

Analysis of the evolution and collaboration networks of citizen science scientific publications

**M. Pelacho, G. Ruiz, F. Sanz,
A. Tarancón & J. Clemente-Gallardo**

Scientometrics

An International Journal for all
Quantitative Aspects of the Science of
Science, Communication in Science and
Science Policy

ISSN 0138-9130

Scientometrics

DOI 10.1007/s11192-020-03724-x



Your article is published under the Creative Commons Attribution license which allows users to read, copy, distribute and make derivative works, as long as the author of the original work is cited. You may self-archive this article on your own website, an institutional repository or funder's repository and make it publicly available immediately.



Analysis of the evolution and collaboration networks of citizen science scientific publications

M. Pelacho^{1,2} · G. Ruiz³ · F. Sanz¹ · A. Tarancón^{3,4} · J. Clemente-Gallardo^{1,3,4} 

Received: 27 February 2020 / Accepted: 16 September 2020
© The Author(s) 2020

Abstract

The term *citizen science* refers to a broad set of practices developed in a growing number of areas of knowledge and characterized by the active citizen participation in some or several stages of the research process. Definitions, classifications and terminology remain open, reflecting that citizen science is an evolving phenomenon, a spectrum of practices whose classification may be useful but never unique or definitive. The aim of this article is to study citizen science publications in journals indexed by WoS, in particular how they have evolved in the last 20 years and the collaboration networks which have been created among the researchers in that time. In principle, the evolution can be analyzed, in a quantitative way, by the usual tools, such as the number of publications, authors, and impact factor of the papers, as well as the set of different research areas including *citizen science* as an object of study. But as citizen science is a transversal concept which appears in almost all scientific disciplines, this study becomes a multifaceted problem which is only partially modelled with the usual bibliometric magnitudes. It is necessary to consider new tools to parametrize a set of complementary properties. Thus, we address the study of the citizen science expansion and evolution in terms of the properties of the graphs which encode relations between scientists by studying co-authorship and the consequent networks of collaboration. This approach - not used until now in research on citizen science, as far as we know- allows us to analyze the properties of these networks through graph theory, and complement the existing quantitative research. The results obtained lead mainly to: (a) a better understanding of the current state of citizen science in the international academic system-by countries, by areas of knowledge, by interdisciplinary communities-as an increasingly legitimate expanding methodology, and (b) a greater knowledge of collaborative networks and their evolution, within and between research communities, which allows a certain margin of predictability as well as the definition of better cooperation strategies.

Keywords Citizen science · Scientometric · Bibliometrics · Complex networks

✉ J. Clemente-Gallardo
jcg@unizar.es

Extended author information available on the last page of the article

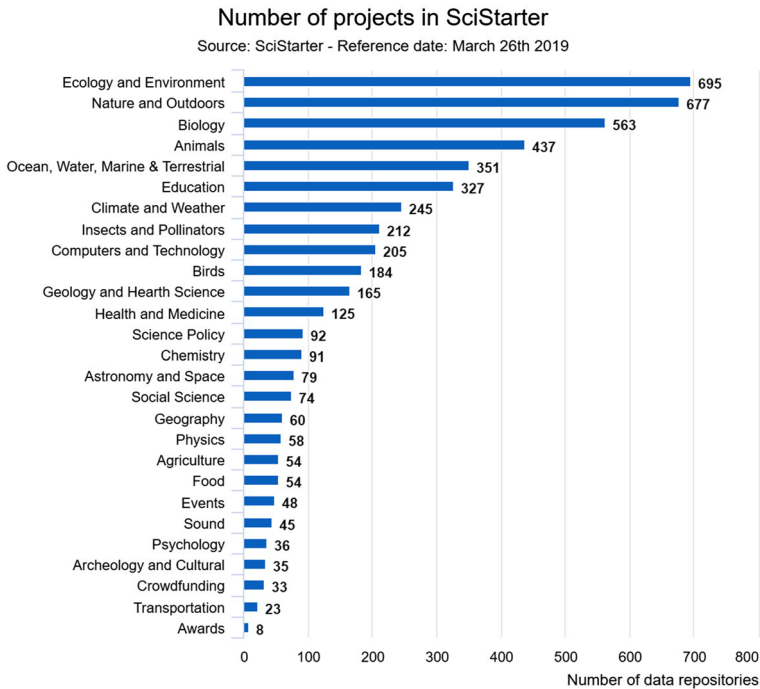


Fig. 1 European Commission (EC) 2019. Open Data: Number of projects in SciStarter. See Open Science monitor available in https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/open-science/open-science-monitor/data-open-collaboration_en

Introduction

The term *citizen science* refers to a broad set of practices developed in a growing number of areas of knowledge (see Fig. 1), characterized by the active citizen participation in one or several stages of the research process. Precise definitions, classifications and terminology remain an open problem, reflecting the fact that citizen science is an evolving phenomenon. The term was simultaneously introduced in the 1990s by Alan Irwin (Irwin 1995) and Ricky Bonney (see Bonney et al. (2009)), and in the last 25 years many other definitions and/or classifications have been proposed and discussed (see for instance, among many other references, (Wiggins and Crowston 2011; Shirk et al. 2012; Haklay 2013; Socien-tize Project 2014; Finquelievich and Fischnaller 2014; Haklay 2015; Pocock et al. 2017; Strasser et al. 2018; Ceccaroni et al. 2019; Heigl et al. 2019)). This *plethora of definitions and classifications* (Kasperowski and Kullenberg 2019) makes more appropriate to speak of a continuum (Cooper et al. 2007; Pocock et al. 2017) or a broad spectrum of citizen science practices.

The aim of this article is to analyze the evolution and collaboration networks of citizen science publications, specifically those published in WoS journals. As it has been demonstrated in previous studies (see Follett and Strezov 2015; Kullenberg and Kasperowski 2016; Bautista-Puig et al. 2019)—and will also be proved in this paper—the expansion of citizen science, as a scientific methodology, is reflected particularly in the exponential growth of citizen science publications in indexed journals for the last two decades. That

expansion is also reflected in the increasing number of research areas where citizen science is playing an active role (see Jordan et al. 2015; Eitzel et al. 2017; Mahr et al. 2018), as well as the exploding number of associations and *communities of practice* (Lave and Wenger 1991) all around the world (Storksdieck et al. 2016). Consequently, the task ahead of us represents a very complex and multifaceted problem. Hence, it is not possible to capture all the necessary information in a single magnitude as the number of publications, for instance. We need new magnitudes which are able to capture as more new dimensions as possible to describe the problem in a more global way. Therefore, the expansion and evolution of citizen science is here characterized in a quantitative and qualitative way by means of the study of co-authorship and the consequent collaborative networks among scientists, within the same community and between different research communities.

Thus, the main novelty of this study, with respect to the previous ones about citizen science publications, lies in the analysis and visualization of the co-authorship networks of those publications. Kumar (2015) has pointed out that studying co-authorship to measure research collaboration has been used since the 1960s, and more recently from the social networks perspective. He adds that the research on co-authorship networks has exponentially grown during the last decade. This approach - not used until now in research on citizen science, as far as we know - allows us to analyse the properties of the corresponding graphs, completing the existing quantitative research. For this goal we use a methodology partially similar to that previously used to study the role of Spanish co-authors networks in Economics (Molina et al. 2018a), or in the analysis of the different types of researchers collaborations (Clemente-Gallardo et al. 2019), both based on the tool *Kampal research tool*, developed by the company Kampal Data Solutions S.L. (<http://www.kampal.com>) and presented in Álvarez et al. (2015).

Our tool creates a database of researchers who have co-authored papers on citizen science and, from the database, we construct graphs where the links between researchers represent the papers they have co-authored. The graphs display the information of the database in a form which can be analyzed from several points of view. With these tools we are able to define an empirical growth law for the total scientific production of citizen science groups. Moreover, we can recognize the collaboration patterns of the different research communities and the most relevant ones from the point of view of centrality and production. In addition, we can study the collaboration structure of the different countries, identify those with a larger production and relevance for the graph, and consider the evolution of the all these properties in the last two decades. These are the major contributions of our work. The complete set of data and the project running on our platform is available as Supplementary Material of this manuscript.

It is important to notice, however, that when undertaking a study related to publications on a given concept, citizen science in this case, two questions arise at the very beginning: the first one, whether that concept is sufficiently unequivocal, and the second one, whether there are different terms to refer to it.

In fact, both the different meanings of citizen science (Cooper and Lewenstein 2016) and the use of various terms (Eitzel et al. 2017) related to this scientific methodology have been frequently discussed, and there remain international debates about the possibility and/or need to define the concept unambiguously (see Strasser et al. 2018; Heigl et al. 2019). To indicate some examples, we can refer to new definitions notably different from the many existing ones, such as that of Ceccaroni et al. (2019), which places special emphasis on the social effects that occur together with scientific impacts, or to new classifications of citizen science such as that of Strasser et al. (2018) in which the *maker* activity is introduced among five types of practices. Regarding the terms used, the transversal character of the

concept, present in very diverse scientific areas, leads to the expressions used being appropriate for a particular project but not for another, and something similar occurs regarding the different ways of alluding to the people involved: citizen scientists, participants, users, volunteers, etc. (see Eitzel et al. 2017).

We should remember, on the other hand, that one of our instrumental objectives is to create the co-authors networks of articles, from the database obtained by searching in WoS. Therefore, we must bear in mind that some studies may remain hidden, since they do not explicitly mention the use of citizen science methodologies (see Cooper et al. 2014; Theobald et al. 2015; Follett and Strezov 2015; Kullenberg and Kasperowski 2016; Turrini et al. 2018; Gadermaier et al. 2018). In this sense, Cooper et al. (2014) discovered how voluntary collaboration was not becoming visible in many scientific articles on ornithology, arguing that the same could be happening in other areas. For that reason, these authors encouraged the use of coherent terminology to facilitate the monitoring of the impact of citizen science in numerous disciplines, and in particular, urged the use of the keyword *citizen science* in the corresponding articles. Also from our experience in the Ibercivis Foundation and the BIFI , leading and promoting citizen science projects since 2006, and especially since the launch of the citizen science Observatory in Spain (<https://ciencia-ciudadana.es/>) in 2016, we are aware that there are people and collectives performing citizen science practices, but they do not always identify them as such (Pocock et al. 2017) or are not familiar with the term (Turrini et al. 2018).

For all these reasons, in order to carry out our search we have included some other expressions or labels that allow us to find articles in which the keyword *citizen science* has not been explicitly used. To define the list of these labels we have also used various classifications of citizen science activities, classifications that, like practices, are neither unique nor static. The Supplementary Material provides a summary of definitions, methodologies and classifications together with terms commonly used to refer to different activities.

In conclusion and as a result of the review of the literature as well as our experience, we will use the expression *citizen science* assuming that it is unequivocal enough to be used as a generic term that includes a wide range of activities (Jordan et al. 2015; Pocock et al. 2017). Along with this, in order to perform the search for articles and authors, we include a list of additional terms, elaborated with the help of the classifications that collect the various practices. We discuss all these issues in Sect. 2.

The paper is organized as follows. As we mentioned above, in Sect. 2 we describe the construction of the database. Firstly, we justify our choice of labels which allows us to characterize the concept of citizen science and summarize the several problems to define it in a closed form. Then, we describe the creation of the database of papers and researchers and compare it with previous approaches in the literature. Finally, we summarize the tools used to study the corresponding co-authorship networks. In Sect. 3 we present and discuss the main results of our analysis of the networks, in particular the quantitative evolution of the global production, the topological properties of the graphs, the evolution of the weight of the different countries in the corresponding communities and the role of the different areas of WoS. Finally, Sect. 4 summarizes our main conclusions and future research lines.

