NEWSERA - Citizen Science as the new paradigm for Science Communication

Deliverable 2.3

Effectiveness of science communication in Citizen Science projects

Revision: v1.0

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 87312

NEWSERA
#CITSCI IS THE NEW #SCICOMM

DELIVERABLE DETAILS

Due date: March 31th, 2021

Actual submission date: April 9th, 2021

Project start date: January 1st, 2020

Duration: 36 months

Work Package concerned: WP2

Concerned work package leader: UNIPD

Deliverable leader: Ibercivis

Dissemination level:

PU: Public (must be available on the website) ☒

CO: Confidential, only for members of the consortium (including the Commission Services) ☐

Cl: Classified, as referred to in Commission Decision 2001/844/EC ☐

Authors:

Francisco Sanz-García, Maite Pelacho (Ibercivis).

Alfredo Ferrer (Kampal).

Revision history:

revision	date	Contributor	Description
v0.1	31.03.2021	Francisco Sanz, Maite Pelacho (Ibercivis). Alfredo Ferrer (Kampal S.L.).	First Draft and final version
v1.0	09.04.2021	Rosa Arias (Science for Change).	Final revision

STATEMENT OF ORIGINALITY

This deliverable contains original unpublished work except where clearly indicated otherwise.

Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

SUMMARY

NEWSERA analyses and evaluates the complex and multidirectional science communication strategies, including digital and non-digital ones, addressed to quadruple helix stakeholders in citizen science projects across Europe as the new paradigm for science communication.

The overall aim of NEWSERA is to demonstrate the virtues of citizen science as an inclusive, broad and powerful science communication mechanism that can allow to increase trust in science communication and, in turn, in science at large, while opening up science and innovation to the whole of society, raising awareness and educating in science, and reducing the chances of incurring in fake news by promoting critical thinking.

This deliverable corresponds to results obtained from Task T2.3 on **Effectiveness and perception of science communication in EU citizen science projects**. The development of this task consisted in carrying out an initial assessment of the communication channels and strategies used in citizen science (CS) projects, as well as the quality and effectiveness of science communication. In addition, the main barriers and challenges for quadruple helix stakeholders were also co-identified through co-creative methodologies.

In T2.3, and after defining preliminary indicators in T2.2, the quality and effectiveness of science communication[1] and its influence in the perception of science has been evaluated in **31 ongoing European CS projects**, which are developed in Italy, Portugal and Spain and were co-identified in T2.1 through the analysis of the European survey on "A portrait of Science Communication Strategies by Citizen Science projects". The survey responses (154 validated replies) and subsequent interviews, with project representatives who indicated their interest in participating in the NEWSERA #CitSciComm Labs, provided general results on science communication in a significant set of ongoing EU CS projects. The #CitSciComm Labs, described in Deliverable 3.1, completed those preliminary results specifically related to the 31 selected pilots.

Using the co-defined indicators - and establishing new ones from the pilot projects analysis - allows us to carry out a preliminary assessment of the current status of science communication within the 31 CS pilots.

[1]In this report we use the term *communication* in a broad sense most of times, although we agree with the distinction made by Rüfenacht et al. 2021 between communication (that which occurs throughout the research process) and dissemination (that which refers to research results). When necessary we distinguish both terms.

We have used and adapted an analytical service and visualization tool, Kampal Social, strongly useful for multiple-variable analysis. Kampal applies different techniques based on complex systems to:

(1) **identify the communities of CS** (size, influencers, quadruple helix stakeholders subcommunities);

(2) **asses the assortativity** (preference for a network's nodes, in this case participants, to attach to others that are similar) **between citizen and professional scientists**;

(3) **evaluate the strength of a community** (cohesion, relevance and distance between community members);

(4) **perform the sentiment analysis**. The results are visualized with charts, maps and evolution graphs; and

(5) identify anti-communities (haters or trolls against CS).

Results obtained through Kampal will also contribute to the selection of the projects to be chosen as NEWSERA pilots in T2.4.

