Chapter 9
A Conceptual Model for Participants and Activities in Citizen Science Projects

Rob Lemmens, Gilles Falquet, Chrisa Tsinaraki, Friederike Klan, Sven Schade, Lucy Bastin, Jaume Piera, Vyron Antoniou, Jakub Trojan, Frank Ostermann, and Luigi Ceccaroni

Abstract Interest in the formal representation of citizen science comes from portals, platforms, and catalogues of citizen science projects; scientists using citizen science data for their research; and funding agencies and governments interested in the impact of citizen science initiatives. Having a common understanding and representation of citizen science projects, their participants, and their outcomes is key to enabling seamless knowledge and data sharing. In this chapter, we provide a conceptual model comprised of the core citizen science concepts with which projects

R. Lemmens (✉) · F. Ostermann
Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands
e-mail: r.l.g.lemmens@utwente.nl

G. Falquet
Centre for Informatics, University of Geneva, Geneva, Switzerland

C. Tsinaraki · S. Schade
European Commission, Joint Research Centre (JRC), Ispra, Italy

F. Klan
Institute of Data Science, German Aerospace Center (DLR), Jena, Germany

L. Bastin
European Commission, Joint Research Centre (JRC), Ispra, Italy

Department of Computer Science, Aston University, Birmingham, UK

J. Piera
Institute of Marine Sciences, Barcelona (ICM-CSIC), Barcelona, Spain

V. Antoniou
Hellenic Army Geographic Directorate, Athens, Greece

J. Trojan
Faculty of Logistics and Crisis Management, Tomas Bata University in Zlin, Zlin, Uherske Hradiste, Czech Republic

L. Ceccaroni
Earthwatch Europe, Oxford, UK

K. Vohland et al. (eds.), *The Science of Citizen Science*,
https://doi.org/10.1007/978-3-030-58278-4_9

and data can be described in a standardised manner, focusing on the description of the participants and their activities. The conceptual model is the outcome of a working group from the COST Action CA15212 *Citizen Science to Promote Creativity, Scientific Literacy, and Innovation throughout Europe*, established to improve data standardisation and interoperability in citizen science activities. It utilises past models and contributes to current standardisation efforts, such as the Public Participation in Scientific Research (PPSR) Common Conceptual Model and the Open Geospatial Consortium (OGC) standards. Its design is intended to fulfil the needs of different stakeholders, as illustrated by several case studies which demonstrate the model's applicability.

Keywords Participation tasks · Dataset description · Data integration · Project description · Project metadata · Interoperability

Introduction

Every citizen science project is unique in terms of its participants, governance model, scientific methodology, measures of quality control, and campaigns conducted, as well as the data and knowledge it generates. It is necessary to determine the current status and trends of citizen science in order to inform relevant decision-makers and to increase the impact of citizen science projects by coordinating their efforts. It is a significant challenge to collate and analyse the fragmented and diverse citizen science data that is generated (e.g. records and observations). The COST Action Working Group 5, tasked with *improving data standardisation and interoperability*, sought solutions to these challenging tasks. The resulting modelling effort was closely linked to the larger objectives of the international Data and Metadata Working Group of the US Citizen Science Association (CSA),[1] which includes members of the European Citizen Science Association (ECSA), and the Australian Citizen Science Association (ACSA). In this chapter, we introduce a model of core citizen science concepts, which is one of the major outcomes from the COST Action working group. This *conceptual model* is implemented using formal and standardised knowledge representation techniques and allows both human interpretation and computer-based processing.

Such a conceptual model fosters the representation of citizen science globally by:

- Enabling a common understanding of the terminology, for example, for indexing literature, outreach and education, and delimiting the field within the generic domains of IT, scientific projects, and data standards
- Forming a basis for facilitating the alignment and integration of data produced in citizen science projects by fostering standardisation and interoperability (being able to share information seamlessly across activities)

[1] http://citizenscience.org/association/about/working-groups/

- Facilitating the creation of software, database *schemas*, and data interchange formats for the development of new citizen science applications
- Supporting potential project participants and other stakeholders to better understand the tasks involved in a particular citizen science project

This chapter first briefly introduces its approach to a conceptual model for citizen science, the stakeholders concerned, and the methodology used. It also defines the concepts, relations, and constraints (axioms) of a *volunteer participation conceptual model*. It then explores connections between these and a traditional *scientific activity conceptual model* that includes the project, funding, outcomes, datasets, and domain. Next, this chapter provides a detailed description of the conceptual model providing the basic concepts about participants and their activities. The conceptual model links to existing standards by adopting and unifying suitable top-level concepts that appear in those data models. The chapter finally demonstrates the applicability of the conceptual model based on case studies before turning to a roadmap for future use and research.

Towards a Conceptual Model for Citizen Science

A conceptual model for citizen science needs to cover three main aspects (and their corresponding metadata):

- Information about citizen science *projects*
- The *people* involved
- Project *outcomes*, typically data and publications

When we refer to citizen science as a domain, we follow the definition outlined by Haklay et al. (this volume, Chap. 2).

Project metadata includes general information such as project name, aim, runtime, the topic or field of science addressed, a contact person or contact point, the organisations involved, and funding sources. In addition, metadata includes information which is specific to citizen science, for example, about the participants (their motivations, skills, knowledge level, and training undertaken). This also includes information that might be important to interested citizens, for example, how to participate and the type and difficulty level of volunteer tasks required.

In addition to project-related metadata, a conceptual model for citizen science needs to provide descriptive elements for project *outcomes*, which are typically data and publications. Data records are usually bundled into datasets following a certain data schema. Typical information about datasets includes name, license, access rights, geographic coverage, access information, submission date, creator, data quality requirements (see Balázs et al., this volume, Chap. 8), information on how data was collected, by whom and with which skills and expertise, and how quality was assessed and verified. Citizen science projects differ from other types of projects in that they employ novel ways of collecting data (e.g. a mobile app specifically designed for a project) and employ data collection protocols that are not common in traditional scientific research projects.

The major difference between a traditional scientific research project and a citizen science project is the participation of non-professionals in scientific activities. Therefore, our formal description of citizen science projects (project metadata) focuses on the representation of the *people* involved, their motivations and skills, the tasks they perform, how they were recruited, how their privacy is protected, how they collect data, and how the quality of their contributions is assessed.