Materials and methods

Delineating the field: the choice of labels

As discussed in the Introduction, it is an open-ended task to define a list of labels to completely characterize the field of citizen science. Yet, in order to carry out this research, it seems necessary to define a set of terms that are as relevant as possible.

To elaborate it we have taken into account:

- The literature on citizen science (see specific references below),
- Some of the many well-known definitions and classifications (see an explicative brief analysis at the Supplementary Material), as well as
- Our own experience leading and promoting citizen science projects since 2006.

The final list of searching terms is as follows:

- Public participation in scientific research (PPSR) (from Shirk et al. (2012), Theobald et al. (2015), Haklay (2015) and <https://scistarter.org/citizen-science>)
- Civic science (from Dillon et al. (2016), Hand (2010), Haklay (2015))
- Participatory science (from Clarke (2003), Haklay (2015))
- Amateur science (from Mims (1999), Alberti (2001), Gura (2013), Haklay (2015) and <https://scistarter.org/citizen-science>)
- Crowd-sourced science (from Hand (2010), Franzoni and Sauermann (2014), Haklay (2015) and <https://scistarter.org/citizen-science>)
- Crowdsourcing science (from Wiggins (2010))
- Crowdsourcing research (from Zhao and Zhu (2014))
- Crowd science (from Hand (2010), Franzoni and Sauermann (2014) Scheliga et al. (2018))
- Collaborative science (from Societize Project (2013), Chan et al. (2015), Hess and Ostrom (2007))
- Community science (from Carr (2004), Theobald et al. (2015), Haklay (2015))
- Volunteer monitoring (from <https://scistarter.org/citizen-science> and Haklay (2013))
- Volunteer-based monitoring (from Maas et al. (1991))
- Volunteer thinking (from Societize Project (2013), Gray (2009), Haklay (2015), Yadav et al. (2018))
- Volunteer computing (from Sarmenta (2001), Andersen et al. (2005), Haklay (2015), Yadav et al. (2018))
- Participatory sensing (from Goldman et al. (2009), Haklay (2015))
- Crowdfunding science (from Ikkatai et al. (2018))
- Contributory science, considering the classifications of Bonney et al. (2009), Shirk et al. (2012), in which they proposed the category *contributory projects*, as later Haklay (2013, 2015) also did.

Creating the database

As we explain below, the Kampal platform allows to build the set of publications, extracted from WoS containing any of the labels above, and published until December 2018. As it

is discussed in next Section, the number of papers is 2645 and the number of researchers co-authoring them is 9955. The previous attempts known to us to explore the network of citizen science publications (Follett and Strezov 2015; Kullenberg and Kasperowski 2016; Bautista-Puig et al. 2019) used a similar approach but with some methodological differences:

- Follett and Strezov (2015) use just the label *citizen science* for a *Topic search* in WoS and Scopus. After a careful analysis, those references which failed to satisfy the definition of citizen science according to the Green Paper of citizen science (Socientize Project 2013) were removed from the list, which, eventually, contained 888 entries. The authors analyzed the evolution of yearly production and the classification of papers by research area among other criteria.
- Kullenberg and Kasperowski (2016) chose a different search procedure. They start by searching for *citizen science* in WoS and extracting all keywords from the original 1281 papers. By means of three "snowball searches" they obtained a total list of 1935 entries. From the analysis of the keywords, they manage to build a network of terms representing the citizen science field besides monitoring also the rate of growth of scientific production, as the previous reference. But they also extracted a list of citizen science projects from Literature, aiming to analyze the type of projects which had the greater WoS publication impact.
- Bautista-Puig et al. (2019) in their study on the relevance of citizen science in open science, analyze the impact on scientific publications but also its visibility in social media. Regarding the impact on publications, they formulate a search strategy in WoS based on an initial search of papers whose title contain the ten more frequently terms they consider relevant to citizen science. This set of publications is then extended obtaining publications by searching for other terms found in similar studies in topics, title, abstract and keywords and a list of synonyms. The total search, including the social strategy, originated 5100 documents in citizen science corresponding to the period between 1956 and 2017, containing all sorts of contributions, not only research papers.

We see that these approaches are similar to ours although the set of papers obtained is smaller (also because of the date, since our analysis contains papers until december 2018) and our list, in principle, may cover a wider range of activities, at least with respect to the first two references. Regarding the third one, it is more difficult to compare the quantitative results since they consider a database with all types of documents, not only research papers. We prefer to restrict the type of documents in order to have the ability to weight the quality of the publication, as we will discuss later. In any case, we can expect some of their qualitative conclusions regarding publications to be similar to ours. Of course, the social impact they consider can not be studied within our framework.

Nonetheless, we can try to quantify the differences of the final databases, in order to estimate the relative weight of the different labels we compiled above. We have performed a *Topic search* in WoS on June 6, 2019. Thus the numbers are slightly different to those presented in next Section, which cover up to December 2018, but the proportions must be similar in both cases. The resulting number of indexed papers for each label is indicated in Table 1.

From the table, we can conclude:

- *Citizen science* is, clearly, the dominant label in the set. All the other 17 terms combined reach just one third of the entries of the first. From that point of view, an analysis

Table 1 Table of relative appearances in the search of June 2019 of the different labels considered in our search.

Label	WoS entries	CS:Yes
Citizen science	2869	–
Public participation in scientific research	45	40
Civic science	46	5
Participatory science	52	17
Amateur science	13	2
Crowdsourced science	4	3
Crowdsourcing science	2	1
Crowdsourcing research	18	3
Crowd science	12	7
Contributory science	1	0
Collaborative science	128	4
Community science	91	9
Volunteer monitoring	68	37
Volunteer-based monitoring	25	7
Volunteer thinking	1	1
Volunteer computing	182	3
Participatory sensing	275	13
Crowdfunding science	2	0

The second column contains the number of entries as journal papers in WoS containing each label. The third column contains the number of entries containing a label and also “Citizen science”

reducing to our first term (as that of Follett and Strezov (2015)) is expected to produce similar results to the ones presented here, since the network analysis is expected to be robust under small changes. Therefore, we can trust that our list represents sufficiently well the system we aim to describe. A proper inclusion of some new search labels would mean the addition of new entries to the database. However, we expect that these entries would not significantly change most of our conclusions.

- Our list contains 9955 researchers. It is important to notice that we have not performed a filtering as that of Follet and Strezov to eliminate those papers which did not fulfilled the requirement of the Green Paper, we have only verified that the resulting list of papers is meaningful. As we saw above, there is an imprecise border between citizen science activities and other similar phenomena and very often it is extremely difficult to classify an activity as *true* citizen science . Hence, we prefer to trust the search results, as Kullenberg and Kasperowski did.

Among the limitations of our study, from the point of view of consider a comprehensive description of the complete citizen science landscape, we can consider the following:

- It may be objected that, by using WoS (or Scopus) as source of data, many other publications would be left out of the analysis, because, in fact, many citizen science activities do not produce results that can be found in scientific academic publications (Theobald et al. 2015; Follett and Strezov 2015; Pettibone et al. 2017; Turrini et al. 2018). In that respect, we want to remember that our interest in this paper is not to analyze the diverse and important impacts of citizen science, but the part that affects directly the academic or *professional* science.

- Follett and Strezov (2015) point out that there are more citizen science indexed publications than shown in WoS and Scopus. One reason - they explain referring a previous analysis Cooper et al. (2014) - is that there are many papers containing data from citizen science, but this is not always indicated, so in a search for articles of citizen science, these are not visible. Of course, our approach is not able to detect those either.
- We have not included some terms in our search list, which, in an initial search in WoS show very important contributions. Terms such as *community based research* (1122 articles in WoS in June 2019), *community based participatory research* (2690 papers), *traditional knowledge* (3114 papers) and *participatory action research* (2755 papers), which produce a very significant publication list. Therefore, apparently including them would multiply the number of indexed publications in our database by at least a factor of three and enlarge significantly the scope of our study. Nonetheless, including those publications is not straightforward, since although some papers in those lists may report very relevant citizen science activities, they contain much more false positive cases than those appearing with the terms of our search list. Hence, we have decided to “play safe” and not to include them in this study and publish a more complete analysis in the near future. The impact of the concept of citizen science as a whole, analyzing both the importance of the second set of terms and that of considering different languages, will be studied in a forthcoming publication. In any case we can expect the qualitative aspects of the total system to be similar to those presented here. Indeed, if we analyze the results of the search of the three labels above, we can see:
 - There is a negligible intersection of the set of papers of those labels and those of our list (below 1%): if we perform a search for those papers containing both sets of labels, the resulting list contains less than 30 papers (in a total of more than 10000).
 - The number of papers per year containing those labels is stable in the last 5 years, i.e., their communities of authors should be relatively stable and disconnected from the ones associated with our list.
 - The papers associated to those four labels belong to a reduced number of areas, essentially those related with public health, environmental issues and anthropology.

We can conclude from here that considering all the labels the resulting system would include a couple of new communities, with respect to ours (Fig. 4), probably with a dense internal structure, but disconnected from the rest of the graph. Therefore, we do not expect the topological aspects to change in a significant way with respect to our results in this paper.

Co-authorship networks

Use of co-authorship networks

As Kumar (2015) explains in his review of the literature on co-authorship networks, their use has been significant since the 1960's, but only recently began the analysis of the role played by co-authorship patterns in scientific literature from the point of view of social networks. From this perspective we can recall the works of Newman (2001a, b, 2003, 2006, 2010). Also, from the point of view of network analysis, the work of Barabási et al. (2002) is considered as one of the main seminal papers of the field. More recently we can consider also more groups focusing on different properties of the networks (Abbasi et al. 2011; Lemarchand 2012; Ding 2010; Biscaro and Giupponi 2014; van Eck and Waltman 2014).

In 2015 a new software platform (Kampal) was introduced to analyze the scientific production from a network theoretic perspective (see Álvarez et al. 2015; Molina et al. 2018a, b; Clemente-Gallardo et al. 2019 for a detailed description of the platform and the technical details concerning the algorithms used in its different components, which for the sake of conciseness we do not include here). It offers the possibility of studying WoS database to build the network of scientific publications satisfying certain criterion (as having researchers from a certain institution among its co-authors, or involving the work of particular researchers), analyze in detail its topological properties and represent the results in an intuitive form. Furthermore, by considering the network built at different years, it allows to study the evolution of the system in a certain time window. This is the platform used in our analysis, which will be described in detail in the next section. In our case the network will be constructed from the publications indexed in WoS which can be considered to represent the use of citizen science in any scientific discipline.