TABLE OF CONTENTS

1. Introduction

NEWSERA analyses and evaluates the complex and multidirectional science communication strategies, including digital and non-digital ones, addressed to quadruple helix stakeholders in citizen science (CS) projects across Europe as the new paradigm for science communication. This report presents the preliminary results on effectiveness and perception of science communication in European CS projects.

In section 2, we briefly present the current state of effectiveness of science communication in CS projects. We mainly review the most commonly used communication strategies and channels, as well as the main barriers and challenges faced by projects. We find elements shared with professional research along with others specific to CS.

Section 3 is devoted to the methodology we used for assessing science communication in CS projects. First, we describe the collection of data on science communication in European CS, and in particular in the projects participating in NEWSERA. We then briefly explain the Kampal Social tool and how we have used it to analyse project data on science communication, as well as its evolution.

Then, in section 4, we report the results on the evaluation of the science communication of the NEWSERA pilot projects. We present - in more or less detail, depending on the information available and provided by the projects - some different analyses for the Italian, Portuguese and Spanish pilots, 31 in all. These analyses allow us to show the communication networks of the projects and their evolution, as well as to discover some strengths and weaknesses that can be used as indicators for better communication strategies.

Among our conclusions, we highlight the need for the pilots to incorporate as soon as possible the recommendations made during the co-creation processes within the NEWSERA Labs. In this way, it will be possible to check the improvement of scientific communication in CS projects, as well as the usefulness of the co-identified indicators.

2. Citizen science communication: channels, strategies, barriers

Over the last decades CS methodologies have been incorporated in an increasing number of research fields in order to achieve its objectives more effectively. In some projects, basic scientific goals seem to have more weight, while in others emphasis is put on social, environmental, policy and educational goals. Clearly, within the same project different objectives co-exist and overlap. In any case, CS is proving remarkably successful in advancing scientific knowledge, allowing more and more people to experience the research process (Bonney et al. 2009). Moreover, it is recognised not only as an approach to research, but also to science communication (Göbel et al. 2019, Vohland et al. 2021).

2.1 Quality and effectiveness of science communication in citizen science

The general situation regarding effectiveness of science communication and CS can be summarised as follows. A wider revision of state of the art is presented within deliverable *2.1 Portrait of Citizen Science Communication Strategies in EU Citizen Science Projects*.

- Different societies face diverse problems, some more specific and/or localised (e.g. rare diseases, endemic diseases, proximity to pollution hotspots), others more general and shared (e.g. natural disasters of various kinds, climate change, cancer). Thousands of research lines are developed - to a greater or lesser extent, however - around the world.
- In some cases, research transfer to the various social fields is not successful. Particularly, many times communication and dissemination leading to action, in one or another way, seems to be neither sufficient nor effective.
- Participatory approaches and *co-production of knowledge and society* (Jassanof 2003) have been proposed in the last decades as necessary means to address the situation, seeking to abandon views such as *scientists in their ivory tower* or, later, the *cognitive deficit model* according to which lay people resist certain research and/or innovations because they do not understand them (Felt, Wynne, Callon et al. 2007).
- Even so, science communication is still too unidirectional (Rüfenacht et al. 2021), that is, from the scientific community to the rest of the stakeholders. Moreover, the different stakeholders' profiles are often not properly considered. Nevertheless, there is a lot of knowledge and expertise on the communities themselves (e.g. farmers' knowledge is key when thinking about climate change, neighbours have the knowledge of their own environment).

- It is here where communication power of CS must be highlighted, although undoubtedly there are also failures of communication. Even so, CS always implies people outside academia *doing research* and *science communication* in different ways, and this is what makes the difference.

Among other emblematic examples illustrating CS as a powerful tool for science communication, we can consider CS projects related to climate change. Two different initiatives, BAYSICS in Germany and Red4C in Spain, clearly refer to CS as a means for achieving better science communication, which is essential for the necessary environmental, social, policy and educational changes.