Stakeholders

The spectrum of stakeholders (as identified by Göbel et al. 2017) who require reliable information about citizen science projects includes:

1. Participants
2. Academic and research organisations
3. Government agencies and departments
4. Civil society organisations, informal groups, and community members
5. Formal learning institutions
6. Businesses or industry

The requirements of the stakeholders listed above vary; for example, a certain level of interoperability is essential for government agencies as well as academic and research organisations. However, in the case of community-driven citizen science projects, the stakeholders are participants or informal groups who do not prioritise interoperability but need data to be provided in a user-friendly format.

Methodology

In this chapter, we define a conceptual model as *a representation of a knowledge domain or system*, with which people can understand the meaning of its underlying concepts and which can be used by computer software to meaningfully process its related data. There are a variety of conceptual models, ranging from simple *mind maps* and *concept maps* (Novak and Cañas 2008) to complex *ontologies* (Simperl and Luczak-Rösch 2014). Commonly, concepts are described in terms of their definitions and the (labelled) relationships between them. In formal models, concepts are often called *classes* (e.g. ‘project’), and classes have specific examples, called *instances* (e.g. ‘OpenStreetMap’). All those elements can be represented visually (for human understanding) and in formal computer language (for data integration). In this chapter, we apply commonly used techniques from *ontology engineering* and *concept map construction*.

The core conceptual model elements and associated metadata presented here draw on previous research and existing vocabularies. In particular, they utilise the Public Participation in Scientific Research (PPSR) Common Conceptual Model (described in Bowser et al. 2017) and the core requirements in the associated conceptual model PPSR-Core.

The conceptual model developed in this chapter is intended to fulfil the needs of different stakeholders, as shown in several case studies. To address this requirement, we refined core elements of the PPSR model based on existing case studies; these informed the identification of additional core concepts.

The conceptual model presented is not the only model that suits the field of citizen science, but it provides a view of the technical aspects of the discipline in order to help stakeholders understand the domain and foster interoperability across applications. It is an evolving model that is becoming established via an international consensus process.

Related Conceptual Models

Conceptual Models of Projects and Participants

A number of models that allow projects to be described in general and scientific projects to be described specifically have been previously developed outside the citizen science community. Those models aim to represent knowledge about a subject domain such as relevant concepts and relationships between those in a very formal way (e.g. in terms of an ontology) or less formally by means of a controlled vocabulary. The following table gives an overview of these models and summarises which facets of projects and their participants they cover. The models listed were carefully considered when designing our conceptual model for the citizen science domain.

We will now summarise the models listed in Table 9.1. FRAPO describes projects and their outputs in terms of publications and datasets. SCoRO models the roles of project participants and their contributions. It allows the linking of individuals' contributions to project outputs. PROV-O can be used to model projects, their outcomes, and how the outputs are produced and by whom. The Project Description Ontology extends PROV-O and is an attempt to model projects in a domain-agnostic way. FOAF can be used to characterise participants of a citizen science project. The FaBiO model is discussed in the next section.

Table 9.1 State of the art of conceptual models of projects and participants

Conceptual model	Aspects related to projects and participants that are covered by the model
The Funding, Research Administration and Projects Ontology (FRAPO) (doi: https://doi.org/10.13140/RG.2.2.26124.92802) (Peroni and Shotton 2018)	Administrative information related to projects (e.g. budget, project partners)
	Information related to project funding
	Project outputs (e.g. in terms of publications and datasets)
The Scholarly Contributions and Roles Ontology (SCoRO) (http://www.sparontologies.net/ontologies/scoro) (Peroni and Shotton 2018)	Roles of people working together on a project (e.g. data creators/managers/curators, principal investigators)
	Contributions of project participants (e.g. intellectual contributions such as conception and design of experiments)
The Friend of a Friend vocabulary (FOAF) (http://www.foaf-project.org)	Interests of participants
	Information about participants (e.g. name, age, home page)
	Relationships between participants (e.g. who knows whom)
The Bibliographic Ontology (FaBiO) (http://www.sparontologies.net/ontologies/fabio) (Peroni and Shotton 2012)	Project outcomes in terms of published or publishable results (e.g. scientific publications)
The PROV Ontology (PROV-O) (http://www.w3.org/TR/prov-o/)	Provenance information about projects (e.g. which project outcomes were produced by whom, with what information and input, and via which project activities)
Project Description Ontology (https://github.com/dr-shorthair/project-ont)	General information about projects that is independent from a specific application domain

Conceptual Models of Project Outcomes

FaBiO models published or publishable project outcomes such as scientific publications. The Project Documents Ontology (PDO) describes other project-related documents such as minutes and status reports.

A number of models provide descriptive elements for datasets. This includes the World Wide Web Consortium (W3C) Recommendation Data Catalog Vocabulary – Version 2 (DCAT)[2] that enables the description of datasets and data services in catalogues. More general specifications, such as Dublin Core,[3] define elements for the description of arbitrary resources, not just publications.

Several conceptual models have been developed for the formal description of observational data and measurements as common outcomes of scientific projects, for example, in the life sciences and geosciences, but also in citizen science. A number

[2]https://www.w3.org/TR/vocab-dcat-2/

[3]https://www.dublincore.org/specifications/dublin-core/dcmi-terms/

of standards with overlapping semantics have emerged: the Semantic Sensor Network (SSN) Ontology,[4] a joint standard of Open Geospatial Consortium (OGC) and W3C, that specifies the semantics of sensors and their observations, and its proposed extensions;[5] the OGC/ISO Observation and Measurement (O&M) conceptual model;[6] and the W3C Data Cube Vocabulary,[7] focusing specifically on the representation of multi-dimensional data. The data model of OGC's SensorThings API[8] is based on the OGC/ISO O&M model and closely resembles it. Although several ongoing community-driven attempts aim to harmonise the description of observational data in order to facilitate data integration, none of the existing data models have been adopted by a scientific community as a whole. However, attempts have been made to link coexisting models by establishing mappings to align different models, for example, the SSN Ontology offers alignments to the OGC/ISO O&M model. An OGC discussion paper (Simonis and Atkinson 2016) gives a helpful overview of standardised information models with relevance to citizen science data and describes a data model for the exchange of citizen science sampling data based on existing standards.