Constructing the publication network

Let us consider the set S of researchers, which in our case correspond to those who have co-authored a manuscript published in a journal indexed in WoS containing any of the terms referred above. We will assign a node of a graph to each researcher in S and define a link between two nodes if the corresponding researchers have co-authored one publication. For the sake of simplicity, we will not consider the different authoring patterns used in different research areas (alphabetical, group-role, etc). In principle, we can consider a different weight for each of these links, depending on different factors:

- We can consider a quality weight for the links by using the JCR Impact factor of the corresponding journal at that particular date i.e., the link representing a paper published in, for instance, 2014 has a weight equal to the impact factor of the corresponding journal in 2014. The link to another paper, published, say, in 2016, will have the weight equal to the impact factor of the corresponding journal in 2016. This may be a good metric if we consider a case where all the researchers belong to the same area, as in the case of a research institute or university department. But, at the same time, it is not a good metric to compare research in very different areas, since the absolute value of the impact factor of the journals of ISI areas changes significantly. For instance according to JCR 2014, an impact factor of 1.6 belongs to the first quartile of the area *Physics Mathematical*, while it is in the last quartile of the area *Biochemistry and Molecular Biology*.
- Instead, we can consider a weight given by the position of the journal in its WoS area. Thus, we could assign
 - 4 points to the paper published in a Q1 journal,
 - 3 points to the paper published in a Q2 journal,
 - 2 points for papers in Q3 journals
 - and 1 point to papers in Q4 journals.

This is a discretized version of the Normalized Journal Position (NJP) index (Costas and Bordons 2007), and has the advantage of being much better adapted to a transversal concept as citizen science. Each area is considered equally then, and the weight assigned reflects the importance of the journal in its area. We will refer to this metric as **Quartile metric**.

- An alternative, which we will also use in the following section, corresponds to the **Excellence metric** which counts the number of papers which belong to their respective first decile in the WoS area, i.e., it assigns 1 point to the papers in the top-decile journals and 0 to all the others.
- Finally, we could also use a flat metric and assign the same weight to each paper. We would be considering thus no quality filter in the type of paper.

We will see examples of these metrics in the next section.

The degree and position of each node

It is also well known from the Theory of Complex Networks that on the interconnected structure we created above, we can also introduce a mechanism to classify the nodes depending on their role in the set. In this sense, the degree of the node and its *betweenness centrality index* become two of the most relevant magnitudes. According to them, the graphical representation of the nodes is chosen. Kampal tool bases his graphical representations in Fruchterman and Reingold force-directed layout algorithm (Fruchterman and Reingold 1991) to achieve it.

Communities

If we consider graph of this type of network, we will identify subgroups of nodes with more publications together than the average around them. They are called *communities* and in this case correspond to the research groups collaborating and publishing together. They appear as denser regions in the graphs, because of the density of links connecting the nodes of the subgroup. The same will hold if we group together the nodes according to some criterion, as for instance those nodes representing researchers working in Institutions from the same country. We will see some examples of this in the following Sections, where the identification of communities of countries will prove to provide us with insightful information.

Results

In this section we will present the results of our study and discuss their meaning. We have divided it into different subsections, covering the different set of tools and the concepts modelled by them:

- First, in Sect. 3.1 we will consider the evolution of the number of publications, using the different metrics presented above.
- Then, in Sect. 3.2 we will discuss the global topological properties of the graphs associated with the complex networks defined by the set of co-authors and their papers.
- The next step is presented in Sect. 3.3, where we will consider the graphs resulting of grouping the authors by the country hosting the Institution from which the papers where signed. This will help us to understand how the collaboration between the different countries which exists now has been built in the last 20 years.
- And finally, in Sect. 3.4 we will analyze the WoS areas of the papers, and consider which are the most prolific, compare the resulting list with existing classifications of

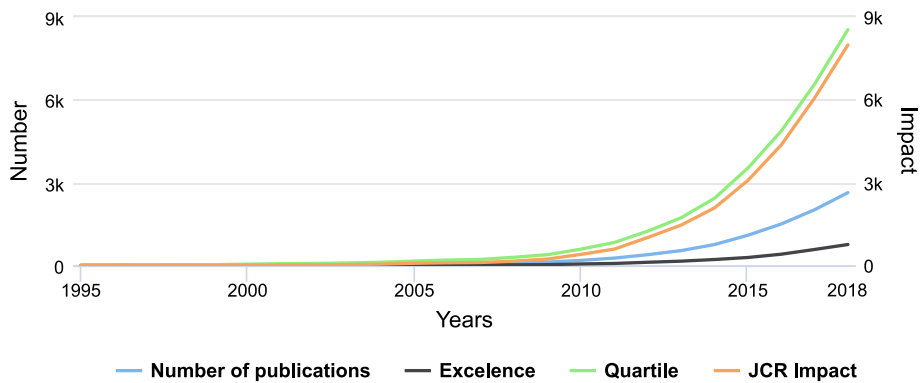


Fig. 2 Evolution of the scientific production of our search list for the different metrics considered in the period 1995–2018

citizen science projects, and study whether we can define a network of areas with a relation based on the existence of authors publishing in them.

This set of results will help us to analyze the multifaceted task under consideration from several complementary perspectives which, we believe, can define a clear description of our complex problem.

The complete project is freely accessible at http://research.kampal.com/citizen_science, where the numerical details of all these sections can be found.

Evolution of scientific production

The first step of our analysis consists in a quantitative analysis of the number of papers published in WoS on citizen science topics. For instance, if we consider the accumulated number of papers, we can observe a spectacular growth rate, particularly after 2010 (see Fig. 2 and Table 2). A similar conclusion can be extracted from the same plot considering different quality metrics, such as the JCR impact, the quartile, or the excellence metric introduced in the previous section.

From Fig. 2 and Table 2 we can conclude that:

- The average JCR impact factor of the publications is around 3, and the journals, in average, are mostly in the second quartile of the respective areas (notice that the quartile value of the line corresponding to the quartile metric which assigns a maximum of 4 points to each paper is close to 8k, and the JCR line to 7.5k, the total number of papers being slightly above 2.5k). The number of excellence papers (first decile of their respective areas) is 758. This implies that more than 1 in 4 papers is published in top journals, what is quite high for many areas.
- The growth observed in the last two decades is truly remarkable. Notice that the growth is much larger than the usual growth of WoS indexing, which is linear in time, as it can be verified in Figure 2 of Kullenberg and Kasperowski (2016). To quantify it better we can compute a fit function for the accumulated production, as it can be seen in Fig. 3. We consider all the accumulated data per year from 1995 until 2018 and determine the best linear fit for the sequence of the logarithms of the number of papers with respect

Table 2 Number of papers published before a given year containing at least one of the terms of our list

Year	Total number of papers
1995	2
1996	4
1997	5
1998	6
1999	9
2000	15
2001	20
2002	24
2003	30
2004	37
2005	50
2006	61
2007	69
2008	95
2009	121
2010	183
2011	265
2012	390
2013	537
2014	757
2015	1095
2016	1504
2017	2022
2018	2645

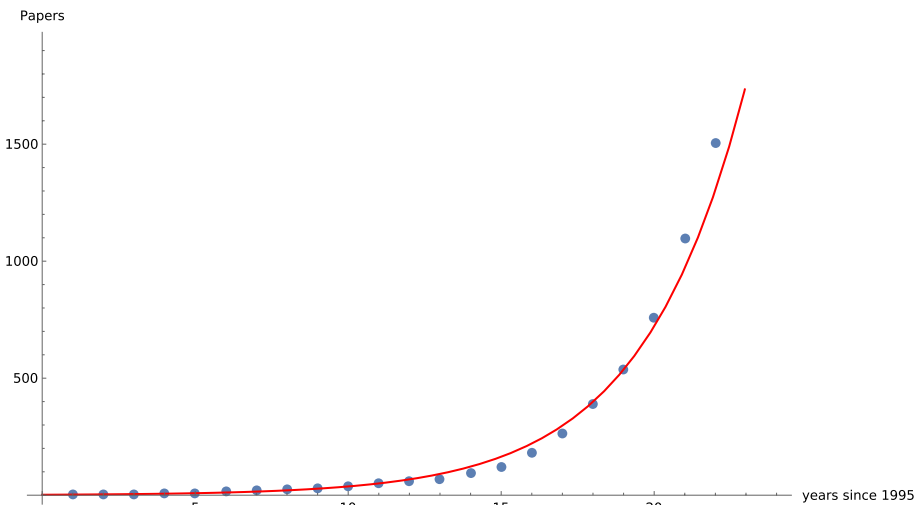


Fig. 3 Evolution of growth of total scientific production (number of papers) versus years passed since 1995 compared with the fit function $p = C \exp(\alpha x)$ with $\alpha = 0.296$ and $C = 1.935$

to the years passed since 1995. Thus, with respect to the number of papers we obtain the exponential $p = C \exp(\alpha x)$ having $\alpha = 0.296$ and $C = 1.935$ as growth rate which is particularly well suited for the last twenty years. This implies a continuous growth of more than 40%, and allows us to predict that, if no drastic changes arise, this robust and stable rhythm should continue in the upcoming years.

Topological properties of the graphs

Now we consider the plot of the graph, colouring automatically by community (Kam-pal uses different algorithms, because their performance depends on the size and in the directed-undirected character of the graph. In our case, we mainly use walktrap (Pons and Latapy 2006) or leading-eigenvector (Newman 2006). From those, we obtain Fig. 4.

We identify in the central part of the graph the largest communities with highly inter-connected researchers. We can consider those researchers as the *the most relevant* for the graph. One important characteristic of our graph is that the communities have few contacts between them, even if some of them are very large and with highly interconnected nodes. Only those communities which are very close to the center exhibit some type of heterogeneity.

We can also characterize the graph from a quantitative point of view by computing different topological indices:

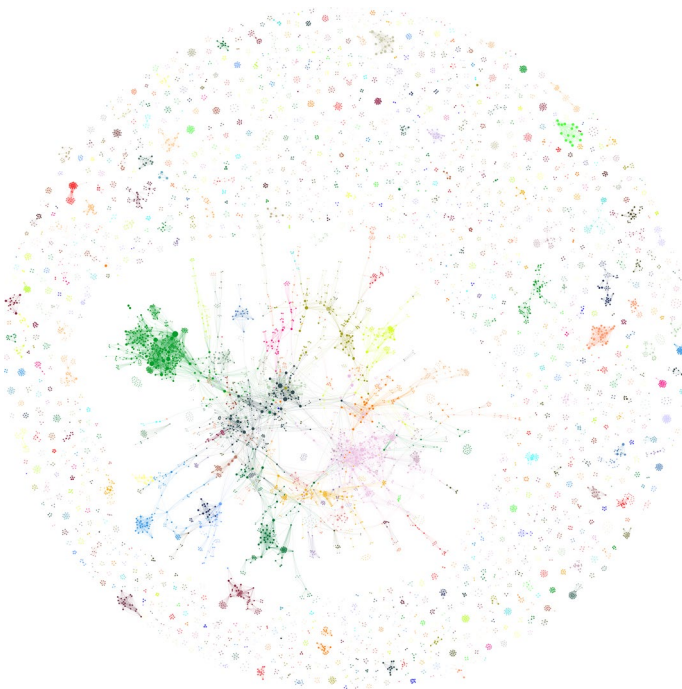


Fig. 4 Example of graph created from the one of publications. The nodes represent researchers and the links represent paper co-authored by them. The different colors represent the communities of researchers publishing together more often

- Number of researchers: 9955
- Number of papers: 2645
- Giant cluster: 2880
- Modularity: 0.967
- Assortativity: 0.6356
- Clustering: 0.7294
- Average distance: 6.14
- Longest path: 12

A few comments are in order regarding these numbers:

- (a) Very small giant cluster: The set of all nodes which are connected with each other, forms the giant cluster of the graph. In this case, the number of researchers contained in it is quite small, less than 30% of the total.
- (b) Huge modularity, i.e., the different communities cooperating together do not interact too much. This aspect also reflects the smallness of the giant cluster and the existence of a huge number of almost isolated communities.
- (c) High clustering: again, the large number of isolated communities and the type of graph considered produce a very high probability for most of the coauthors of a given author to be co-authors themselves.

