across organizations to ensure that projects have sufficient support to succeed. Projects may be broken down into **task forces**, which have designated leaders, specific objectives, and a time frame for completion. Any member of the CSA Data and Metadata WG may propose or lead a task force.

The working group will hold an in-person meeting at each biennial CSA conference, quarterly virtual meetings, and regional meetings on an ad hoc basis. Meetings are intended to give updates on projects and task forces. Outside of meetings, virtual coordination and collaboration will be centralized through two primary platforms: (1) The Member 365 platform, accessible to all CSA members; and (2) Basecamp, accessible to all CSA working group members by invitation. The CSA website will serve as the public face of the Data and Metadata WG.

Citizen scientists learn water sampling techniques on the Mississippi River. Photo Credit: Wisconsin Department of Natural Resources, CC BY-ND.

Working group members

FOUNDING CO-CHAIRS

Name	Affiliation
Anne Bowser	CSA Board of Directors, The Wilson Center
Greg Newman	CSA Board of Directors, Colorado State University, CitSci.org

CITIZEN SCIENCE ASSOCIATION LIAISONS

1. The European Citizen Science Association (ECSA). CSA and ECSA have an active Memorandum of Understanding (MoU) to work together on web-based services and resources including project databases, initiatives, and guidelines. In addition, ECSA supports a Projects, Data, Tools, and Technology Working Group.[5] ECSA liaisons:

Name	Affiliation
Luigi Ceccaroni	ECSA Board of Directors, 1000001 Labs
Jaume Piera	ICM-CISC

2. The Australian Citizen Science Association (ACSA). CSA and ACSA have an active MoU to work together on web-based services and resources including project databases, initiatives, and guidelines. ACSA liaisons:

Name	Affiliation
Peter Brenton	Atlas of Living Australia
Jessie Oliver	ACSA Management Committee, Queensland University of Technology

STANDARDS DEVELOPMENT ORGANIZATION (AND SIMILAR) LIAISONS

1. Open Geospatial Consortium (OGC). OGC supports a Citizen Science Domain Working Group (DWG).[6] CSA and OGC have an active MoU to work together on developing citizen science data standards.

Name	Affiliation
Ingo Simonis	Open Geospatial Consortium

2. International Council for Science: Committee on Data for Science and Technology (CODATA) and World Data System (WDS). A joint CODATA-WDS Task Group (TG) addresses Citizen Science and the Validation, Curation, and Management of Crowdsourced Data.7 CODATA-WDS TG liaison:

Name	Affiliation
TBD	TBD

3. Biodiversity Information Standards (TDWG). TDWG supports a Citizen Science Interest Group.[8] TDWG liaison:

Name	Affiliation
Rob Stevenson	University of Massachusetts - Boston

4. World Wide Web Consortium (W3C). W3C is responsible for the Data Catalog vocabulary (DCAT) which is currently under revision by the Data eXchange Working Group (DXWG) W3C liaison:

Name	Affiliation
Simon Cox	CSIRO

Name	Affiliation
Lucy Bastin	European Commission, Joint Research Centre
Jonathan Brier	University of Maryland, SciStarter
Hillary Burgess	University of Washington
Caren Cooper	North Carolina State University
Pam DiBona	Massachusetts Bays National Estuary Program
Claudia Goebel	Museum für Naturkunde Berlin
Annette Olson	American Association for the Advancement of Science
Alison Parker	ORISE Fellow hosted at the U.S. Environmental Protection Agency
Hampapuram (Rama) Ramapriyan	Science Systems and Applications, Inc./ NASA Goddard Space Flight Center
Russell Scarpino	CitSci.org
Sven Schade	European Commission, Joint Research Centre
Justin Schell	University of Michigan
Carrie Seltzer	AAAS Science & Technology Policy Fellow at the National Science Foundation
Hannah Specht	University of Minnesota
Peter Tango	USGS at the Chesapeake Bay Program Office
Chrisa Tsinaraki	European Commission, Joint Research Centre
Julie Vastine	Alliance for Aquatic Resource Monitoring

Agenda for Advancing the Working Group's Mission

Project #1: Advancing work on PPSR-Core

PPSR-Core is a set of global, transdisciplinary data and metadata standards describing contextualized details about PPSR **projects** (Project Data Model, or PDM), **datasets** (Dataset Data Model, or DDM), and **data** (Observation Data Model, or ODM). These standards are united, supported, and underlined by a common framework, the PPSR-Core **common data model** (CDM), which illustrates how information is structured within the citizen science domain (figure 1).

Introducing PPSR CORE:
A Common Data Model (CDM)
with three main schemas

Overview

1 Project Data Model (PDM)

A coordinated cluster of citizen science and data. Information describes the context and purpose for activities.

Key components of this schema are used to share data between SciStarter, ALA, CitSci.org, and the Federal Catalog.

May Have 0:n

Will Have 1:n

3 Data Model (ODM)

Data records using a common method and data schema.

(eg. Physical/chemical water quality records over a series of data collection events at one or more sites)

2 Data Set Model (DDM)

Descriptive & contextual information about collection of data records using a common method and data schema.