In parallel, practitioners such as data managers of research data infrastructures have developed their own vocabularies and models that do not rely on existing standards. In the biomedical domain, several domain-specific data models have been developed. Those include the Extensible Observation Ontology (OBOE)[9] (Madin et al. 2007) and the Biological Collections Ontology.[10] There are hundreds of domain-specific metadata standards and data models facilitating the description of scientific data in specific scientific domains, for example, BioPortal[11] currently lists 838 ontologies in the biomedical domain. Finally, the catalogue of the Digital Curation Centre[12] lists numerous disciplinary metadata standards.

The Proposed Conceptual Model for Citizen Science

As a starting point, we considered the top-level model of the CSA report (Bowser et al. 2017) (see Fig. 9.1), which proposed a grouping of the existing attributes into a set of modules. The titles of the modules were adapted by Working Group 5 (see COST Action CA15212 Working Group 5 2018a). The Project Metadata Model

[4] https://www.w3.org/TR/vocab-ssn/
[5] https://www.w3.org/TR/vocab-ssn-ext/
[6] https://www.iso.org/standard/32574.html
[7] https://www.w3.org/TR/vocab-data-cube/
[8] https://www.ogc.org/standards/sensorthings
[9] https://github.com/NCEAS/oboe
[10] http://www.obofoundry.org/ontology/bco.html
[11] http://bioportal.bioontology.org/
[12] http://www.dcc.ac.uk/resources/metadata-standards

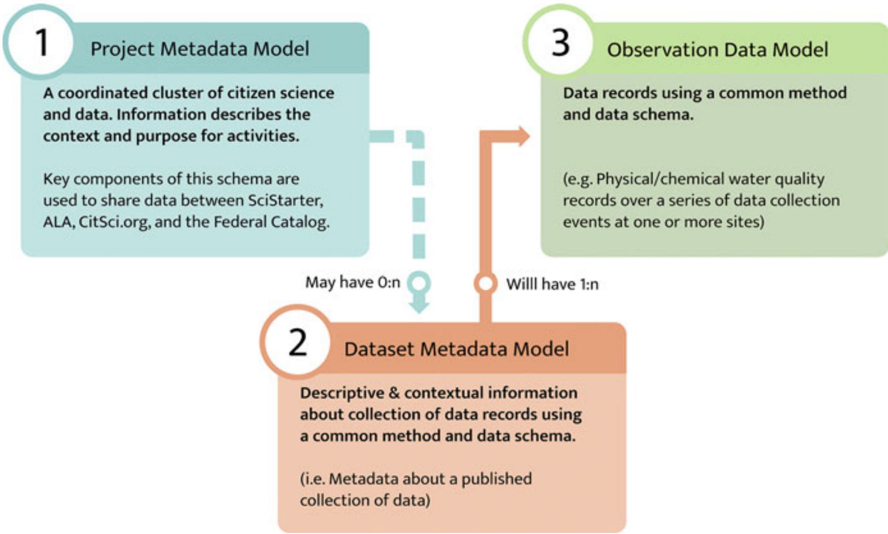

Fig. 9.1 The PPSR-Core conceptual model adapted from the Public Participation in Scientific Research (PPSR) Common Conceptual Model (Bowser et al. 2017). The 0:n (and the dashed arrow) means that a Project Metadata model may have zero or more Dataset Metadata Models. The 1:n (and the solid arrow) means that a Dataset Metadata Model will have one or more Observation Data Models

describes the key components of a citizen science project. The Dataset Metadata Model characterises a dataset as an output of a project and describes its geographic coverage, data collection method, and access rights. The Observation Data Model contains a detailed description of the data elements that are used in a dataset, for example, the meaning of specific sensor observations (such as nitrogen/nitrate concentration in a water quality measurement).

Project Description

The development of the PPSR-Core model was driven by the requirements of the implementations available at the time. As a consequence, it is tied to these implementations, and a conceptual model allowing for better project content representation is still not available. In addition, PPSR-Core still includes some domain-specific properties, especially from the biodiversity domain. Since citizen science activities take place in different disciplines and focus on specific aspects that vary across activities, a model that tries to capture everything in the domain can become complicated and difficult to manage.

In order to exploit the citizen science knowledge encoded in PPSR-Core and, at the same time, overcome the above-mentioned drawbacks, we have developed a

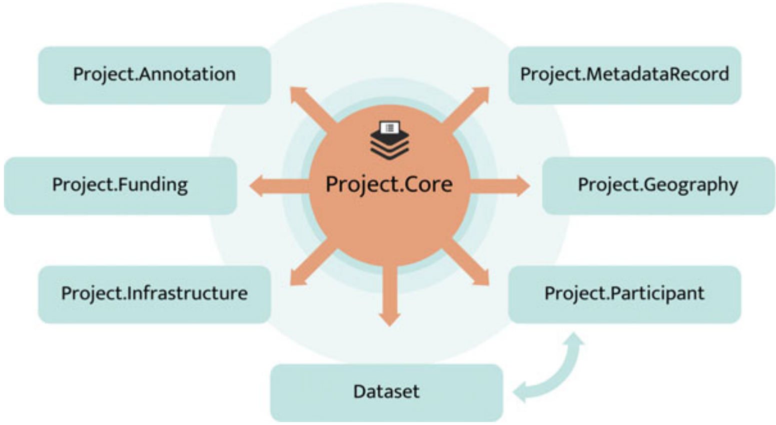

Fig. 9.2 An overview of the structure of the main conceptual model, highlighting the different modules. The arrows indicate the dependency between the modules. The connection between the Dataset and Project.Participant modules indicates that there are relationships between concepts across these modules

modular conceptual model for the representation of citizen science knowledge (see COST Action CA15212 Working Group 5 2018b). This model comprises different modules that are all linked to the Project.Core module that captures essential project information (see Fig. 9.2 for an overview of the structure of our conceptual model). The Project.Core module includes many properties imported from PPSR-Core, like project name, website, start and end date, etc., and unifies the other modules. These modules include:

- The Project.MetadataRecord module, which captures general information about the project, including its provenance
- The Project.Annotation module, which captures information, like tags, used for annotating project descriptions
- The Project.Funding module, which captures project funding information
- The Project.Infrastructure module, which captures information about project infrastructure (hardware, software, services, etc.)
- The Project.Geography module, which captures geographical information about the project
- The Dataset module, which captures information about project datasets
- The Project.Participant module, which captures information about project participants and their activities within a project

Due to the wide scope of the main conceptual model for citizen science, it was developed in phases. In this chapter, our attention is focused on *participation* and *participant activities* in citizen science projects. Related initiatives from CSA, ECSA, ACSA, and OGC are accounted for and the model is tested with case studies.