We can conclude then that the set of researchers publishing these papers are working in medium sized groups, with few contacts with the other groups. This was to be expected taking into account the transversality of the citizen science notion, and hence the many different research areas considered. Nonetheless, those nodes with high centrality values represent people with a large influence in their communities, in many cases because they have addressed conceptual problems in citizen science (citizen motivation, citizen engagement, good practice, etc) which are common to most scientific areas where citizen science is being used as a tool. To some extent they are creating *citizen science theory* and therefore they transcend their original field and collaborate with researchers of different fields.

Country-based analysis

If we consider the coloring of the graph based on the country hosting the institution of each researcher, we obtain Fig. 5.

We see how:

- Most small communities correspond to researchers of the same country.
- In the most relevant communities (those in the center of the graph), this is also true for most of the cases but some communities are a bit more heterogeneous.

Both properties were to be expected in such a young discipline (before 1995 there are no references and the first years there were 2-3 papers a year) which is transversal and therefore is publishing in many different areas at the same time. Excepting the large heterogeneous communities of the center of the graph consolidated in the last years, it is difficult to find collaborations between the different groups, since even those which are geographically close may be working in very different scientific areas.

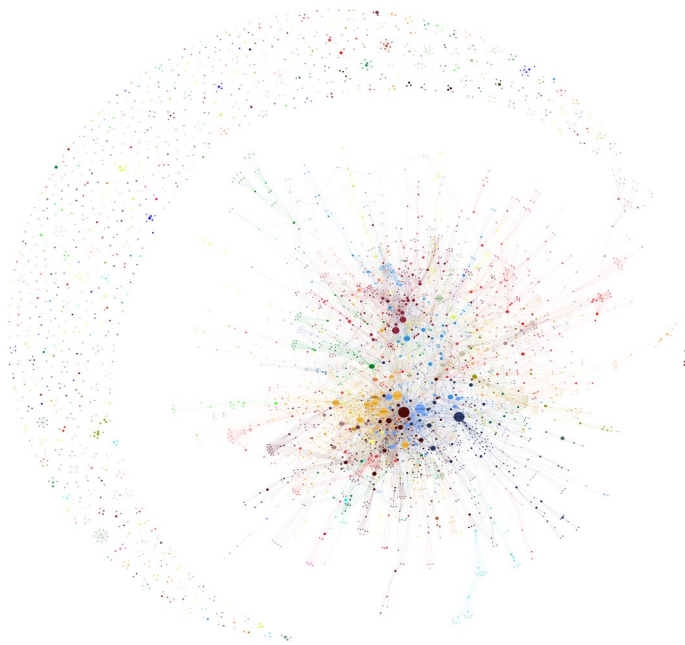


Fig. 5 Example of graph created from the one of publications grouping the researchers by Research Institution. We can appreciate a much more connected graph with respect to Fig. 4 where the nodes represent individual researchers

Furthermore, our software platform allows us to make groups of researchers attending to different criteria. For instance, we can study the relevance of the different countries, considered as the place where the research institution which hosted the researcher when he published the paper is located. All researchers signing papers in the same country are represented by a single node, whose degree is obtained combining the degrees of the individual researchers. Our graph becomes now a graph of countries instead of a graph of researchers. Again the relative diameter of the nodes and the position with respect to the center represent the degree and the centrality of each node (now a country). If we classify the countries by production, we obtain the following graph (Fig. 6) and table (Fig. 7):

The most relevant property is the huge dominance of USA and UK over the rest of the world. Far below these two, Australia takes the third position in the rank. In fourth position appears Germany, and afterwards Canada, Italy, China, France, The Netherlands and Spain. We see thus the prominent role of Anglo-Saxon countries and far below the appearance of several European countries and China. This is the picture at the end of 2018.

Had we done the same exercise in 2005, the situation would have been quite different. Indeed, the graph and rank by that date look like as Figs. 8 and 9.

USA is already in first position but the contributions are so few that the differences are not very relevant, only a factor 180 from the top (USA) to the bottom country (Philippines). Moreover, there are no collaboration between countries, all links are self-links.

Five years later, in 2010, the situation has changed, or it is in the process of changing. Indeed, the graph and rank by that date look like as Figs. 10 and 11.

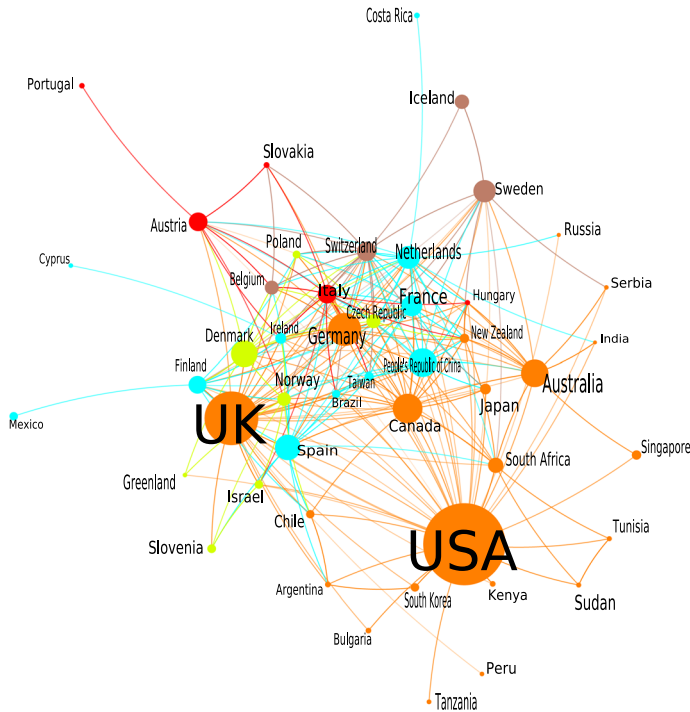


Fig. 6 Example of graph created from the one of publications grouping by countries

There are already some collaborations between countries, even if the top ones appear still as independent entities. USA and UK are already at the top, but some European countries appear in the top-10.

In 2015 the situation was already similar to the present, but with a notably less connected graph. Indeed, the graph and rank by that date look like as Figs. 12 and 13.

It is worth mentioning the appearance of China in the top-10 countries, but with a smaller centrality and relevance values than the European countries.

Collaborations are already frequent and have built communities of countries which collaborate more often, in a similar way to the analogous concept for researchers. Indeed, it is interesting to see how those communities have been built and how the different countries have been creating the relations with the others. If we do that in 2018, we find the following three communities (Figs. 14, 15, 16):

The first contains the three overall largest contributors (USA, UK and Australia) and their main collaborators (the most relevant being Canada, China, South Africa and Japan), while the second and the third gather most European countries and their main collaborators. We see thus a structure of two main poles, one organized around the Anglo-Saxon countries and the other in continental Europe, the first being much larger than the second. Hence, we can suggest an influence of political and cultural affinity in the creation and growth of these communities. We can expect the evolution to lead to a situation with a large dominant node containing the Anglo-Saxon countries and other(s) smaller node(s) corresponding to the rest of European countries with their

Outcome ?

✕

Name	Total outcome ▼	Centrality	Relevance	Contributors
USA	15 350.53	1.00	1.00	66
UK	6 747.99	0.63	0.65	57
Australia	2 296.19	0.10	0.35	42
Germany	2 119.61	0.22	0.35	50
Canada	1 746.01	0.10	0.25	44
People's Republic of China	1 521.90	0.05	0.22	38
Spain	1 225.76	0	0.20	45
France	1 205.29	0.12	0.35	47
Netherlands	1 164.65	0.10	0.36	48
Italy	1 136.67	0.10	0.26	41
Denmark	891.58	0.02	0.16	24
Sweden	800.85	0.02	0.17	38
Switzerland	774.77	0.02	0.21	42
Austria	626.19	0.02	0.16	33
Finland	534.19	0.02	0.16	41
South Africa	489.54	0.05	0.16	27
Belgium	467.78	0.02	0.17	33
Japan	388.92	0.07	0.12	16
Czech Republic	271.54	0	0.12	29
Norway	251.78	0.02	0.12	28

Fig. 7 Top-20 countries by production

main international collaborators. A deeper analysis of these aspects will be performed in a future paper, incorporating a topic-based analysis of the projects of the different countries.

Other quantitative aspects

It is well known that the different areas in WoS have very different characteristics regarding publication patterns, number of authors of each publication, etc. In this section we will analyze the role of each area in the global set of citizen science publications. In order to do that, we will extract from our publication database the following data:

- The number of papers published in the different areas of WoS,
- The number of authors publishing in those areas,
- The number of papers published by each author,
- The areas where the different authors publish papers.

Fig. 8 Example of graph created from the graph of publications grouping by countries in 2005

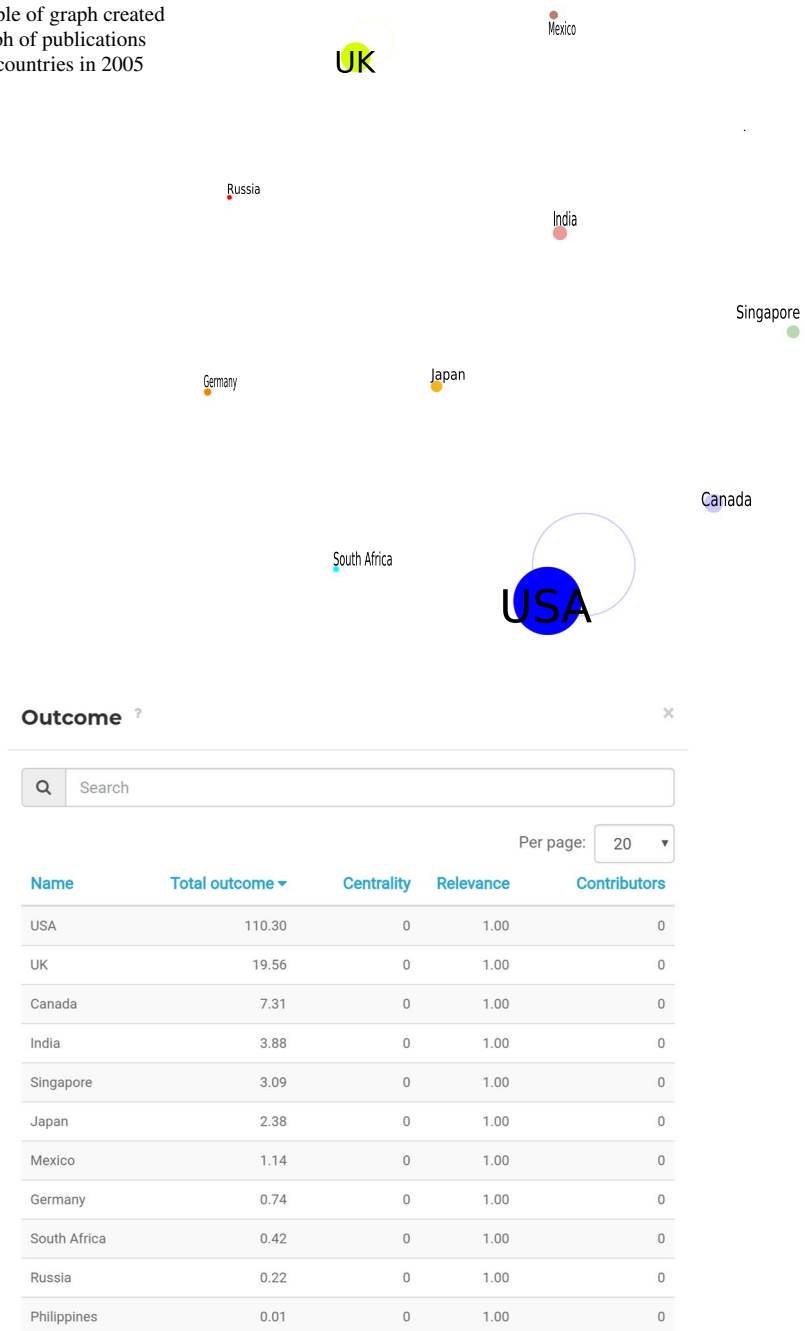
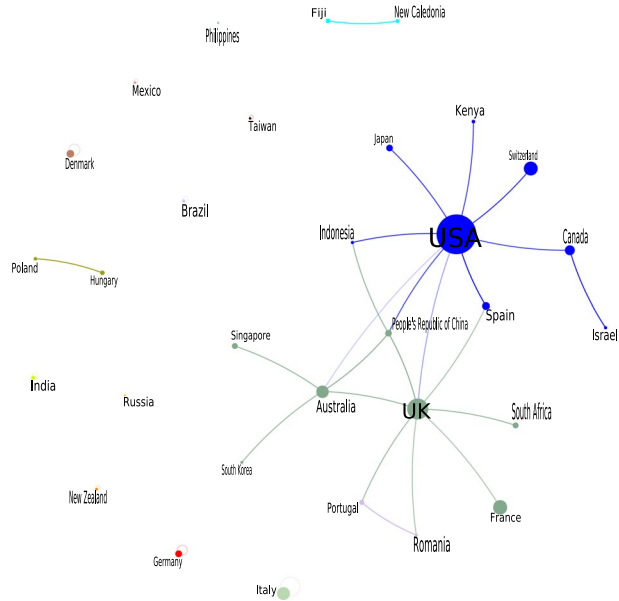