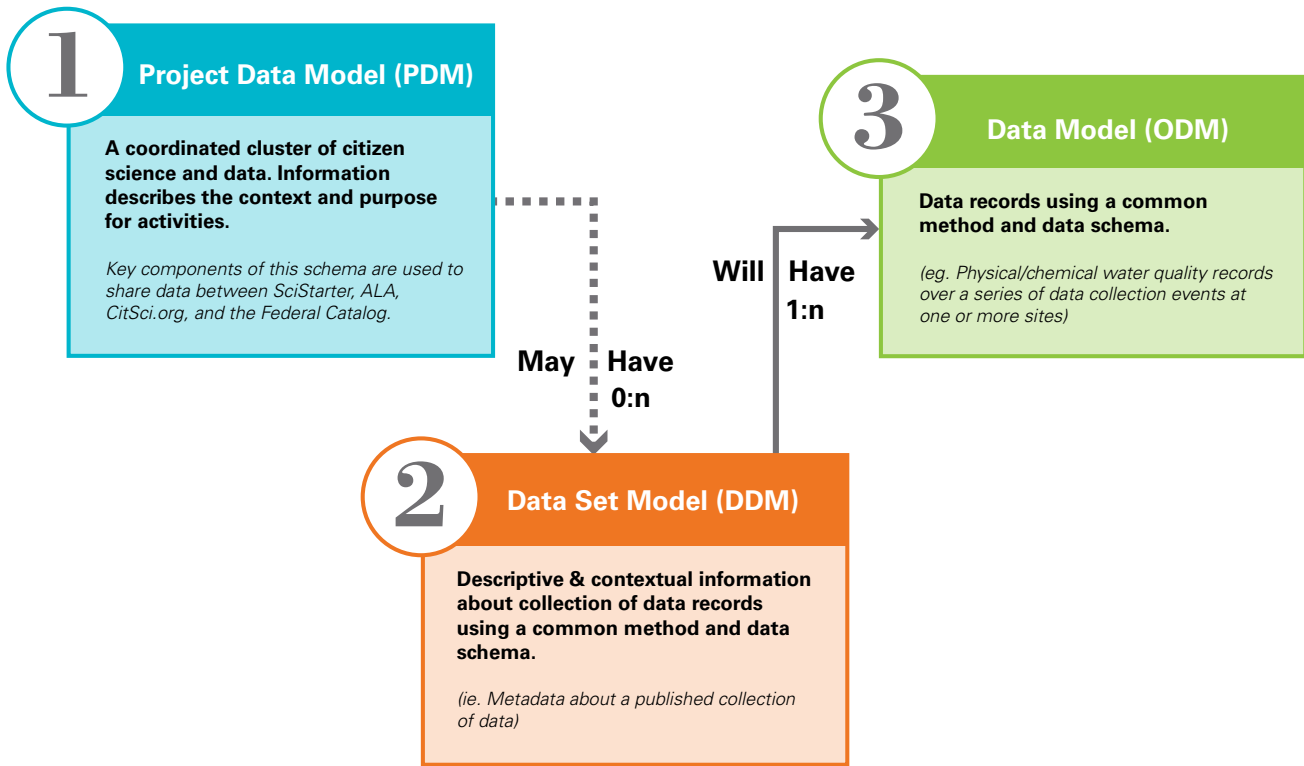*(ie. Metadata about a published collection of data)*

Figure 1: The PPSR-Core data model framework: A common data model with three main schemas.

Building off ongoing information exchanges between four databases of citizen science projects-- SciStarter, The Atlas of Living Australia, CitSci.org, and the Federal Catalog -- the PDM will facilitate cross-platform search and navigation to detailed project information. The DDM will enable datasets to be consistently described with metadata that help potential data users make informed decisions about fitness for purpose. This includes information about data quality, such as how the data were collected, the processes taken to ensure data quality, and the resulting data quality assessment. The ODM will define a core set of attributes which are common to most, if not all, citizen science projects, extendable with a range of common disciplinary or topical data collection protocols (e.g., for air and water quality monitoring, biodiversity, or genetic data).

Where appropriate, PPSR-Core will leverage and reference existing recognised standards for particular attributes, augmenting them with new PPSR-Core attributes to fulfill requirements that are unique to citizen science. The CSA Data and Metadata WG will work closely with standards development organizations (SDOs) including the Open Geospatial Consortium (OGC), the World Wide Web Consortium (W3C), and the Taxonomic Data Working Group (TDWG) throughout this process.

To date, some progress has been made on each of the three data models associated with PPSR-Core, particularly the Project Data Model and the Dataset Data Model (through the evolving OGC standard SWE4CS[9] and various existing implementations). However, the working group has reached consensus that agreement on the Common Data Model (figure 1) should be reached before work on the three supported data models continues. Reaching agreement on the Common Data Model requires

(1) understanding and articulating user requirements and use cases, and

(2) understanding the opportunities and limitations associated with existing citizen science data and supporting standards and ontologies.

In supporting the WG priority of advancing PPSR-Core, two task groups will advance each of these goals. A fourth task group will finalize a high-level Common Data Model to shape future development of the standard.

TASK FORCE 1.1: PPSR-CORE USER REQUIREMENTS ANALYSIS

Citizen science is a complex transdisciplinary and interdisciplinary approach to conducting scientific research. Activities bridge disciplines including biology; health; ecology & environmental science; astronomy, astrophysics & cosmology; agriculture & nutrition; energy; chemistry; mathematics; digital technologies; social sciences; economics; and, education, among others.[10] Citizen science also engages a range of stakeholders in civil society organizations; academic or research organizations; government agencies and departments; formal learning institutions; businesses and industry; and, public participants.[11] With this diversity in mind, any approaches to standards development and data integration should be complemented with a holistic and comprehensive view of what is actually needed.

Objectives of this task group include:

- Better understand community needs relating to standardization through a series of stakeholder engagement roundtables.

- Based on stakeholder needs analysis, develop a series of use cases to help test and validate the Common Data Model. These cases should be technically challenging. Some use cases may require combinations of attributes from components not adequately considered by independent data models – for example, a search for citizen science projects which use a certain sampling protocol, and are suitable for children. In this case, "Suitable for children" should be an attribute of the project data model (PDM) whereas sampling protocol is an attribute of the dataset data model (DDM). Critically, use cases should be grounded in and driven by actual user needs.

  Interested WG members: Sven Schade (lead), Anne Bowser (lead), Lucy Bastin, Peter Brenton, Jessie Oliver. Note: The objectives of this task force are shared by the Open Geospatial Consortium (OGC) Citizen Science Domain Working Group (DWG).

TASK FORCE 1.2: PPSR-CORE OPPORTUNITIES AND LIMITATIONS ANALYSIS

As described above, an initial PPSR-Core project data sharing protocol was developed to support the exchange of records on citizen science projects between platforms including SciStarter, CitSci.org, the Atlas of Living Australia (ALA), and the Federal Catalog. Information about the initial PPSR-Core project data sharing protocol may be found in Appendix A. However, the actual development of the PPSR-Core project data *model* is limited beyond a small number of core fields.

In addition to PPSR-Core, there are three other mapping models under discussion: One is based on the architecture of the Atlas of Living Australia's BioCollect system.[12] A second is based on the architectural model underpinning the CitSci.org system,[13] and the third is based on the outcomes of a European COST Action Project.[14] In addition to these mapping models, there are a range of existing standards and implementations relevant to the project data model (e.g., Dublin Core, PROVO-O, DCAT, ISO, VIVO, DBPedia ontology and others.); the dataset model (e.g., Darwin Core, Dublin Core, ISO, and others); and, the data model (e.g., Darwin Core, OGC, various domain specific standards, and various ISO standards). There is a need to analyze the opportunities presented by existing efforts as well as the limitations of each. One initial mapping exercise for the dataset model is presented in Appendix B.