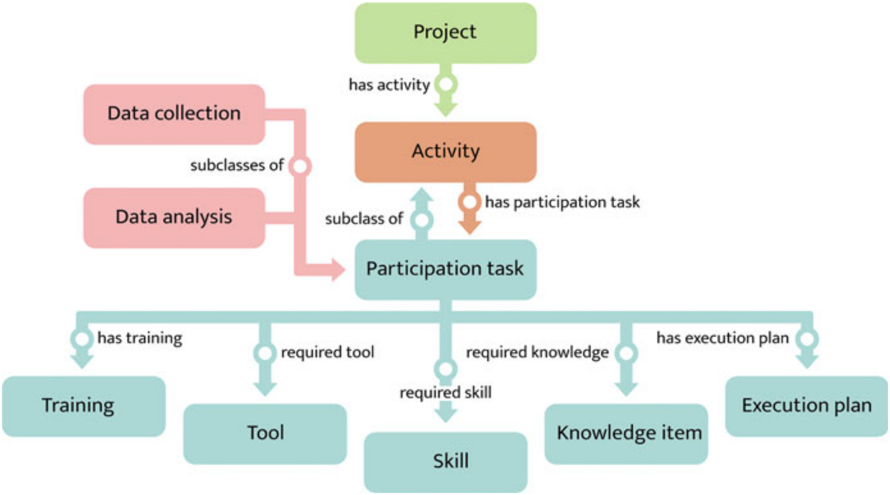

Fig. 9.3 Excerpt (part a) of the conceptual model on citizen participation. The different boxes represent concepts; the arrows represent relationships

Since the role of the citizens as participants is the main difference between citizen science projects and traditional research projects, in the following section, we will discuss the Project.Participant module in more detail.

Together, Table 9.1 and Figs. 9.2 and 9.3 outline all the concepts and relationships in the conceptual model related to the Project.Core and Project.Participant modules. Here, we describe a selection of the concepts; a full list of descriptions is currently under development and available in the model repository.[13]

Participation and Activity Description

At the heart of the Project.Participant module lie the relationships between the participants, their activities, their outputs, and the skills, knowledge, and tools required to perform them. A project has one or more activities, and these are performed by participants with a variety of roles and motivations, during a specified time range.

In the model, the Activity concept (see Figs. 9.3 and 9.4) represents activities that belong to a Project. A general *activity*, such as ‘Collecting data about bird migration’, may contain a number of tasks. A *task* is an activity with a specific goal and a limited duration (a kind of transaction), such as ‘Taking a picture of a bird and storing it in an image collection’ or ‘Validating a bird identification’. The description

[13]Doi: https://doi.org/10.5281/zenodo.3695444

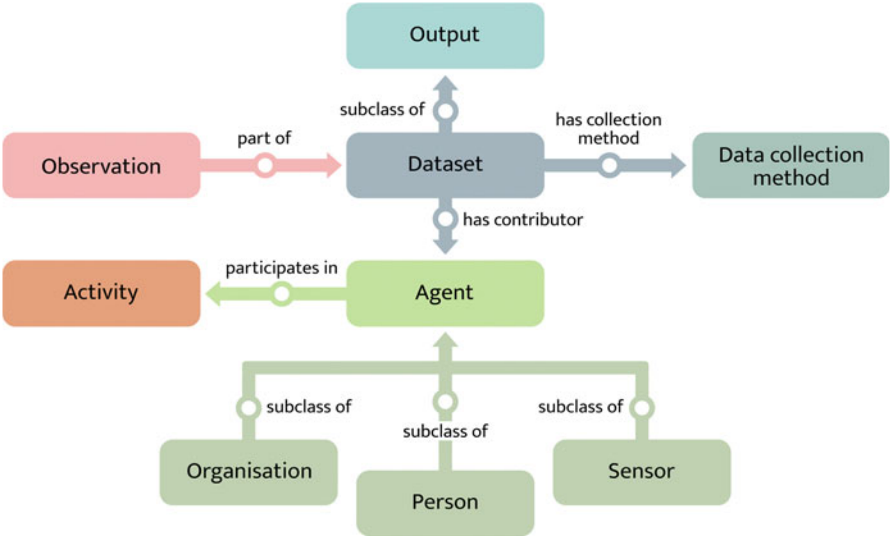

Fig. 9.4 Excerpt (part b) of the conceptual model part on citizen participation

of a task includes details of the knowledge, skills, and tools required as well as the training available and its *execution plan*.

The Agent concept in Fig. 9.3 generalises the idea of participants to groups of people in particular organisations and to machines, such as sensors. An *instance* of the Agent concept represents a type of agent, for example, 'registered Zooniverse user' or 'mapping agency'.

The Activity description includes its output (e.g. dataset, publication, software) that can be composed of a number of *output items*. A project may acknowledge the participation of an actor in the production of an output item. In this case, the description of an output item includes a link to the role played by the actor its production. The description of a Project also includes its participant recruitment technique and its privacy protection policy. The dataset as an entity is handled in our model as a specific type of output. Its details are described in a separate module (see Fig. 9.1), and although they are required for interoperability, they are beyond the scope of this chapter. The same holds true for the semantics of a dataset's content, which is described in the Data Model (Fig. 9.1). Here we make use of existing standards, such as the underlying data models of the SensorThings API (Footnote 8) and the SSN Ontology (Footnote 4).

The concepts depicted in Fig. 9.4 cover participation and its requirements. The model does not claim to be exhaustive, but rather serves as a backbone. Each of the branches, such as tools and skills, can themselves be described by external models. The subclassification is also not exhaustive. Part a (Fig. 9.3) and part b (Fig. 9.4) are connected through the Activity concept.

Application in Case Studies

This section explains how the conceptual model for citizen science can be used in specific case studies, that is, how the different characteristics of a project – its participants, its data, etc. – can be described by using the model. The case studies represent projects with different domains, community sizes, and types of participation in order to demonstrate the breadth of citizen science applications that the model can accommodate. The first sub-section highlights four different projects. Here we demonstrate how they can be described with the help of our model in order to understand project content and metadata. The second sub-section illustrates another use for our model: the application of its concepts and structure for (1) creating project descriptions in a specific inventory and (2) structuring data collection.