Fig. 9 Top-10 countries by production in 2005

Fig. 10 Example of graph created from the one of publications grouping by countries in 2010



Outcome ?

Search Per page: 20

Name	Total outcome ▼	Centrality	Relevance	Contributors
USA	699.57	0.75	1.00	9
UK	185.29	1.00	0.94	8
France	79.36	0	0.36	1
Switzerland	76.52	0	0.34	1
Italy	65.11	0	0.35	0
Australia	60.75	0.33	0.50	5
Canada	36.30	0.17	0.44	2
Spain	20.38	0	0.28	2
Denmark	19.88	0	0.35	0
Germany	15.25	0	0.35	0
People's Republic of China	14.62	0.09	0.35	4
Japan	14.48	0	0.33	1
South Africa	10.98	0	0.33	1
Singapore	10.84	0	0.34	1
Portugal	6.69	0.17	0.47	2
Fiji	6.20	0	0.55	1
Hungary	5.71	0	0.57	1
New Caledonia	4.13	0	0.45	1
India	3.88	0	0.35	0
Kenya	3.17	0	0.25	1

Fig. 11 Top-20 countries by production in 2010

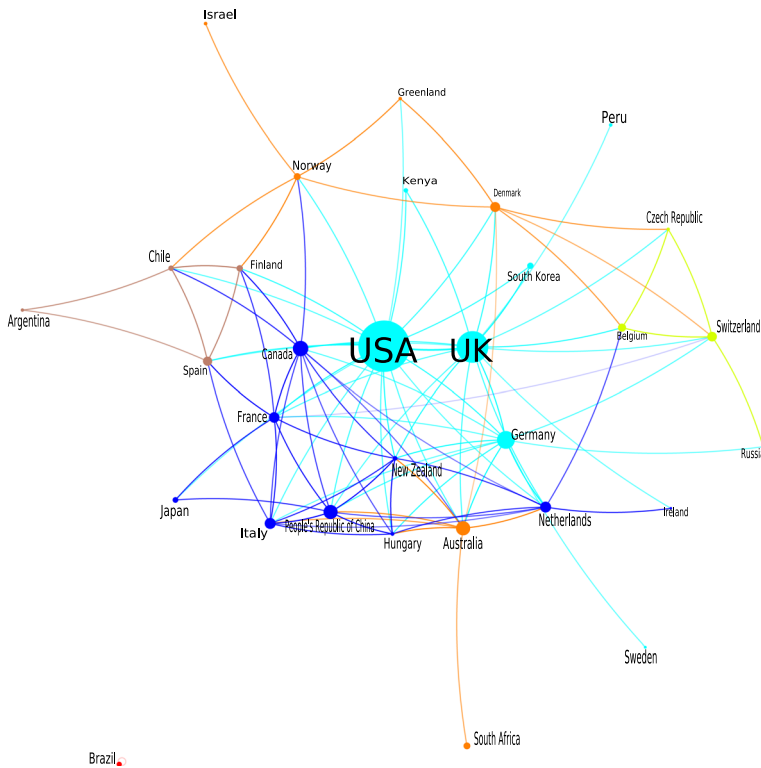


Fig. 12 Graph created from the one of publications grouping by countries in 2015

Let us consider each point separately:

Number of papers per area of knowledge

From the WoS data our platform selects the areas where WoS classifies the journal which publishes the paper. If a journal is listed in more than one area, we consider all of them. From it, we can conclude which are the most active areas publishing papers based on citizen science activities.

There are papers published in journals of 175 different areas of WoS. The complete list can be found in the Supplementary Material (database DB-Areas.xlsx), but the top-20 of them can be found in Table 3 and Fig. 17. Results are remarkable: Ecology, Environmental Sciences and Biodiversity Applications represent half of the total number of papers.

These results are to be compared with the distribution per areas of citizen science projects which we discussed above (See Fig. 1). It is clear that there is a correlation between the number of projects and publications in each area. Even if it was to be expected that many of the projects are not reflected in WoS publications, this phenomenon seems to affect all areas in a similar way. In that way, among the top areas in the Zooniverse platform (see the list of Zooniverse projects available in <https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/>

Outcome ?

Search Per page: 20

Name	Total outcome ▼	Centrality	Relevance	Contributors
USA	6 923.25	1.00	1.00	45
UK	2 776.42	0.59	0.62	39
Australia	676.45	0.18	0.38	26
Germany	653.10	0.14	0.28	30
Canada	587.69	0.07	0.20	29
People's Republic of China	448.10	0.04	0.17	18
Italy	412.30	0	0.17	26
France	405.95	0.07	0.26	32
Netherlands	323.64	0.07	0.25	27
Spain	246.90	0.07	0.25	30
Switzerland	237.23	0	0.18	29
Denmark	141.02	0.04	0.18	20
Belgium	139.93	0	0.12	12
South Africa	111.36	0	0.11	11
Sweden	105.10	0	0.11	17
Finland	95.34	0	0.11	24
Japan	92.96	0	0.10	10
New Zealand	86.95	0	0.10	13
Brazil	74.55	0	0.10	11
Norway	63.64	0	0.11	20

Fig. 13 Top-20 countries by production in 2015

[open-science/open-science-monitor/data-open-collaboration_en](#)), for instance, *Nature* and *Biology* combined have five times more projects than *Space*, while the WoS areas *Ecology*, *Environmental Sciences* and *Biodiversity conservation* contain five times the number of papers published in *Astronomy & Astrophysics*.

It is important to discuss the differences of these results with previous similar analysis such as Kullenberg and Kasperowski (2016) or Bautista-Puig et al. (2019). We can see, comparing Fig. 18 with Fig. 6 of Kullenberg and Kasperowski (2016) that there are similarities but also some differences. First of all, we must take into account that in our figure papers belonging to more than one area, are counted in all of them. This explains the remarkable difference in the total number of papers in each category, combined with the difference in years in both studies. *Ecology* and *Environmental Sciences* are in both cases the leading categories, with a remarkable distance with respect to the next one. The most significant differences come from the categories *Geography* and *History and Philosophy of Science* which appear in the top-6 in Kullenberg and Kasperowski (2016) and they are not contained in the top-20 of ours. There is also an important difference in the category *Astronomy & Astrophysics* which in our study appears in the top-5 and in the 14th position in Kullenberg and Kasperowski (2016). In the last case we consider that the time difference and the different accounting systems may explain the difference. Regarding the appearance of *History and Philosophy of Science*, if one checks the papers associated to that category contained in the Supplementary Material of that paper, there are several associated to the study of the phenomenon of public engagement, which of course is related to citizen science, but not completely equivalent. The different search mechanisms of both studies can justify this difference.

Community ?

Click on a community to see more details

Per page: 10

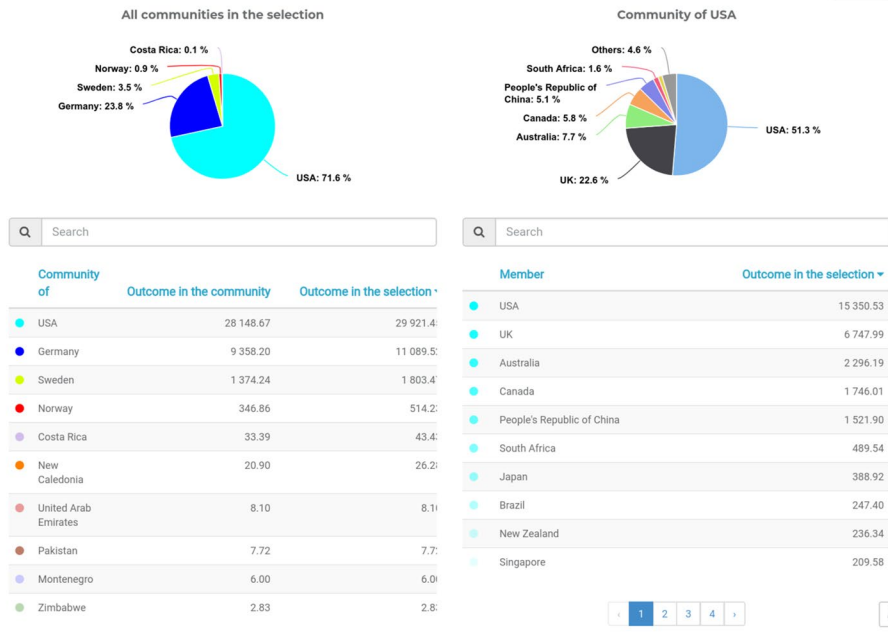


Fig. 14 Communities of countries in 2018: Top community by production

With respect to Bautista-Puig et al. (2019), there are also similarities and differences. The most important difference may be the fact that our analysis restricts to research papers and their retrieve any document type. This can introduce important differences in the relative number of documents. In what regards the results, in their Figure 4 the top-20 categories listed appear as percentages of the total and it is therefore complicated to do a direct comparison. From their number of documents, most certainly the authors chose to assign each document to only one category and that complicates the problem even more. The main difference appears to be the appearance of the category *Public, enviromental and occupational health* in the top-3, while in our list appears in 17th position. In this case, the inclusion of those labels which we did not consider, in particular *participatory action research* and similar ones included in the last paragraph of page 5 of Bautista-Puig et al. (2019) explains the different result. This issue will be addresses in a future paper.

Number of authors per area of knowledge

Another interesting result from our analysis of the database refers to the number of authors publishing in the different areas. The results are presented in Table 4 and Fig. 18, while the complete list can be found in the Supplementary Material.