Objectives of this task group include:

- Collect, document, and publish a list of the most important disciplinary and interdisciplinary standards for advancing PPSR-Core.

- Assess the strengths and limitations of each relevant standard for the citizen science context.

- Complete and publish an initial mapping of a handful of standards relevant to the PPSR-Core Project Data Model (see Box 1 for more details), Dataset Data Model, and Observational Data Model.

Interested WG members: Peter Brenton (lead), Luigi Ceccaroni (lead), Ingo Simonis, Lucy Bastin.

Box 1: Guiding principles for utilizing existing standards in PPSR-Core

The value of existing standards for PPSR-Core will be assessed using a "hierarchy of precedence" approach. All rules are applicable. In case of conflicting rules, the earlier rule will apply.

1. Universal, domain agnostic standards (e.g., Dublin Core) will take precedence over domain or industry-specific standards.

2. Curated or managed standards that are published by recognised publishing channels and/or standards development organizations will take precedence over "standards" which have not been published through such channels.

3. Original or "root" terms will take precedence over derivative or alternate implementations where term equivalents exist.

4. Widely used terms (e.g., common to multiple standards) will take precedence over existing bespoke or uncommon terms.

5. Where any of the above points are in contention, working group members *including liaisons to partner organizations* will make a majority decision on the appropriate term and namespace to use.

In the absence of an existing appropriate standard for a term, PPSR-Core will define a new standard term in the **"ppsr:" namespace**. This may or may not take into account existing bespoke or uncommon terms.

Sharing Project Information

The Atlas of Living Australia (ALA), CitSci.org, and Federal Catalogue of Crowdsourcing and Citizen Science all collect different types of information on citizen science projects. All of these projects share records with SciStarter as the "canonical" citizen science project database. Sharing records across databases is possible through the PPSR-Core Interim Project Data Sharing Protocol (Appendix A). New project databases are encouraged to use PPSR-Core standards to ensure interoperability and participate in this exchange.

TASK FORCE 1.3: PPSR-CORE COMMON DATA MODEL

Further developing PPSR-Core requires a common data model that clarifies the relationship between the Project Data Model (PDM), Dataset Model (DDM), and Observational Data Model (ODM). Development of the common data model will build on initial work with the PPSR-Core standard; experiences learned from existing systems; stakeholder engagement and use case development; and, the opportunities and limitations analysis.

Objectives of this task group include:

- Synthesize existing data models into a single data model that members of the CSA Data & Metadata WG agree is an accurate representation of the field and which adequately addresses the requirements of all use cases identified by Task Force 1.1.

- Share the common data model with broader members of the citizen science community for additional feedback and refinement.

Interested WG members: Peter Brenton (lead), Luigi Ceccaroni (lead), Sven Schade, Lucy Bastin.

Project #2: Triangulating and integrating citizen science data.

Citizen science can be used to fill geographic, temporal, and other data gaps in local, national, and global monitoring efforts and frameworks such as Essential Biodiversity Variables (EBVs).[15] Achieving this promise requires being able to find and assess citizen science data, combining it with other scientific data, as well as the technical capacity to integrate data from a range of sources. Triangulating and integrating citizen science data will help advance research while also demonstrating the value of data standards and interoperability.

TASK FORCE 2.1: INTEGRATING CITIZEN SCIENCE DATA

The CSA Data & Metadata WG will develop a number of proof of concept data integration and/or visualization demonstrations. These may combine data from multiple projects within a domain, multiple projects in different research domains, and/or combine citizen science data with different types of data, for example, information collected through remote sensing. The benefits and challenges of each demonstration will also be written up as case studies published to the CSA website.

Objectives of this task force include:

- Developing a template for data integration case studies that demonstrates the goal of each exercise, the relevant projects and stakeholders, the added value of data integration, the use of any relevant standards, technical challenges experienced, and other important information.

- Executing and documenting a series of data integration case studies including:

  - Global Mosquito Alert.[16] This demonstration will integrate different types of citizen science mosquito monitoring data with data from other sources, such as remote sensing. This project is supported by the NASA DEVELOP Program and will begin in Fall 2017.

  - Volunteer water quality monitoring. There are numerous organizations coordinating volunteer water quality monitoring efforts, but few data integrations have been broadly publicized within the CS community.

  - Data Refuge[17] climate data integration. This exercise will triangulate citizen science data related to climate and the environment from federally-sponsored projects. This project could leverage or link to data from projects in the Federal Catalog.[18]

  - CitSci.org and iNaturalist exchange. Observations submitted to CitSci.org need to be vetted for proper taxonomic identifications, and the iNaturalist community has a good way to obtain research grade identifications through crowdsourcing with expert naturalists. This integration will share observations made on CitSci.org with iNaturalist and, in return, get a response when the shared observation obtains research grade status on the iNaturalist platform, to enable CitSci.org to improve their record. This process also enables iNaturalist to in turn share the newly-minted CitSci.org research grade observation with GBIF.

  - CitSci.org and CoCoRaHS exchange.[19] Researchers are working to see if is feasible to share CoCoRaHS data with CitSci.org to enable projects to correlate precipitation data with other observations made through the CitSci.org platform.

  - Invasive Alien Species (IAS) integration. A test case that integrates observation reports from three different apps (Invasive Alien Species in Europe,[20] IASTracker,[21] and Natusfera[22]). This test case also provides a harmonized data validation support tool and feeds the validated data into EASIN,[23] a scientific database that is used for science and policymaking.

- Protected Areas state of knowledge integration. Within the framework of JRC's Digital Observatory for Protected Areas, it is proposed to collate and mobilize contributed data on the state, pressures, and management of protected areas. This demonstration will evaluate the feasibility of a linked data annotation approach to aggregating heterogeneous resources for selected sites, relating to poaching, species trends, designation and management, recreation and socioeconomic concerns. Key aspects of best practice will be identified in order to support the publication and mobilization of additional resources from citizen science and Volunteered Geographic Information (VGI).

Interested WG Members: Anne Bowser (Global Mosquito Alert), Justin Shell (Data Refuge), Greg Newman (CitSci.org test cases), Sven Schade (IAS test case), Lucy Bastin (Protected Areas).