Instantiation of Projects

After providing a short introduction to the four selected citizen science projects, we use our conceptual model as a skeleton for each specific project. Where applicable, the concepts (as depicted in Figs. 9.3 and 9.4) have been *instantiated* for each project; see Tables 9.2 and 9.3. In other words, a concept is assigned a project-specific value where possible and applicable. This means that specific projects, their activities, participants, data outputs, etc., are described with the help of the conceptual model. Using this common model allows the projects to be compared and combined, thus increasing interoperability between the projects and their elements. It should be emphasised that in the tables only a few examples are provided and that each entry in the table corresponds to a concept in the model, which is more than just a flat table. For example, a project can have multiple participation tasks, each using different tools; and a project can produce multiple, different datasets, and so on. We will now introduce our case studies.

OpenStreetMap. OpenStreetMap (OSM) is a well-known *crowdsourcing* project in which thousands of volunteers maintain an online map of the world. OSM has all the characteristics of participation and data handling we see in many other citizen science projects. In addition, OSM is an essential geographical reference for many citizen science projects.

Bash the Bug (Zooniverse). The objective of the Bash the Bug project is to improve tuberculosis diagnosis. The task of the volunteers is to accurately determine which antibiotics are effective for each of the collected tuberculosis samples. This is carried out by analysing pictures of plates showing the effects of several antibiotics on the tested sample.

Mars in Motion (Zooniverse). Mars in Motion was created to look for and identify geological changes on the surface of Mars over time by gathering in-depth data on the type of features that are detected. It is part of the i-Mars.eu project, which includes several European partners, and is focused on developing tools and datasets to increase the exploitation of space-based data from the US National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA) Mars mission beyond the scientific community.

Table 9.2 Instantiation of the conceptual model with OSM, Bash the Bug, Mars in Motion, and MICS

	Project instantiation	Project instantiation
Concept	OpenStreetMap	Bash the Bug
Activity	OSM mapping event OR data capture facilitation	Assessment of antibiotic effect (a)
ActivityAgentDependency		(a) – (u)
Agent		Zooniverse user (u)
AssociatedPublications		
Consortium		
ContactDescription		
ContactPoint		
DataCollectionMethod	On-screen digitising	Web crowdsourcing (on-screen recognition)
Dataset	OSM change set	Antibiotics sensitivity dataset
Description		
Email		
EndDate		
ExecutionPlan	Mapping campaign, HOT task	
GeographicExtent		
GroupOfAgents		
Hardware		
Initiative		
KnowledgeItem	GIS fundamentals	
Machine		
MeansOfContact		
Motivation	Contribute free map data	
Name		
Observation	OSM feature	
Origin		
OGC:Datastream		
OGC:Sensor		
Organisation		
OrganisationCategory		
Output		
OutputItem	Geometric primitive (point, lines)	Record (infection sample –>sensitivity to antibiotics)
ParticipantAcknowledgment	OSM user id	Authorship recognition
Participation	OSM contribution	
ParticipationTask		Image classification
Person	OSM contributor	
PrivacyProtectionPolicy	Terms of use in signup	https://www.zooniverse.org/privacy
PrivacyProtectionTechnique	Person is behind user id	Zooniverse informed consent
Project	OSM.org	

(continued)

Table 9.2 (continued)

Concept	Project instantiation OpenStreetMap	Project instantiation Bash the Bug
Publication	OSM upload	
RecruitmentMethod	Mapping event invitation, conferences, courses	
Role	OSM mapper; OSM validator	
Sensor		
Skill	Image interpretation	
Software (as tool)	Mobile/web editors	Zooniverse classification web app
Software (as output)		
SpatialAreaofInterest		
StartDate		
Status		
TemporalExtent	OSM mapping event duration	
Tag		
Tool	OSM editor, e.g. iD, JOSM, etc.	
Training	Online self-training, Wiki specs	8-slide tutorial
WebPage		

Concept	Project instantiation Mars in Motion	Project instantiation MICS Project
Activity	Online classification of Martian surface (a)	Online assessment of project impact (a)
ActivityAgentDependency	(a) – (u)	(a) – (u)
Agent	Human (u) and machine	Human (u) and machine
AssociatedPublications	https://doi.org/10.1016/j.jag.2017.05.014	
Consortium	FP7 i-Mars project	Horizon 2020 Project
ContactDescription	Project contact	Project contact
ContactPoint	James Sprinks	Luigi Ceccaroni
DataCollectionMethod	On-screen analysis	On-screen survey
DataSet	ESA Mars Express HRSC Camera	MICS Impact Assessment Corpus of Knowledge
Description		
Email	james.sprinks@nottingham.ac.uk	lceccaroni@earthwatch.org.uk
EndDate		
ExecutionPlan	Project deliverables/timeline	Project deliverables/timeline
GeographicExtent		
GroupOfAgents		
Hardware	N/A	N/A

(continued)

Table 9.2 (continued)

Concept	Project instantiation Mars in Motion	Project instantiation MICS Project
Initiative	iMars	MICS
KnowledgeItem	Scientific concept	Project outputs
Machine	Change detection algorithm	Impact assessment algorithm
MeansOfContact	Email	Email
Motivation	Contribute to understanding of Martian surface evolution	To understand the impact of their project
Name	Jo Bloggs	Joanna Blogson
Observation	Martian geomorphological feature	Impact measurement
Origin		
OGC:Datastream	N/A	N/A
OGC:Sensor	N/A	N/A
Organisation	iMars project consortium	MICS project consortium
OrganisationCategory	European FP7 Project Funded	European H2020 Project Funded
Output	Martian geomorphological features that evolve temporally	Impact assessment
OutputItem	Distance, speed, typology	Report
ParticipantAcknowledgment	Authorship, acknowledgement in publication	Acknowledgement on website
Participation	Mars in Motion contribution	MICS contribution
ParticipationTask	To detect changes on the Martian surface, through comparison of two images	To evaluate the impact of their citizen science project
Person	Mars in Motion participant	MICS participant
PrivacyProtectionPolicy	Zooniverse privacy policy	MICS privacy policy
PrivacyProtectionTechnique	Zooniverse informed consent	MICS informed consent
Project	https://www.zooniverse.org/projects/imarsnottingham/mars-in-motion/	https://mics.tools/
Publication	Database of Martian feature change	Project impact assessment report
RecruitmentMethod	Online correspondence to existing Zooniverse community	
Role	Project participant	Project coordinator
Sensor		
Skill	Image interpretation	Assessing project processes/output that have impact
Software (as tool)	Web platform (Zooniverse)	Web platform
Software (as output)	N/A	N/A
SpatialAreaofInterest		
StartDate		
Status		