Obviously, we could expect a remarkable correlation between this list and that of the previous section. Only a few changes of order between the two lists suggest that the behavior is common for both parameters. Again, the correlation of the number of

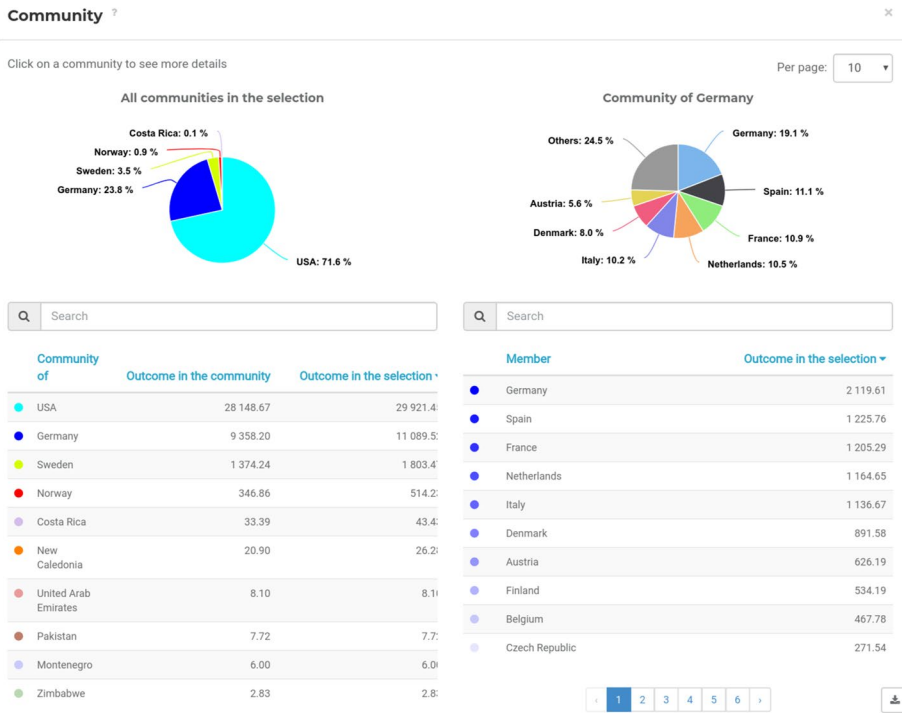


Fig. 15 Communities of countries in 2018: Second community by production

authors with the projects in the Zoouniverse platform is good, a bit worse with the projects in Scistarter (both can be compared in the list of citizen science projects available in https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/open-science/open-science-monitor/data-open-collaboration_en). Indeed, if we analyze and compare the top categories in the three lists, we see that

- The first two categories “Ecology” and “Enviromental Sciences” in Table 4 represent the 35% of the total. These categories can be considered analogous to the categories “Ecology and Environment” and “Nature and Outdoors”, of Scistarter, representing 28% and the category “Nature” of Zoouniverse, which represent the 30% of the total.
- “Biodiversity conservation” in Table 4 represents the 12% of the total while “Ocean, Water, Marine & Terrestrial” and “Insects and Pollinators” in Scistarter represents the 6.5% of the total projects and “Climate” represents the 7% of the total projects in Zoouniverse.
- “Astronomy & Astrophysics” represents the 5% of the total in Table 4 while “Astronomy and Space” contain the 0.6% of the total in Scistarter and “Space” the 9.4% in Zoouniverse.

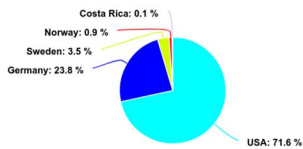
We see thus that, even if they are not identical, the correlation is quite good, being slightly better with Zoouniverse.

Community ?

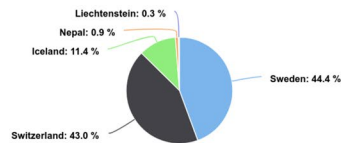
Click on a community to see more details

Per page: 10

All communities in the selection



Community of Sweden



Community of	Outcome in the community	Outcome in the selection
USA	28 148.67	29 921.4
Germany	9 358.20	11 089.5
Sweden	1 374.24	1 803.4
Norway	346.86	514.2
Costa Rica	33.39	43.4
New Caledonia	20.90	26.2
United Arab Emirates	8.10	8.1
Pakistan	7.72	7.7
Montenegro	6.00	6.0
Zimbabwe	2.83	2.8

Member	Outcome in the selection
Sweden	800.85
Switzerland	774.77
Iceland	206.30
Nepal	15.77
Liechtenstein	5.78

Fig. 16 Communities of countries in 2018: Third community by production

Table 3 Top-20 areas in WoS classified by the number of publications including citizen science activities

Areas	Number of papers
ECOLOGY	2211
ENVIRONMENTAL SCIENCES	1744
BIODIVERSITY CONSERVATION	1315
MULTIDISCIPLINARY SCIENCES	739
ASTRONOMY & ASTROPHYSICS	584
COMPUTER SCIENCE, INFORMATION SY	477
MARINE & FRESHWATER BIOLOGY	380
TELECOMMUNICATIONS	361
ZOOLOGY	360
ENVIRONMENTAL STUDIES	354
BIOTECHNOLOGY & APPLIED MICROBIO	344
GEOSCIENCES, MULTIDISCIPLINARY	315
GEOGRAPHY, PHYSICAL	305
ENGINEERING, ELECTRICAL & ELECTR	298
EVOLUTIONARY BIOLOGY	259
WATER RESOURCES	254
PUBLIC, ENVIRONMENTAL & OCCUPATI	248
GENETICS & HEREDITY	248
COMPUTER SCIENCE, THEORY & METHO	228
ENTOMOLOGY	225

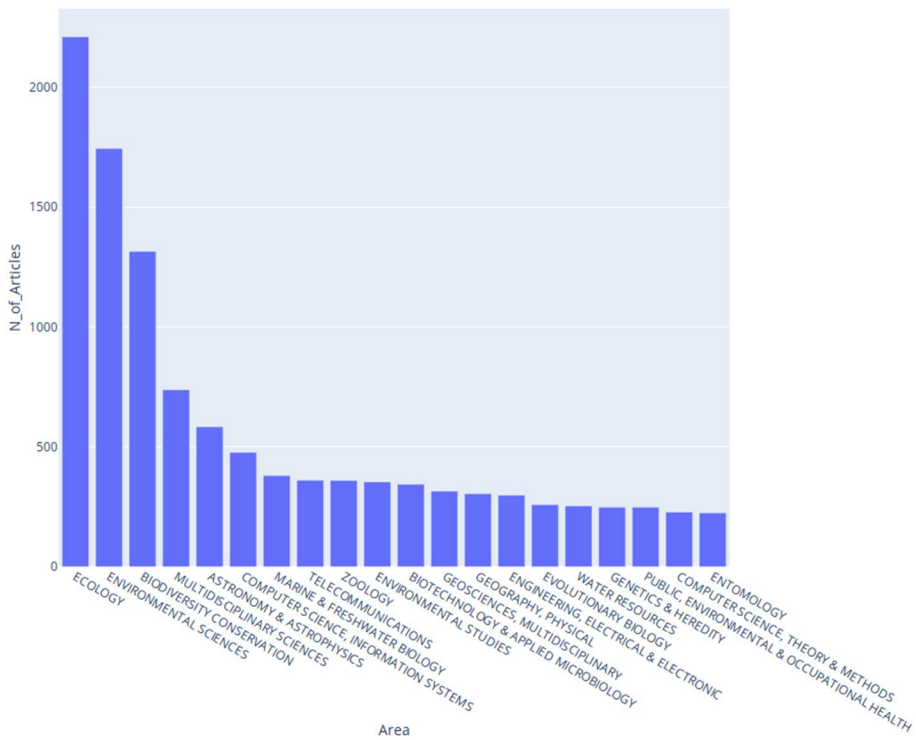


Fig. 17 Top-20 areas in WoS classified by the number of publications including citizen science activities

Number of papers per author

A final aspect corresponds to the number of publications per author in the graph. From our database it is possible to extract the number of publication co-authored by each researcher in our graph. Results are summarized in Table 5 and Fig. 19.

It is clear that most of the authors have published very few papers containing citizen science activities. Those having published more than three papers are less than 2.5% of the total. Citizen science seems to be an accessory tool for them, and not their main research framework. Most probably they correspond to the large peripheral cloud of small clusters in Fig. 2, with small-radius nodes. On the other hand, the central part of the graph contains those researchers with large nodes and larger values of centrality.

Areas where the different authors publish papers

Finally, as a partial summary of this section we can also consider the relation between the different areas which is associated with the co-authoring network. Therefore, we can construct a new graph where the nodes are the WoS areas. Two of those nodes will share a link if there are authors who have published papers in both areas. This defines a closeness criterion between areas of research from the point of view of citizen science and allows us to build a scheme of the different research areas of citizen science and

Table 4 Top-20 areas in WoS classified by the number of authors of the publications including citizen science activities

Areas	Number of authors
ECOLOGY	1777
ENVIRONMENTAL SCIENCES	1517
BIODIVERSITY CONSERVATION	1138
MULTIDISCIPLINARY SCIENCES	706
COMPUTER SCIENCE, INFORMATION SY	419
ASTRONOMY & ASTROPHYSICS	402
MARINE & FRESHWATER BIOLOGY	379
ZOOLOGY	341
ENVIRONMENTAL STUDIES	339
TELECOMMUNICATIONS	309
GEOSCIENCES, MULTIDISCIPLINARY	309
BIOTECHNOLOGY & APPLIED MICROBIO	309
GEOGRAPHY, PHYSICAL	284
ENGINEERING, ELECTRICAL & ELECTR	270
WATER RESOURCES	245
PUBLIC, ENVIRONMENTAL & OCCUPATI	239
EVOLUTIONARY BIOLOGY	236
BIOCHEMISTRY & MOLECULAR BIOLOGY	214
COMPUTER SCIENCE, THEORY & METHO	210
GENETICS & HEREDITY	207

their relations, at least from the point of view of their impact in *professional* science. The result is presented in Fig. 20.

It is particularly interesting to identify the communities of the graph, i.e., those areas with a higher density of interconnections, which appear with their nodes in the same color. Our software does this automatically. We can distinguish the following communities:

- Environmental Sciences: containing also Ecology, Biodiversity and Biology, among several others. They are represented with pink nodes.
- Geography, Geo-sciences and Photographic technology. They are represented with red nodes.
- Biotechnology, together with Biophysics, Chemistry (analytical) and Food Science, among others. They are represented as orange nodes.
- Computer Science, with Electrical & Electronic Engineering and Telecommunications among others. They are represented as blue nodes.
- Astronomy and Astrophysics, with Meteorology and Atmospheric Sciences belong to another community represented by black nodes, which also include several other areas.
- Finally, Public, Environmental and Occupational Health define, with a few other less relevant areas, the community represented by green nodes.

The detailed list of areas and the corresponding communities can be found in the Supplementary Material.