TASK FORCE 2.2: SCOPING, BUILDING, AND POPULATING A "DATA BAZAAR"

In order to take full advantage of PPSR-Core and related efforts, the citizen science community need supporting cyberinfrastructure, dubbed "Data Bazaar", to help users find, access, analyze, and share diverse types of citizen science data.

Objectives of this ambitious and long-term task force include:

- Scoping, building, and populating an inventory and search engine of citizen science data. This could be a common platform or, initially, a distributed catalog that links out to other data sources. Each data set in this inventory should be documented in a manner consistent with the PPSR-Core dataset data model.

- Scoping, building and using solutions for harmonized data access from citizen science projects based on the PPSR-Core dataset data model and observational data model. These solutions would allow data not only to be discovered but also accessed in a standardized way and in a standardized format.

- Scoping, building and using a platform for data integration workflows, data modeling, data visualization, and other important research processes. This might provide the methods, tools and best practices for data integration, but also services that can directly be used for integrating data from multiple sources. This platform will use the PPSR-Core data standards. It may be based on or integrate with existing platforms such as CitSci.org for data storage and management, and RawGRAPHS[24] for data visualization.

- Building on test cases and existing or new cyberinfrastructure, develop a growing number of integrated, analyzed, and/or visualized data sets that use citizen science to address wicked global problems.

Interested WG Members: Anne Bowser, Sven Schade, Jessie Oliver, Lucy Bastin, Peter Brenton.

Project #3: Developing a communications strategy, including a suite of tools and messages, around data and metadata in citizen science.

Science communication is a challenge for many researchers, including those utilizing citizen science methodologies. Communicating the wide range of concepts and practices associated with citizen science data and metadata may be especially challenging; however, there is a pressing need to develop language to help researchers in and beyond the citizen science community understand key concepts and ideas.

TASK FORCE 3.1: DATA AND METADATA ELEVATOR PITCHES

Both within and outside of the citizen science community, many people question the benefits of data standardization and interoperability[25] or challenge the quality of citizen science data.[26] Members of the Data & Metadata Working Group will develop a series of "elevator pitches" to help the citizen science community articulate key concepts about citizen science data. Relevant audiences include:

(1) Citizen science researchers and practitioners from a wide range of projects;

(2) Scientists not yet using citizen science;

(3) Funders;

(4) Policymakers; and,

(5) Volunteers.

Objectives of this task force include:

- Developing language around the value and importance of citizen science data and metadata standards and interoperability. This should cover basic information on what data and metadata are. It should also address what PPSR-Core is, why it is helpful, and how it is being and should be used.

- Developing language around issues related to data quality in citizen science. There is a need for the community to be able to succinctly establish the high quality of many citizen science data. There is also a need to elaborate on the concept of "fitness for use" or "fitness for purpose" based on existing examples from numerous application areas (science, policy) and research domains.

Interested WG Members: Julie Vastine (lead), Lucy Bastin, Alison Parker.

TASK GROUP 3.2: THE STATE OF THE DATA IN CITIZEN SCIENCE

Members of the CSA Data & Metadata WG will publish a series of journal articles, columns, white papers, or blog posts on the state of the data in citizen science. Goals of these publications include educating the citizen science community on data issues and evaluating/ tracking the state of CS data in regard to metrics like data openness, data accessibility, data documentation, data interoperability, and data quality.

Objectives of this task force include:

- Publish a number of peer-reviewed research publications or white papers on the state of the data in citizen science.

- Investigate the opportunity for an ongoing publication series such as a column in *Citizen Science: Theory and Practice* or a series of regular posts on the CSA Blog.

Interested WG Members: Andrea Wiggins (lead), Anne Bowser, Caren Cooper, Rob Stevenson, Hillary Burgess, Sven Schade, Peter Brenton, Jessie Oliver. Note: This is a shared objective with the WDS-CODATA Citizen Science TG.

TASK GROUP 3.3: CHOOSE YOUR OWN DATA QUALITY ADVENTURE

There are a number of established tools and processes for supporting data quality in citizen science.[27] Yet, there is no obvious way for citizen science practitioners to understand which best practices may suit their needs. A decision support tool for understanding different data collection concerns and solutions based on different user inputs would help support good practices around data quality in citizen science.

Objectives of this task force include:

- Understand and develop a common vocabulary for discussing the range of data quality practices in citizen science.

- Match different data quality practices to different types of citizen science projects and activities.

- Build a decision support tool to make this information easy for users to access.

Interested WG Members: Pam DiBona (lead), Hillary Burges (lead), Russanne Low, Lucy Bastin, Hampapuram Ramapriyan.

Conclusion

The recent proliferation of citizen science exposes a need for coordination between people, projects, and data. The Citizen Science Association's Data & Metadata Working Group seeks to advance and facilitate data interoperability between citizen science projects, and other data repositories. The Working Group also should promote collaboration in citizen science via the development and/or improvement of international standards for data and metadata. Such coordination is important for helping citizens science projects maximize their impact on research and decision making, and for allowing information collected by local communities to "scale" in order to have national and even global impact. Through our work, we hope to help citizen science reach its full potential as a paradigm that both advances and democratizes knowledge production in all areas of scientific research.

Citizen Science Association
Data and Metadata
Working Group:
Report from CSA 2017 and
Future Outlook

Authors
Anne Bowser
Peter Brenton
Rob Stevenson
Greg Newman
Sven Schade
Lucy Bastin
Alison Parker
Jessie Oliver

Appendix

APPENDIX A:
INTERIM PROJECT METADATA SHARING PROTOCOL

APPENDIX B:
PPSR-Core -- DATASET DATA MODEL (v.0.2.0 Sept 11, 2017)

December 2017

APPENDIX A: INTERIM PROJECT METADATA SHARING PROTOCOL

The PPSR-Core Interim Project Data Sharing Protocol (v0.2.0 Sept 11, 2017) is a precursor to the PPSR-Core Project Data Model which extracts **only** the fields required for the sharing project information between different catalogs. The fields described in this model are already implemented by four citizen science project catalogs hosted by SciStarter, the Federal Catalog, the Atlas of Living Australia, and CitSci.org.