(continued)

Table 9.2 (continued)

Concept	Project instantiation Mars in Motion	Project instantiation MICS Project
TemporalExtent	30 years + of historical data	Unlimited
Tag		
Tool	Online Zooniverse image annotation	Online survey data entry
Training	Online training (compulsory)	Online guidance/ examples
WebPage		

Table 9.3 Instantiation of the conceptual model with the JRC Citizen Science Project Inventory and the Participatory Toponym Handling Project

Concept	Project instantiation JRC Citizen Science Project Inventory	Project instantiation Participatory Toponym Handling Project
Activity		Toponymic data handling OR toponymic field survey
ActivityAgentDependency		
Agent		Citizens, local government, university, Badan Informasi Geospasial (BIG)
AssociatedPublications		doi: https://doi.org/10.3390/ijgi7060222; doi: https://doi.org/10.3390/ijgi8110500
Consortium		
ContactDescription		
ContactPoint		
DataCollectionMethod		Geographic data collection: fieldwork and office treatment
DataSet		Toponymic files and gazetteers
Description	Brief description	
Email		
EndDate	End year	
ExecutionPlan		Workshops (toponymic field survey campaign), citizen science project on toponym, local government project on toponym, HOT (Indonesia) task
GeographicExtent	Geographical extent	West Java province
GroupOfAgents		
Hardware		Tablets
Initiative		
KnowledgeItem		Toponymy
Machine		

(continued)

Table 9.3 (continued)

Concept	Project instantiation JRC Citizen Science Project Inventory	Project instantiation Participatory Toponym Handling Project
MeansOfContact	Contact	
Motivation		Contribute toponymic data, preserve embedded knowledge on toponyms, collect toponyms in their surrounding areas
Name	Name	
Observation		Place names, coordinates, history, pronunciation
Origin	Source	
OGC:Datastream		
OGC:Sensor		
Organisation		
OrganisationCategory	Lead organisation category; project initiator category	
Output		Dataset (toponyms, gazetteers)
OutputItem		Geometric primitive (point), audio (pronunciation of toponym)
ParticipantAcknowledgment		Reward-based approach, incentive, capacity building opportunities
Participation		Local toponyms
ParticipationTask		Providing place name and related information
Person		CitSciTopon contributor, LocalGovt contributor, OSM contributor
PrivacyProtectionPolicy		Terms of use in signup
PrivacyProtectionTechnique		Person is behind user id
Project	Project	
Publication		SAKTI upload
RecruitmentMethod		Toponymic training event invitation, Toponymic survey invitation
Role		Data collector; data verificator/validator
Sensor		
Skill		Interview, communication with local people
Software (as tool)		Mobile/web editors
Software (as output)		
SpatialAreaofInterest	Geographic coverage	
StartDate	Start year	

(continued)

Table 9.3 (continued)

Concept	Project instantiation JRC Citizen Science Project Inventory	Project instantiation Participatory Toponym Handling Project
Status	Still active	
TemporalExtent		National naming authority event programme duration
Tag	Primary environmental domain; primary environmental field; primary category of project	
Tool		EpiCollect, SAKTI application, ODK and OSM OpenMapKit
Training		Training on toponymy, workshops, focus group discussions
WebPage	Website	

MICS. The MICS project provides an integrated platform of metrics and instruments to measure both the costs and the benefits of citizen science. These metrics and instruments consider the impacts of citizen science on the following domains: society, governance, the economy, the environment, and science.

Deployment of the Conceptual Model

In addition to the basic metadata provision outlined in the previous section, the conceptual model can be used as a structure for project-related activities. Two case studies are provided here.

JRC Citizen Science Project Inventory

The European Commission Joint Research Centre (JRC) has developed a multidisciplinary data infrastructure (Friis-Christensen et al. 2017) to facilitate open access to its research data, in line with the recent *open data* trend (Trojan et al. 2019). The JRC Data Infrastructure[14] has helped establish requirements for dataset metadata. The JRC datasets are published in the JRC Data Catalogue and are described by metadata that follow a modular metadata schema. The schema consists of (1) a core profile which defines the common elements of metadata records, based on the reference standards DCAT-AP (ISA DCAT-AP 2015) and DataCite (2016), and (2) a set of extensions, which defines elements specific to given domains (geospatial, statistical, etc.), based on existing metadata standards.

[14] http://data.jrc.ec.europa.eu/

In addition, the JRC Citizen Science Project Inventory has supported the JRC in describing projects. The JRC Citizen Science Project Inventory was initially developed as one of the outcomes of the study *Citizen Science for Environmental Policy: Development of an EU-wide Inventory and Analysis of Selected Practices* (Bio Innovation Service 2018; Turbé et al. 2019). This project was executed by the European Commission (DG Environment), with the support of the JRC. The project also included additional contracted partners: the Bio Innovation Service (France), the Fundacion Ibercivis (Spain), and the Natural History Museum (UK). The main objective was to build an evidence base of citizen science activities to support environmental policies in the European Union (EU). Specifically, the goal was to develop an inventory of citizen science projects relevant to environmental policy and assess how these projects contribute to the United Nations Sustainable Development Goals (SDGs). To this end, a desk study and an EU-wide survey were used to identify 503 citizen science projects of relevance to environmental policy. The resulting project inventory has been published in the JRC Data Catalogue[15] and is updated on a regular basis (it also considers new entries suggested via an online survey).[16]

The Citizen Science Explorer,[17] a dynamic catalogue provided as part of the JRC GitHub space, has been developed to provide more visibility to the JRC Citizen Science Project Inventory and to showcase the opportunities for knowledge sharing and management. The inventory is available in the form of *comma-separated values* (CSVs),[18] JSON,[19] and JSON-LD.[20] Therefore, the conceptual model described in this chapter does not allow us to represent all the information available in the inventory but does allow us to structure its core entities in a standardised way.