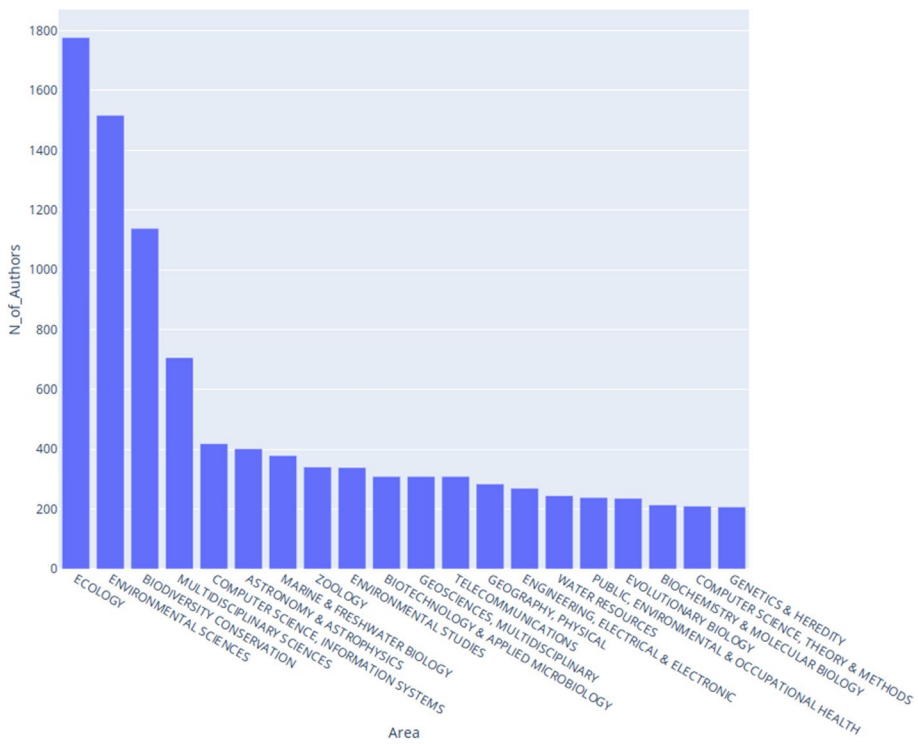


Fig. 18 Top-20 areas in WoS classified by the number of authors of the publications including citizen science activities

Conclusions and future work

In this paper we have presented an analysis of the citizen science publications in WoS, both from a quantitative and a qualitative point of view, considering many different aspects. The complexity of the problem and the large number of elements on which it depends, lead us to treat it from the point of view of the theory of complex networks, and to use a dedicated software platform (<http://www.kampal.com>Kampal). Thus, we have constructed a database of papers and authors obtained from WoS by using a list of labels representing most of the different aspects of citizen science. From that database, our platform is able to construct a network of coauthoring which can be used to study, at the same time, the quantitative aspects of scientific production and the qualitative aspects based on the collaboration patterns between the authors.

As for the researchers and papers increasing, the main conclusions are the following:

- There has been an exponential growth in the number of papers per year, with a exponent close to 0.3. If nothing changes at a global scale, we can expect this growth to continue in the following years.
- The average paper published is of high quality, with an average impact factor close to 3.
- The number of researchers publishing papers with citizen science activities is almost 10k.

Table 5 Number of papers per author with publications including citizen science activities

Authors	Number of papers
8460	1
994	2
262	3
122	4
56	5
24	6
18	7
9	8
8	9
5	10
1	11
5	12
5	13
2	14
0	15
1	16
1	17
2	19
1	24
1	30

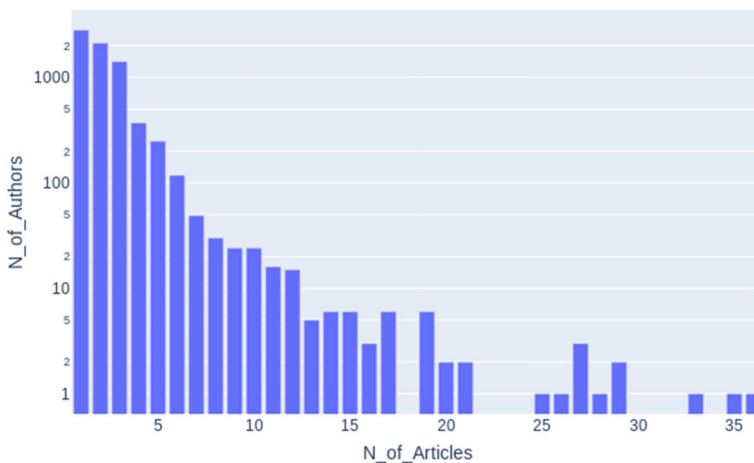


Fig. 19 Number of papers per author (in log scale) with publications including citizen science activities

We have analysed the role of citizen science as a transversal concept and its relationship with the areas of WoS. We have calculated the number of articles and researchers per area, and the number of articles per researcher. Combining this analysis with the topological properties of the graph, we conclude that:

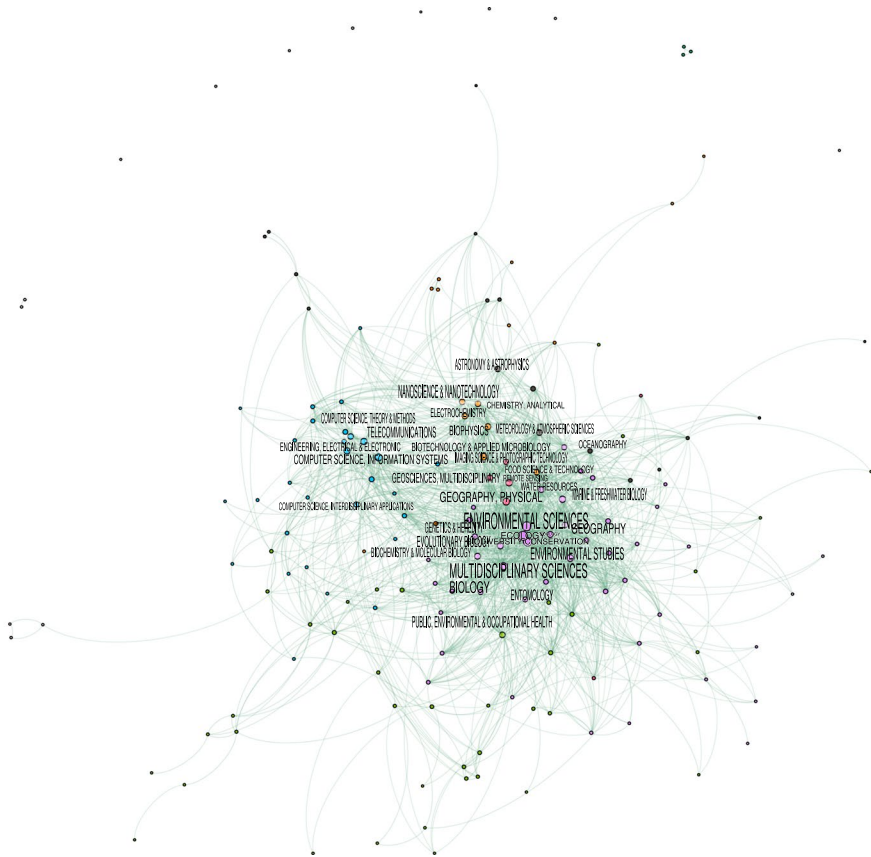


Fig. 20 Graph of the different WoS areas and their relations

- There is a minority of authors who have conducted research in many different areas. These are mostly the authors contained in the small giant cluster of the graph in Fig. 4, with high centrality index.
- Complementarily, there is a large number of researchers -most of the total- who have resorted to citizen science in their respective fields, not on a regular basis but occasionally. These authors, with very low centrality values, are represented in the cloud of nodes outside the giant cluster defining very small communities.
- The modularity is huge, suggesting that the different communities are isolated from each other, with very few contacts between them. Precisely because of that, the clustering coefficient is also very high.
- The structure, therefore, seems to reflect a big number of professional scientists considering citizen science to be an applicable methodology in their respective areas of research. At the same time, the fact that some authors, albeit a minority, have carried out citizen science activities in very different areas of study also seems to show the capacity of certain methodologies to be extended from one area to another.

Regarding the evolution of the different countries in citizen science, we have seen how a situation of isolated authors at the beginning of the century is changing with time. The evolution seems to lead to a structure with a huge dominant node, represented by the Anglo-Saxon countries (USA, UK and Australia), and one or two smaller nodes represented by the European countries, and their respective partners.

As far the label searching is concerned, the expression “citizen science” is the most relevant one for finding scientific papers related to active citizen participation in science, but the discussion on other different useful terms remains open. Some of the terms turn out to be not very relevant (e.g. crowdsourcing science), while it may be interesting to consider some others, albeit using restrictions to avoid false positives (e.g. “participatory action research”).

As we explained above, our analysis exhibits some limitations, which we expect to be able to consider in future papers:

- Consider the network associated with the keywords of papers instead of the areas of the WoS. This issue requires a deep transformation of the software platform and will be considered in the future.
- Study in detail the ratio between the authors in the different areas of the WoS, which could not be completed in this work.
- Analyse the number of authors per paper in the different areas, comparing it with the areas average values. This can give indications about the possible differences of the research groups working with citizen science tools.
- Adding certain search terms, with the necessary tools to avoid introducing a significant number of false positives.

With respect to the last point on addition of terms, its influence is expected to be restricted to particular research areas - mainly health and social sciences - with little connection to most of the articles considered in this paper, from the point of view of the co-authors network. An increase in multi-, inter- and transdisciplinary research -favoured by various ‘COST actions’ proposed by the ‘European Cooperation in Science and Technology’, or by the Sustainable Development Goals proposed by the United Nations, among other examples- may certainly lead, in the future, to a variation in those connections. Even so, considering the total set of areas and publications, we think that most of the conclusions of this research should remain valid.

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

References

- Abbasi, A., Altmann, J., & Hossain, L. (2011). Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics*, 5(4), 594–607.