AIMS:

- To document the who/what/when/where/how about projects so that a range of stakeholders can understand the objectives, reasons, spatial, temporal, social, and other contexts for citizen science activities;

- To provide information about projects to the public in order to help potential volunteers make informed choices about projects to contribute to;

- To facilitate effective sharing of information between project catalogs for more comprehensive and personally relevant searchability of projects; and,

- To collect consistent information about projects worldwide and facilitate more effective research on the paradigm of citizen science than has previously been possible.

NOTES:

- The following metadata schema is only a subset of the overall PDM and is provided as a minimum set of PDM attributes required to meet all of the current use cases for data sharing of project information between project catalog repositories.

- This metadata schema describes the objectives, participants, scope and operational aspects of projects. Projects will typically involve undertaking a range of activities, some of which may involve the creation of one or more datasets.

- While there is some consensus amongst Task Group 1.2 and 1.3 members on these fields in their current form, it should be expected that some fields may change as full consensus evolves. The W3C Data eXchange project is also currently working on a standard specification for a PDM as part of a revision of DCAT standards which, as it evolves, may also influence this PDM schema.

- This schema specification indicates several fields as "required." As a general rule, all "required' fields are considered mandatory for databases of citizen science projects seeking to share information. These fields are needed in order to:

  a. Ensure that a minimum set of key information is consistently recorded for every project; and,

  b. Ensure that every project can be found in project catalogs through the most common discovery approaches and requirements.

- Fields marked as "not required" do not need to be included in databases to achieve the above goals, but add value to the information provided and are recommended for consideration as inclusions in information sharing agreements.

RELEVANT INFORMATION IN THE PPSR-CORE DATA SHARING PROTOCOL INCLUDES:

- PPSR-Core field name. A short, common language descriptive term for the field. This may be used as the field label displayed in user interfaces or it may be changed to better represent the needs of the audience of a database. If the field name is changed, the field description should be kept verbatim and acknowledged to ensure that the field is correctly and consistently applied.

- PPSR-Core database term. The name used for a field in a database. It is important that the database term is applied exactly in the format shown here (ie. camel case with no spaces, underscores, etc.), including the namespace prefix and ":". The "namespace" refers to the abbreviation for the applicable standard ontology in front of the term itself. This standardized database term allows project catalogs to exchange information automatically (e.g., via APIs).

- Required. Indicates whether a field/term is mandatory or optional in the database. Fields with required = Y are essential, and are required for different project catalogs to effectively share project and dataset information. These fields must be included in a compliant database and must be used to record data in that database (e.g., they cannot be null or empty fields). Fields with required = N are not essential, but including them will allow for more information about the project to be recorded and will help people understand the project.

- Data type. A data type is a special class/category which determines what kind of information can be entered or recorded into a particular data field. For example, some fields can only accept numeric values, not text, or images, or any other kind of information. Every database element must have a specified data type.

- Multiplicity. Multiplicity indicates whether a particular field is allowed to have zero, only one, or many usages for a given record. For example, "Aim" is mandatory and must have one usage for every record; a zero value is not permitted. But, because "Aim" seeks to document the primary aim or goal of a project, it does not make sense for this field to have multiple usages in a single record. However, citizen science projects must relate to at least one area of science and could potentially address many areas at the same time, so a multiplicity of 1:many is appropriate for the "scienceType" field.

- Related Standard / Mapping. Many of the properties (fields) used in PPSR-Core are applicable beyond citizen science, and may have already been specified as standards in other domains. Where these do apply, the existing standard will specify the namespace which holds the ontology. Note: Some existing standards pertain to the definition of a property and others describe the format or structure of that property. Therefore, in some cases such as dateTime, two standards may apply to the same property- one to define the property, and one to define the format or structure. Multiple related standards for defining a property or format will not be offered (see Box 1, Guiding principles for utilizing existing standards in PPSR-Core).

Namespace references. Relevant standards include:

- @prefix dcterms: <http://purl.org/dc/terms/> .

- @prefix foaf: <http://xmlns.com/foaf/0.1/> .

- @prefix org: <http://www.w3.org/ns/org#> .

- @prefix proj: <http://www.w3.org/ns/project#> .

- @prefix prov: <http://www.w3.org/ns/prov#> .

- @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.

- @prefix sosa: <http://www.w3.org/ns/sosa/>.

- Description/comments. The description is the definition of the field, term, or property. This is important because in order for fields to be consistently interpreted and used in all situations the meaning of the field must be clearly understood by everyone who uses it, and must be consistently applied.

PPSR-Core field name	PPSR-Core database term	Required	Data type	Multiplicity	Related standard/ term mapping	Description/ comments
Database information						
GUID	projectId	Y	text	1		Globally unique identifier (GUID) for the project; system generated.
External Id	projectExternalId	N	text	0:1		The identifier of the project in an external database or repository.
Origin	project OriginalRepository	Y	text	1		The name of the project database where a project was first registered. Allows traceability of a project in multiple databases to its original registration.
Date Created	projectDateCreated	Y	dateTime	1	dcterms:created <u>ISO 8601</u>:2004 (E)	The date and time that the record was created in the database.
Date Updated	project Last UpdatedDate	Y	dateTime	1	dcterms:modified <u>ISO 8601</u>:2004 (E)	The date and time that the record was last updated in the database.

Basic Project Information						
Name	projectName	Y	text	1	proj:name	Short name or title of the project.
Aim	projectAim	Y	text	1	proj:objective	Primary aim, goal, or objective.
Description	projectDescription	Y	text	1	rdfs:comment	Abstract or description of the project.
Tags	projectTags	N	vocabulary	0:many		Controlled vocabulary terms, supplied by the person who entered the project, to assist with search and filtering.
Keywords	dcatKeyword	N	text	0:many	dcat:keyword	Keywords (comma separated) which are indexed and aid in searching for and finding projects. Data Catalogue Vocabulary (DCAT)
Status	projectStatus	Y	vocabulary	1		The activity status of the project.
State date	projectStartDate	Y	dateTime	1	prov:startedAtTime ISO 8601:2004 (E)	The actual date that a project began.
End date	projectEndDate	N	dateTime	0:1	prov:endedAtTime ISO 8601:2004(E)	The actual date that a project ended.
Project topic	projectScienceType	N	vocabulary	0:many	proj: hasFieldOfResearch	The project topic or field of science.
Intended outcomes	project IntendedOutcomes	N	vocabulary	0:many	proj:objective	A project's goals, or intended outcomes of participation.
Images and communications						
Image	projectImage	N	image	0:1	foaf:img	An image to represent a project
Image credit	projectImage Credit	N	text	0:1		A credit for the image used to represent a project.
URL	projectUrl	N	http uri	0:1		URL to an external web site for the project.