There are other initiatives which can be considered as case studies for identifying stakeholders needs. These include activities covered by Earthwatch (e.g. the MICS project, in which the impact of citizen science projects is measured) and COST Actions throughout Europe.

Participatory Toponym Handling Project

One application case where the citizen science conceptual model had a direct influence, and which in turn can be used to shape future developments of the conceptual model, concerns the collection and maintenance of place names (or *toponyms*) in Indonesia.

This particular case study was motivated by the fact that many national mapping agencies (and agencies responsible for the naming of places in databases and gazetteers) have scarce or insufficient resources. At the same time, many citizens

[15]http://data.jrc.ec.europa.eu/dataset/jrc-citsci-10004

[16]https://ec.europa.eu/eusurvey/runner/CSProjectInventory

[17]https://github.com/ec-jrc/citsci-explorer/

[18]https://tools.ietf.org/html/rfc4180

[19]https://tools.ietf.org/html/rfc8259#section-4

[20]https://www.w3.org/TR/json-ld/

have rich local and traditional knowledge of toponyms. Indonesia, in particular, has many regional and local languages and a varied topography. Including local and traditional knowledge is also relevant from a research point of view, because it can, for example, uncover yet unwritten histories.

The Geospatial Information Agency of Indonesia (Badan Informasi Geospasial, BIG[21]) is responsible for toponyms in Indonesia. BIG conducted two pilot projects in 2015 (Yogyakarta) and 2016 (Lombok) on the involvement of citizens in toponym handling. The Indonesian approach includes many stakeholders, combining both *top-down* and *bottom-up* elements: national legislation provides regulations and procedures, while their implementation relies on local actors. However, local governments tasked with the implementation often lack the capacity to provide the required skills and resources.

The pilot projects led to the development of a *participatory toponym handling framework* (Perdana and Ostermann 2018). More importantly for this chapter, the framework adopted several concepts from an early version of Working Group 5's citizen science conceptual model. Thus, although the framework has been subsequently improved and significantly expanded through collaborative learning, including focus group discussions with stakeholders and workshops (Perdana and Ostermann 2019), this example shows the utility of an early version of the conceptual model for designing a project involving citizens.

The concrete participatory toponym handling approach that was developed is also expected to influence ongoing legislation processes. Furthermore, it resulted in three experimental toponym collection projects in late 2018 (their outcomes will soon be published).

Using this chapter's conceptual model, we can describe the participatory toponym handling. The main Activity is the collection of place names, either entirely new ones or updating existing ones. The Agents carrying out this activity are citizens, local government officials, experts from the national mapping agency, and academics/researchers. The DataCollectionMethod is field surveys using tablets, supplemented by office-based processing. The created Datasets are initially forms completed by participants (Observations) with multimedia elements (e.g. audio recordings of pronunciation) and ultimately enriched gazetteers. Therefore, the ParticipationTask is to provide place names and related information. The Motivation is to contribute toponymic data, preserve embedded knowledge on toponyms, and collect toponyms in their surrounding areas.

Roadmap for Future Research and Use

The benefits of using the conceptual model presented in this chapter are twofold: *human understanding* of citizen science project characteristics and *machine processing* of these characteristics. Further technical development and

[21] https://big.go.id/en

documentation of best practices will be required to support the model in use. Humans wishing to discover, evaluate, and contribute to projects will require intuitive visualisation of the conceptual model and well-designed tools for search and query. Machines that use the model for data alignment will require well-designed APIs, and repositories of standards, schemas, and agreed terms, with reliable access mechanisms.

An example of the context in which this conceptual model could be used is the EU Horizon 2020 Framework Programme project EU-Citizen.Science, which aims 'to build a central platform for citizen science in Europe, a place to share useful resources about citizen science, including tools and guidelines, best practices and training modules'.[22] By utilising a metadata schema such as this conceptual model, a greater understanding of data types, their structure, and their relationships can be achieved. Adopting the conceptual model will also ensure that the tools, guidelines, and training developed are as widely applicable and usable as possible.

The following recommendations are designed to foster the uptake of the conceptual model by the citizen science community in order to increase citizen science interoperability:

- Develop procedures to respond to existing regulatory or legal frameworks related to citizen science, such as the implementation of the INSPIRE Directive (in Europe) and the provision of related best practices and tools.[23]
- Involve the ECSA, CSA, ACSA, and the Citizen Science Global Partnership (CSGP) in the definition of an agenda for the model's practical implementation and possibly as hosts for interoperable catalogues of citizen science projects and data. They could also provide guidelines on the use of existing solutions.
- Include a dedicated section on ECSA, CSA, ACSA, EU-Citizen.Science, and CSGP websites to explain the conceptual model and provide introductory information.
- Develop extensions related to more diverse outcomes, such as mathematical theorems, hardware, and policy and societal impacts.
- Develop communication approaches to help practitioners navigate through the various standards and concepts (e.g. a 'choose your own adventure' approach; see also the Digital Curation Centre[24] for additional ideas).

Implementing the proposed recommendations will take some time and also require collaboration across communities. The publication of the conceptual model outlined in this chapter should support this process. In addition, some of the work needed to fulfil the recommendations is already in progress and will ultimately be disseminated through citizen science community channels.

[22]https://eu-citizen.science/about/

[23]https://inspire-reference.jrc.ec.europa.eu/vocabularies/geospatial-standards/sensorml

[24]http://www.dcc.ac.uk/

References

Bio Innovation Service. (2018). *Citizen science for environmental policy: Development of an EU-wide inventory and analysis of selected practices*. Final report for the European Commission, DG Environment, in collaboration with Fundacion Ibercivis and The Natural History Museum. https://doi.org/10.2779/961304.

Bowser, A., Brenton, P., Stevenson, R., Newman, G., Schade, S., Bastin, L., Parker, A., & Oliver, J. (2017). *Citizen Science Association Data & Metadata Working Group: Report from CSA and Future Outlook* (Workshop Report). Washington, DC: Woodrow Wilson International Center for Scholars.

COST Action CA15212 Working Group 5. (2018a). *On the citizen-science ontology, standards & data.* Minutes of WG5 workshop in Geneva, June 6, 2018 ('Geneva Declaration').