- Alberti, S. (2001). Amateurs and Professionals in One County : Biology and Natural History in Late Victorian Yorkshire Amateurs and Professionals in one county : biology and natural history in late Victorian Yorkshire. *Journal of the History of Biology*, 34(1), 115–147.
- Álvarez, R., Cahué, E., Clemente-Gallardo, J., Ferrer, A., Íñiguez, D., Mellado, X., et al. (2015). Analysis of academic productivity based on complex networks. *Scientometrics*, 104(3), 651.
- Andersen, D. P., Korpela, E., & Walton, R. (2005). High-performance task distribution for volunteer computing. In: Proceedings- first international conference on e-science and grid computing, e-Science 2005, vol 2005, (pp. 196–203).
- Barabási, A. L., Jeong, H., & Néda, Z. (2002). Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, 311, 590–614.
- Bautista-Puig, N., De Filippo, D., Mauleón, E., & Sanz-Casado, E. (2019). Scientific landscape of citizen science publications: dynamics, content and presence in social media. *Publications*, 7(1), 12.
- Biscaro, C., & Giupponi, C. (2014). Co-authorship and bibliographic coupling network effects on citations. *PLoS ONE*, 9(6), e99502.
- Bonney, R., Ballard, H., Jordan, R., McCallie, E., Phillips, T., Shirk, J. and Wilderman, C.C. (2009). Public participation in scientific research : defining the field and science education. Technical Report July. <https://www.informalscience.org/public-participation-scientific-research-defining-field-and-assessing-its-potential-informal-science>.
- Carr, A. J. L. (2004). Why do we all need community science? *Society and Natural Resources*, 17(9), 841–849.
- Ceccaroni, L., Bowser, A. & Brenton, P. (2019). Civic education and citizen science. In *Civic engagement and politics*. IGI Global. ISBN 9781522509622, pp. 1–23.
- Chan, L., Okune, A. & Sambuli, N. (2015). What is open and collaborative science and what roles could it play in development? In Albagli S, Maciel ML and Abdo AH (eds.) *Open Science, open issues*.
- Clarke, C.K. (2003). Space exploration advocacy in the 21st century : the case for participatory science submitted in partial fulfillment of the requirements of the university of North Dakota for the Degree of Master of Science. PhD Thesis, PhD Thesis University of North Dakota.
- Clemente-Gallardo, J., Ferrer, A., Íñiguez, D., Rivero, A., Ruiz, G., & Tarancón, A. (2019). Do researchers collaborate in a similar way to publish and to develop projects? *Journal of Informetrics*, 13(1), 64–77.
- Cooper, C. & Lewenstein, B. (2016). Two meanings of citizen science. In *The rightful place of science: Citizen science*. pp. 51–62.
- Cooper, C. B., Dickinson, J., Phillips, T., & Bonney, R. (2007). Citizen science as a tool for conservation in residential ecosystems. *Ecology and Society*, 12(2), art11.
- Cooper, C. B., Shirk, J., & Zuckerberg, B. (2014). The invisible prevalence of citizen science in global research: Migratory birds and climate change. *PLoS ONE*, 9(9), e106508.
- Costas, R., & Bordons, M. (2007). The h-index: Advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics*, 1(3), 193–203.
- Dillon, J., Stevenson, R. B., & Wals, A. E. (2016). Introduction to the special section moving from citizen to civic science to address wicked conservation problems. *Conservation Biology*, 30(3), 450–455.
- Ding, Y. (2010). Applying weighted pagerank to author citation networks. *Journal of the American Society for Information Science*, 62(2), 236–245.
- Eitzel, M. V., Cappadonna, J. L., Santos-Lang, C., Duerr, R. E., Virapongse, A., West, S. E., et al. (2017). Citizen science terminology matters: Exploring key terms. *Citizen Science: Theory and Practice*, 2(1), 1.
- Finquelievich, S., & Fischnaller, C. (2014). Ciencia ciudadana en la Sociedad de la Información: Nuevas tendencias a nivel mundial. *Revista Iberoamericana de Ciencia, Tecnología y Sociedad - CTS*, 9(27), 11–31.
- Follett, R., & Strezov, V. (2015). An analysis of citizen science based research: usage and publication patterns. *PLoS ONE*, 10(11), e0143687.
- Franzoni, C., & Sauermann, H. (2014). Crowd science: The organization of scientific research in open collaborative projects. *Research Policy*, 43(1), 1–20.
- Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, 21(11), 1129–1164.
- Gadermaier, G., Dörler, D., Heigl, F., Mayr, S., Rüdiger, J., Brodschneider, R., et al. (2018). Peer-reviewed publishing of results from citizen science projects. *Journal of Science Communication*, 17(3), 1–5.
- Goldman, J., Shilton, K., Burke, J., Estrin, D., Hansen, M., Ramanathan, N., et al. (2009). Participatory sensing: A citizen-powered approach to illuminating the patterns that shape our world. Woodrow Wilson International Center for Scholars: Technical Report May.
- Gray, F. (2009). The age of citizen cyberscience. Technical report, CERN. <https://cerncourier.com/a/viewpoint-the-age-of-citizen-cyberscience/>.

- Gura, T. (2013). Amateur experts. *Nature*, 496, 259–261.
- Haklay, M. (2013). Citizen science and volunteered geographic information: overview and typology of participation. In D. Sui, S. Elwood, & M. Goodchild (Eds.), *Crowdsourcing Geographic Knowledge*. Dordrecht: Springer.
- Haklay, M. (2015). Citizen science and policy: a European perspective by. Wilson Center-Commons Lab: Technical report.
- Hand, E. (2010). Volunteer army catches interstellar dust grains. *Nature*, 466, 685–687.
- Heigl, F., Kieslinger, B., Paul, K. T., Uhlik, J., & Dörler, D. (2019). Opinion: Toward an international definition of citizen science. *Proceedings of the National Academy of Sciences*, 116(17), 8089–8092.
- Hess, C., & Ostrom, E. (2007). *Understanding knowledge as a commons: From theory to practice*. Cambridge: MIT Press.
- Ikkatai, Y., McKay, E., & Yokoyama, H. M. (2018). Science created by crowds: A case study of science crowdfunding in Japan. *Journal of Science Communication*, 17(3), 1–14.
- Irwin, A. (1995). *Citizen science: A study of people, expertise and sustainable development.environment and society*. London: Routledge.
- Jordan, R., Crall, A., Gray, S., Phillips, T., & Mellor, D. (2015). Citizen science as a distinct field of inquiry. *BioScience*, 65(2), 208–211.
- Kasperowski, D., & Kullenberg, C. (2019). The many modes of citizen science. *Science & Technology Studies*, 32(2), 2–7.
- Kullenberg, C., & Kasperowski, D. (2016). What is citizen science?-A scientometric meta-analysis. *PLoS ONE*, 11(1), e0147152.
- Kumar, S. (2015). Co-authorship networks: A review of the literature. *Aslib Journal of Information Management*, 67(1), 55–73.
- Lave, J., & Wenger, E. (1991). *Situated learning: legitimate peripheral participation. learning in: doing social, cognitive and computational perspectives*. Cambridge: Cambridge University Press.
- Lemarchand, G. A. (2012). The long-term dynamics of co-authorship scientific networks: Iberoamerican countries (1973–2010). *Research Policy*, 41(2), 291–305.
- Maas, R. P., Kucken, D. J., & Gregutt, P. F. (1991). Developing a rigorous water quality database through a volunteer monitoring network. *Lake and Reservoir Management*, 7(1), 123–126.
- Mahr, D., Göbel, C., A. I. and K. V. (2018). Watching or being watched: Enhancing productive discussion between the citizen sciences, the social sciences and the humanities. In: *Citizen science: innovation in open science, society and policy*. UCL Press.
- Mims, F. M. I. I. I. (1999). Amateur science-strong tradition, bright future. *Science*, 284, 2–3.
- Molina, J. A., Ferrer, A., Iñiguez, D., Rivero, A., Ruiz, G., & Tarancón, A. (2018a). Network analysis to measure academic performance in economics. *Empirical Economics*. <https://doi.org/10.1007/s00181-018-1546-0>.
- Molina, J. A., Iñiguez, D., Ruiz, G., & Tarancón, A. (2018b). The nobel prize in economics: Individual or collective merits? *Boston college working papers in economics 966*. Boston College Department of Economics.
- Newman, M. E. (2001a). The structure of scientific collaboration networks. *PNAS*, 98(2), 404–409.
- Newman, M. E. J. (2001b). Scientific collaboration networks, II: Shortest paths, weighted networks, and centrality. *Physical Review E*, 64(1), 16132.
- Newman, M. E. J. (2003). Structure and function of complex networks. *SIAM Review*, 15(3), 247–262.
- Newman, M. E. J. (2006). Finding community structure in networks using the eigenvectors of matrices. *Physical Review E*, 74(3), 36104.
- Newman, M. E. J. (2010). *Networks: An introduction*. Oxford: Oxford University Press.
- Pettibone, L., Vohland, K., & Ziegler, D. (2017). Understanding the (inter)disciplinary and institutional diversity of citizen science: A survey of current practice in Germany and Austria. *PLoS ONE*, 12(6), e0178778.
- Pocock, M. J. O., Tweddle, J. C., Savage, J., Robinson, L. D., & Roy, H. E. (2017). The diversity and evolution of ecological and environmental citizen science. *PLoS ONE*, 12(4), e0172579.
- Pons, P., & Latapy, M. (2006). Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications*, 10(2), 191–218.
- Sarmenta, L.F.G. (2001). Volunteer computing. PhD Thesis, MIT PhD Thesis.
- Scheliga, K., Friesike, S., Puschmann, C., & Fecher, B. (2018). Setting up crowd science projects. *Public Understanding of Science*, 27(5), 515–534.
- Shirk, J. L., Ballard, H. L., Wilderman, C. C., Phillips, T., Wiggins, A., Jordan, R., et al. (2012). Public participation in scientific research: A framework for deliberate design. *Ecology and Society*, 17(2), art29.

- Socientize Project (2013) Green paper on citizen science. citizen science for Europe: Towards a society of empowered citizens and enhanced research. Technical report.
- Socientize Project. (2014). White paper on citizen science for Europe. Retrieved from:<http://www.socientize.eu/?q=eu/content/white-paper-citizen-science>.
- Storksdieck, M., Shirk, J. L., Cappadonna, J. L., Domroese, M., Göbel, C., Haklay, M., et al. (2016). Associations for citizen science: Regional knowledge, global collaboration. *Citizen Science: Theory and Practice*, 1(2), 1–10.
- Strasser, B.J., Baudry, J., Mahr, D., Sanchez, G. and Tancoigne, E. (2018). “Citizen Science”? rethinking science and public participation. *Science & Technology Studies* (X): 52–76.
- Theobald, E. J., Ettinger, A. K., Burgess, H. K., DeBey, L. B., Schmidt, N. R., Froehlich, H. E., et al. (2015). Global change and local solutions: Tapping the unrealized potential of citizen science for biodiversity research. *Biological Conservation*, 181, 236–244.
- Turrini, T., Dörler, D., Richter, A., Heigl, F., & Bonn, A. (2018). The threefold potential of environmental citizen science-Generating knowledge, creating learning opportunities and enabling civic participation. *Biological Conservation*, 225, 176–186.
- van Eck, N.J. and Waltman, L. (2014) Visualizing bibliometric networks. In *Measuring scholarly impact*. Cham: Springer International Publishing.
- Wiggins, A. (2010). Crowdsourcing science. In: *Proceedings of the 16th ACM international conference on Supporting group work - GROUP '10*. New York, New York, USA: ACM Press. ISBN 9781450303873, p. 337.
- Wiggins, A. and Crowston, K. (2011). From conservation to crowdsourcing: A typology of citizen science. In *2011 44th Hawaii international conference on system sciences*. IEEE. ISBN 978-1-4244-9618-1, pp. 1–10.
- Yadav, P., Charalampidis, I., Cohen, J., Darlington, J., & Grey, F. (2018). A collaborative citizen science platform for real-time volunteer computing and games. *IEEE Transactions on Computational Social Systems*, 5(1), 9–19.
- Zhao, Y., & Zhu, Q. (2014). Evaluation on crowdsourcing research: Current status and future direction. *Information Systems Frontiers*, 16(3), 417–434. <https://doi.org/10.1007/s10796-012-9350-4>.

Affiliations

M. Pelacho^{1,2} · G. Ruiz³ · F. Sanz¹ · A. Tarancón^{3,4} · J. Clemente-Gallardo^{1,3,4} 

¹ Fundación Ibercivis, Edificio I+D-Campus Río Ebro, Universidad de Zaragoza, C/ Mariano Esquillor s/n, 50018 Zaragoza, Spain

² Departamento de Lógica y Filosofía de la Ciencia, Universidad del País Vasco UPV/EHU, Av. Tolosa 70, 20018 Donostia-San Sebastián, Spain

³ Instituto de Biocomputación y Física de los Sistemas Complejos (BIFI), Edificio I+D-Campus Río Ebro, Universidad de Zaragoza, C/ Mariano Esquillor s/n, 50018 Zaragoza, Spain

⁴ Departamento de Física Teórica, Universidad de Zaragoza, Campus San Francisco, 50009 Zaragoza, Spain