Geography						
Project Latitude	projectPinLatitude	Y	floating point	1	GeoAPI (OGC 09-083r3)	Latitude coordinate of the center of the project area. Typically this is where the project is hosted, e.g., a home institution.
Project Longitude	projectPin Longitude	Y	floating point	1	GeoAPI (OGC 09-083r3)	Longitude coordinate of the center of the project area. Typically this is where the project is hosted, e.g., a home institution.
Geographic extent	project GeographicCoverage	N	GeoJSON, WKT, SHP	1	proj: hasAreaOfInterest GeoAPI (OGC 09-083r3)	User-defined geospatial representation of the project area footprint. Coverage is typically represented in a GeoJSON object which has a centroid coordinate ("centre") and a definition of the boundary of the shape.
	projectGeographic CoverageCentroid Latitude	N	floating point	1	GeoAPI (OGC 09-083r3)	Latitude coordinate of the centroid of the project extent area. Latitude coordinate in geographic decimal degrees for the center or home base of the project best representing the project's location as a point.
	projectGeographic CoverageCentroid Longitude	N	floating point	1	GeoAPI (OGC 09-083r3)	Longitude coordinate of the centroid of the project extent area. Longitude coordinate in geographic decimal degrees for the center or home base of the project best representing the project's location as a point.

UN regions	unRegions	Y	vocabulary	1:many	Modified M49 coding classification	Select list of United Nations regions. Used for enabling regionalized views of projects in order to make search-ability and applicability of projects more regionally relevant for users.
Project location	projectLocality	N	text	0:1		Free text location specification.
Personal and Organizational Affiliates						
Project host	projectResponsible PartyName	Y	text	1	proj:hasLeader foaf:name prov:entity	Name of the primary organization responsible for hosting or implementing the project.
Project host contact	projectResponsible PartyContact	N	text	0:1	proj:hasLeader foaf:{various}	The contact details for a party responsible for the project. This may include postal and street address, telephone and email details. Note that this is not parsed into individual properties.
Project host contact email	projectResponsible PartyEmail	N	text	1		Email address for the person responsible for keeping the record up to date.
Project contact	projectContact Name	N	text	1	proj:hasLeader Foaf:name prov:entity	The name of the person responsible for keeping the project up to date. Traditionally the primary project coordinator.
Project contact email	project ContactEmail	N	text	1	proj:hasLeader foaf:mbox	Email address for the person responsible for keeping the record up to date.
Associated organisation	projectAssociated PartyName	N	text	0:many	proj:activity Participation Association prov:agent rdfs:label	Short text name or title of a party (eg. Organisation) associated with the project and performing a role in the project.

Supplementary Information for Citizen Science							
How to participate	project HowToParticipate	N	text	0:1			Free text description of how people can get involved in the project. Textual instructions for joining the project.
*Participation tasks	projectTasks	N	Vocabulary	0:1			The full list of tasks that volunteers may do to contribute to a project.
Difficulty	difficultyLevel	N	vocabulary	0:1			
Participant Age	projectParticipantAge	N	vocabulary	0:1			The intended age groups or populations recruited for participation.
Equipment	projectEquipment	N	text	0:1	prov:used (sosa:Sensor, sosa: Platform)		Required or suggested equipment
Country specific terms							
US Federal sponsor	project UsFederalSponsor	N	vocabulary	0:many			Name of U.S. federal agency offering a project direct sponsorship, funding, or other support.
Funding Programme Description	programDescription	N	text	0:1			General description/ summary of the Research Program or funding initiative.
Funding Programme ID	programId	N	text	0:1			The identifier of the initiative (e.g., overarching research program or funding initiative) encompassing the project
Funding Programme Name	programName	N	text	0:1			The initiative (e.g., overarching research program) encompassing the project (e.g., Horizon 2020)

Standard Vocabularies

The following sets of vocabulary terms are intended for use with the corresponding database term in the above table.

Box 2. General guidelines for use of vocabularies

- Terms should be mutually exclusive and unambiguous.

- Terms should be used exactly as provided to ensure compatibility between repositories sharing this information.

- Implementations may use a subset of terms from the vocabulary list, but should not add terms which are not in the published standard list as this would break compatibility with other systems.

If a particular implementation needs to add new terms to the standard published vocabulary lists, it should be done in consultation via the 'authority' responsible for managing the PPSR-Core standard-- eg., the Citizen Science Association Data & Metadata WG-- and in consultation with all existing implementations of controlled vocabularies.

Database Term	Vocabulary List of Terms	Comments
projectTags	Fees applicable Suitable for children Teaching materials available Do-it-yourself Participate from home Etc.	These are sample project terms. The SciStarter database alone includes more than 100 tags, which are not listed for reasons of brevity.
difficultyLevel	Easy Medium Hard	
projectStatus	Not yet started Active Periodically active On hold Completed Abandoned	Similar terms mapping a • Pending = Not yet started • Active but seasonal = Periodically active • Hiatus = On hold • Complete = Completed • Terminated = Abandoned

projectScienceType	Agriculture & Veterinary science	This current list represents an aggregation of terms from existing implementations and does not yet comply with the guidelines for mutual exclusivity and non-ambiguity.
	Animals	
	Archaeology & Cultural	
	Astronomy & Space	
	Biodiversity	
	Biogeography	
	Biology	
	Birds	
	Chemical sciences	
	Climate & Weather	
	Ecology & Environment	
	Education	
	Food science	
	Genetics	
	Geography	It is also known to be incomplete in terms of full representation/ coverage of areas of science and is therefore likely to change over time. Implementations should be mindful of this and should engage actively with the PPSR-Core community to monitor changes and contribute to the discussion.
	Geology & Earth science	
	Health & Medicine	
	Indigenous culture	
	Information & Computing sciences	
	Insects & pollinators	
	Long-term species monitoring	
	Ocean, Water, Marine & Terrestrial	
	Nature & outdoors	
	Natural resource management	
	Physics	
	Psychology	
	Science policy	
	Social sciences	
	Sound	
	Transportation	