COST Action CA15212 Working Group 5. (2018b). *Towards a new version of the PPSR core conceptual model*. Workshop Report, including Annex.

DataCite. (2016). *DataCite metadata schema for the publication and citation of research data*. Version 4.0. DataCite Specification. DataCite Metadata Working Group. https://doi.org/10.5438/0012.

Friis-Christensen, A., Perego, A., Tsinaraki, C., & Vaccari, L. (2017). *The JRC multidisciplinary research data infrastructure*. In Proceedings of IiWAS 2017, pp. 338–342. https://doi.org/10.1145/3151759.3151810.

Göbel, C., Martin, V., & Ramirez-Andreotta, M. (2017). *Stakeholder analysis: International citizen science stakeholder analysis on data interoperability* (Final Report). https://doi.org/10.13140/RG.2.2.26124.92802.

ISA DCAT-AP. (2015). *DCAT application profile for European data portals*. Version 1.1. ISA Specification. ISA DCAT-AP Working Group. EU ISA Programme. https://joinup.ec.europa.eu/asset/dcat_application_profile/asset_release/dcat-ap-v11

Madin, J., Bowers, S., Schildhauer, M., Krivov, S., Pennington, D., & Villa, F. (2007). An ontology for describing and synthesizing ecological observation data. *Ecological Informatics, 2*(3), 279–296. https://doi.org/10.1016/j.ecoinf.2007.05.004.

Novak, J. D., & Cañas, A. J. (2008). *The theory underlying concept maps and how to construct and use them* (Technical Report IHMC CmapTools 2006-01 Rev 2008-01). Pensacola: Institute for Human and Machine Cognition.

Perdana, A. P., & Ostermann, F. O. (2018). A citizen science approach for collecting toponyms. *ISPRS International Journal of Geo-Information, 7*(6), 222. https://doi.org/10.3390/ijgi7060222.

Perdana, A. P., & Ostermann, F. O. (2019). Eliciting knowledge on technical and legal aspects of participatory toponym handling. *ISPRS International Journal of Geo-Information, 8*(11), 500. https://doi.org/10.3390/ijgi8110500.

Peroni, S., & Shotton, D. (2012). FaBiO and CiTO: Ontologies for describing bibliographic resources and citations. *Journal of Web Semantics, 17*(December), 33–43. https://doi.org/10.1016/j.websem.2012.08.001.

Peroni, S., & Shotton, D. (2018). *The SPAR ontologies*. In Proceedings of the 17th International Semantic Web Conference (ISWC 2018), pp. 119–1360. https://doi.org/10.1007/978-3-030-00668-6_8.

Simonis, I., & Atkinson, R. (2016) *Standardized information models to optimize exchange, reusability and comparability of citizen science data*. OGC Discussion Paper OGC 16-129. http://www.opengis.net/doc/DP/16-129

Simperl, E., & Luczak-Rösch, M. (2014). Collaborative ontology engineering: A survey. *The Knowledge Engineering Review, 29*, 101–131.

Trojan, J., Schade, S., Lemmens, R., & Frantál, B. (2019). Citizen science as a new approach in geography and beyond: Review and reflections. *Moravian Geographical Reports, 27*(4), 254–264. https://doi.org/10.2478/mgr-2019-0020.

Turbé, A., Barba, J., Pelacho, M., Mugdal, S., Robinson, L. D., Serrano-Sanz, F., et al. (2019). Understanding the citizen science landscape for European environmental policy: An assessment and recommendations. *Citizen Science: Theory and Practice, 4*(1), 34. https://doi.org/10.5334/cstp.239.

Rob Lemmens works in the Faculty of Geo-Information Science and Earth Observation (ITC) at the University of Twente, with a research focus on semantic web technologies and the improvement of geo-crowdsourcing methods.

Gilles Falquet is an associate professor at the Center for Computing at the University of Geneva where he leads the Knowledge Engineering research group. His main research interests include knowledge graphs, knowledge visualisation, and the application of knowledge engineering techniques to geographic information systems.

Chrisa Tsinaraki works at the Joint Research Centre of the European Commission. Her studies and professional experience focus on knowledge and data engineering, and her interests cover data ecosystems, multidisciplinary interoperability, open data, digital libraries, and public participation in science for policy.

Friederike Klan heads a research group on citizen science at the Institute of Data Science, German Aerospace Center (DLR). She is an expert on knowledge management and semantic web technologies. Her research interests include (meta)data standardisation for citizen science, data interpretability, and interoperability.

Sven Schade works at the European Commission's scientific and knowledge management service, the Joint Research Centre. With a background in geospatial information science and knowledge engineering, his interests cover data ecosystems, multidisciplinary interoperability, and public participation in science for policy.

Lucy Bastin has a background in spatial ecology and remote sensing, software development for environmental applications, and interoperable standards to communicate uncertainty in scientific models and workflows. She leads development on DOPA (the Digital Observatory for Protected Areas) at the EC Directorate for Knowledge for Sustainable Development and Food Security.

Jaume Piera is leader of ENBIMOS (ENvironmental and Biodiversity monItoring systeMs based on citizen science and participatOry approacheS). His focus is on systems and tools to optimise research based on participatory systems. He coordinates the H2020 Cos4Cloud project, which co-designs services for citizen observatories according to the European Open Science Cloud framework.

Vyron Antoniou is an officer in the Hellenic Army Geographic Directorate. His background is in GIS, VGI, web mapping applications, survey engineering, cartography, and terrain analysis. His current interests are centred on machine learning and deep learning for geospatial applications.

Jakub Trojan is an assistant professor in regional geography and regional development focused on participatory research at Tomas Bata University in Zlin, Uherske Hradiste, Czech Republic. He also works at the Institute of Geonics of the Czech Academy of Sciences, where he investigates new technologies applicable in spatial research and citizen science.

Frank Ostermann joined the Faculty of Geo-Information Science and Earth Observation (ITC) at the University of Twente in 2014. His research interests are collaborative and participatory approaches to creating and processing geographic information ('citizen GI science') and their impact on reproducibility and representativeness of results.

Luigi Ceccaroni manages innovation and strategic research at Earthwatch, an independent research organisation based in Oxford, UK. He is co-vice-chair of the European Citizen Science Association.