projectIntendedOutcomes	Civic and community benefit Conservation Individual learning Programmatic Research advancement	In some cases, sub-terms are also used. This vocabulary is being developed by the DEVISE Project.
projectParticipantAge	Elementary school (6-10) Middle school (11-13) Youth/teen (up to 17) Families General public Targeted group	Consistent with CAISE controlled vocabularies.
projectTasks	Annotation Audio or video recording Classification or tagging DIY hacking/making Data analysis Data entry Download software for distributed computing projects Finding entities Geolocation Identification Learning Measurement Observation Photography Problem solving Sample analysis Site selection and/or description Specimen/sample collection Transcription	

unRegions	Africa Americas – South America Americas – Central America Americas – Northern America Asia Europe Oceania	
projectUsFederalSponsor	Department of Interior Department of Commerce Department of Health and Human Services Central Intelligence Agency Department of Defense Department of Education Department of Energy Department of Homeland Security Department of Housing and Urban Development Department of Justice Department of State Department of Transportation Department of Veterans Affairs Department of Labor Federal Communications Commission General Services Administration Intelligence Advanced Research Projects Agency National Aeronautics and Space Administration National Archives and Records Administration National Endowment for the Arts National Science Foundation White House Office of Personnel Management Peace Corps Smithsonian Institution Environmental Protection Agency US Department of Agriculture	Many terms have sub-terms (e.g., child agencies).

APPENDIX B: PPSR-CORE -- DATASET DATA MODEL (V.0.2.0 SEPT 11, 2017)

The PPSR-Core Interim Dataset Data Model (v0.2.0 Sept 11, 2017) is still under discussion as a precursor to the PPSR-Core Dataset Data Model which extracts only the fields required for the sharing project information between different metadata data repositories. The fields described in this model have been mapped to a series of common data set metadata schemes, such as the The DCAT Application Profile for data portals (DCAT-AP). Currently, possible extensions and profiles are under discussion in order to capture additional information as, for example, included in the metadata schemas of ISO, the OGC or DataCite.

FINAL AIMS:

- To consistently and comprehensively describe sets of data records which use the same data collection method and belong to the same project so that users of the data can understand its origin, spatial and temporal scope, biases, usage limitations, etc.

- To enable downstream data consumers to make informed decisions about fitness for intended purpose or use, particularly in respect to scientific analysis, aggregation with other equivalent datasets (eg. big-data analytics), and use in decision-making.

NOTES:

- The following metadata schema is only a subset of the overall DDM and is provided as a minimum set of PDM attributes required to meet the current use cases for data sharing of dataset information between data repositories.

- While there is some consensus amongst Task Group 1.2 and 1.3 members on these fields in their current form, it should be expected that some fields may change as full consensus evolves.

- Researchers conducted an in-depth mapping exercise of this model to the DCAT-AP and DataCite based metadata schema that is used by the European Commission's Joint Research Centre (JRC) in order to analyze completeness with respect to this particular model and possible shortcomings.

- The table below includes only those elements for which we could find an agreement.

- The following items still require further discussion, which may lead to a revision and extension of the current model:

  a. Possible distinction between a dataset and its distributions (e.g. as different sub-sets provided via separate access services). The results of this discussion will also affect the presence/cardinalities of some of the core elements, such as dataset external URL, download URL, encoding language, etc.

  b. Representation of the essentially required information about responsible parties as humans, organizations, or possibly only via e-mail addresses.

    c. Distinction between access and use conditions, as well as, licenses.

    d. Options to include data quality information at different levels of detail (possibly in form of profiles, see also below).

    e. Representation of contributors to a dataset (possibly allowing for individual persons or organizations).

    f. Possibilities to refer to a related publication, e.g. in a scientific journal.

    g. Addition of keywords as free-text or controlled vocabularies.

    h. Inclusion of status information for the dataset.

- For dataset metadata, there are several elements that are most commonly catalogued within institutions. However, there are several purpose dependent standards that vary between different institutions, such as the encoding of quality information and the inclusion of additional reference material. With the core model, we aim to represent the common elements. Furthermore, we consider the additional provision of extensions (profiles) that allow for a common modeling of more specific requirements. Ultimately, this approach should enable the integration of different implementations, each of which represents overlapping parts of the overall model.

Suggested PPSR-Core Field Name	PPSR-Core Database Term	Required	Data type	Multiplicity	Applicable standard(s) / term mapping	Property description/ comments
GUID	dcterms:identifier	Y	text	1	Dublin Core terms, Darwin Core terms - (dwc:datasetId)	Persistent identifier of a dataset (associated to the project). Should equate to the datasetExternalId if data is stored in an external repository.
Dataset name	dcterms:title	Y	text	1	Dublin Core terms. Equivalent to - dwc:datasetName and ISO 19115 (CI_Citation.title)	The name of the dataset for citation purposes
Description	dcterms:abstract	Y	text	1	Dublin Core terms	Abstract or description of the dataset.
URL	datasetExternalUrl	N	http uri	0 : many		Web location where the project data will be published

Download URL	dcat:downloadURL	N	http uri	0 : 1	Data Catalog Vocabulary (DCAT)	A URL from which dataset observation records can be downloaded
Access rights	dcterms: accessRights	Y	vocabulary	1	Dublin Core terms	Category of rights to use IP contained in the dataset or a type of use applied to the dataset
License	dcterms:license	Y	vocabulary	1	Dublin Core terms	License applied to the dataset
Encoding language	dataset EncodingLanguage	N	vocabulary	0 : 1	ISO 19115 (MD_DataIdentification.characterSet)	The machine language the dataset and associated metadata is encoded in
Geographic coverage	dataset GeographicCoverage	Y	GeoJSON, WKT,SHP	1	GeoAPI (OGC 09-083r3)	Geographic/spatial scope of coverage of the collection sites of data records within the dataset.
Status	datasetStatus	Y	vocabulary	1		Indicator of the current status of a dataset, e.g. if it is already published
Date submitted	dcterms: dateSubmitted	N	dateTime	0 : many	Dublin Core terms, ISO 8601:2004(E), ISO 19115 (CI_Citation.date)	The date a dataset submission was published into a receiving system
Date modified	dcterms:modified	N	dateTime	0 : 1	Dublin Core terms, ISO 8601:2004(E)	The most recent date-time on which the resource was changed.
Update frequency	dataset UpdateFrequency	N	vocabulary	0 : 1		How often the project information or dataset is updated

Standard Vocabularies

The following sets of vocabulary terms are intended for use with the corresponding database term in the above table.

Database Term	Vocabulary List of Terms	Comments
datasetUpdateFrequency	Daily Weekly Monthly Annually Occasionally Continuous	Taken from DC Terms (http://dublincore.org/groups/collections/frequency/)
dcterms:license	Creative Commons zero (CC 0) Creative Commons Attribution (4.0) International (CC-BY 4.0) Creative Commons Attribution Non-commercial (CC-BY-NC) Creative Commons Attribution Share Alike (CC-BY-SA) Creative Commons Attribution Non-commercial Share Alike (CC-BY-NC-SA)	This list is possibly incomplete and should allow for addition of terms as required. As a general rule though people are encouraged to apply non-restrictive licenses to enable effective and manageable data sharing and aggregation – CC 0 or CC-BY 4.0 are preferred. More restrictive licenses such as SA, NC and ND become very difficult to manage in a distributed information sharing context.
dcterms:accessRights	Open access Embargoed access Restricted access Metadata only access	See also http://vocabularies.coar-repositories.org/documentation/access_rights/

Endnotes

1   iNaturalist. "iNaturalist." https://www.inaturalist.org/ (accessed November 7, 2017)

2   GBIF | Global Biodiveristy Information Facility. "Free and open access to biodiversity data." https://www.gbif.org/ (accessed November 7, 2017)

3   Citizen Science.org. "Association." http://citizenscience.org/association/ (accessed November 7, 2017)

4   "Addressing Societal and Scientific Challenges through Citizen Science and Crowdsourcing." Executive Memorandum: Office of Science and Technology Policy. (September 30, 2015). https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/holdren_citizen_science_memo_092915_0.pdf

5   Europeans Citizen Science Association. "Projects, Data, Tools, and Technology." https://ecsa.citizen- science.net/taxonomy/term/206 (accessed November 7, 2017)

6   Open Geospatial Consortium. "Citizen Science DWG." http://www.opengeospatial.org/projects/groups/citizenscience (accessed November 7, 2017)

7   International Council for Science: Committee on Data for Science and Technology. "Citizen Science and the Validation, Curation, and Management of Crowdsourced Data (with WDS)." http://www.codata.org/task-groups/citizen-science-and-crowdsourced-data (accessed November 7, 2017)

8   TDWG. "Citizen Science." https://github.com/tdwg/citizen-science (accessed November 7, 2017)

9   Ingo Simonis. "Citizen Science Data Models" Github. https://github.com/opengeospatial/swe4citizenscience/blob/master/ER/clause-7-models.adoc along with note that it was (accessed on November 7, 2017)

10 Francois Houllier and Jean-Baptiste Merilhou-Goudard, Rapporteur. (2016). Citizen Science in France. Paris, France: INRA. Available Online at http://www.sciences-participatives.com/en/Report

11 Claudia Goebel, Victoria Y. Martin, and Mǒnica Ramìrez-Andreotta. 2017. International Citizen Science Stakeholder Analysis on Data Interoperability. Washington, DC: Woodrow Wilson International Center for Scholars

12 Atlas of Living Australia. "BioCollect." https://www.ala.org.au/biocollect/ (accessed November 7, 2017)

13 CitSci.org. "CitSci.org." http://www.citsci.org/ (accessed November 7, 2017)

14 Cost: European Cooperation in Science & Technology. "Citizen Science to promote creativity, scientific literacy, and innovation throughout Europe." http://www.cost.eu/COST_Actions/ca/CA15212 (accessed November 7, 2017)

15 Mark Chandler et al., 2017. Contribution of citizen science towards international biodiversity monitoring. Biological Conservation, 213, Part B, 280-294

16 The Wilson Center: Science and Technology innovation Program. "Global Mosquito Alert." https://www.wilsoncenter.org/article/global-mosquito-alert (accessed November 7, 2017)

17 DataRefuge. "Building refuge for federal climate & environmental data." https://www.datarefuge.org/ (accessed November 7, 2017)

18 Catalog: citizenscience.gov "Federal Crowdsourcing and Citizen Science Catalog" https://ccsinventory.wilsoncenter.org (accessed November 7, 2017)

19 Community Collaborative Rain, Hail & Snow Network. "Welcome to CoCoRaHS!" https://www.cocorahs.org/ (accessed November 7, 2017)

20 European Union. "Invasive Alien Species Europe" Google Play Store, Vers 1.0.5 (2017) https://play.google.com/store/apps/details?id=eu.europa.publications.mygeossias&hl=en (accessed November 7, 2017)

21 IC5Team. "IASTracker." Google Play Store, Vers. 0.1 (2016). https://play.google.com/store/apps/details?id=com.ic5team.iastracker (accessed November 7, 2017)

22 Natusfera "Natusfera." http://natusfera.gbif.es/ (accessed November 7, 2017)

23 Joint Research Centre. "European Alien Species Information Network – EASIN." http://easin.jrc.ec.europa.eu/ accessed November 7, 2017)

24 RAWGraphs. "RAWGraphs." http://rawgraphs.io/ (accessed November 7, 2017)

25 Claudia Goebel, Vicki Martin, and Monica Ramirez-Andreotta. 2017. "Stakeholder Analysis: International Citizen Science Stakeholder Analysis on Data Interoperability." Final Report. Woodrow Wilson International Center for Scholars. https://doi.org/10.13140/RG.2.2.26124.92802

26 Anon. 2015. "Rise of the citizen scientist." Nature, 524, 262, p. 7565

27 See, for example, Margaret Kosmala, Andrea Wiggins, Alexandra Swanson, and Brooke Simmons. 2016. Assessing data quality in citizen science. Front Ecol Environ 14 (10): 551-560; Andrea Wiggins, Greg Newman, Robert D. Stevenson, and Kevin Crowston. 2011. Mechanisms for Data Quality and Validation in Citizen Science. In Proceedings of the 2011 IEEE Seventh International Conference on e-Science Workshops (ESCIENCEW '11). IEEE Computer Society, Washington, DC, USA, 14-19

The Wilson Center

@TheWilsonCenter

facebook.com/WoodrowWilsonCenter

www.wilsoncenter.org

Woodrow Wilson International Center for Scholars
One Woodrow Wilson Plaza
1300 Pennsylvania Avenue NW
Washington, DC 20004-3027

Wilson Center