The first project, BAYSICS (BAYSICS, n.d.), poses research and understanding together with acting on climate change. Their motto is: *"Convey knowledge - promote perception - communicate complexity".*

> *Far-reaching consequences of climate change make extensive adaptation and climate protection measures necessary in Bavaria. These can only be successfully put into practice in democratic societies if they are seen by citizens as legitimate, acceptable and feasible. Under the motto "Convey knowledge - promote perception - communicate complexity" selected target groups (e.g. recreation seekers, pollen allergy sufferers, citizens interested in nature, pupils) are addressed in a Citizen Science approach via the innovative BAYSICS portal. Thus, climate change is made tangible through their own observations in their concrete environment.*

The second example, *Red4C: A network to tackle climate change through citizen science (Red4C n.d.),* starts from the following situation:

> *"Climate change (CC) is the main challenge currently facing humanity. However, the perception of it by Spanish society does not seem to be adequate, since the dissemination of information on CC is not generating coherent and rational changes in people's behavior, i.e., in their lifestyle".*

Therefore, the main objective of Red4C is to create a national network of entities and organisations working in the field of CS and/or CC. This network is elaborating the *Red4C Guide for the monitoring and evaluation of Climate Change in biological systems through Citizen Science,* aimed at those interested in addressing CC mitigation and adaptation through citizen participation. The findings and tools provided by Red4C will serve as a basis for incorporating CS into the policy framework of public participation and governance programmes.

Thus, **CS is starting to be understood as the new paradigm for science communication**, which is the core of the NEWSERA project. Improving the quality and effectiveness of science communication in CS projects is one of our main objectives. The co-creation approach allows all of us to better understand the issue in order to achieve our goals, which are shared with those all involved in CS and/or science communication. #CitSciComm Labs constitute the main tool to share knowledge and experience with representatives of CS projects and

quadruple helix stakeholders. During the Labs, different *design thinking* methodologies (accurately described in deliverable 3.1) are addressed around the communication needs of each one of the 31 participant CS projects.

The NEWSERA #CitSciComm Labs are based on a bottom-up approach to test the concepts of CS communication and CS journalism by involving the targeted stakeholders in a co-creation process. NEWSERA seeks an inclusive process in which all the involved stakeholders have a voice. In this way, from a multidirectional communication, mutual learnings can be achieved and communication strategies co-created and improved.

In order to define specific actions and indicators for each one of the participant projects, the first #CitSciComm Lab activity mixes three different methodologies:

- SWOT Analysis: Strengths, Weaknesses, Opportunities and Threats.
- Utopian thinking: Wishes, in addition to the SWOT analysis.
- Six thinking hats: Six different colors for six different perspectives.
    - White: Facts. Information and data. Neutral and objective.
    - Red: Feelings/emotions. Intuition, hunches, gut instinct.
    - Yellow: Benefits. Positives, plus points. Why an idea is useful.
    - Blue: Planning/processes. Organizing and planning.
    - Green: Ideas/creativity. Alternatives, possibilities and solutions.
    - Black: Judgement. Cautions, difficulties, weaknesses, dangers and risks.

Diagnoses are clearly necessary, and it is equally needed to be able to implement concrete communication actions as well as assessing their effectiveness, keeping in mind the diverse stakeholders profiles.

2.2 Channels and strategies

To adequately address how to improve and evaluate communication actions in CS projects, we consider very briefly the evolution and current state of communication channels and strategies.

More participatory governance models have given rise in the last decade, also in science, to the need for two-way communication beyond the more traditional one-way modes (see Fig. 1) (Rüfenacht et al. 2021). Moreover, not only dissemination - about research results - but also communication during the whole research process is essential for project success (Rüfenacht et al. 2021).

NEWSERA proposes also such a dialogue among all stakeholders - according to the quadruple helix model plus data journalism - in its own methodologies.

Therefore, improving communication in CS projects also requires bidirectional, or better, a multidirectional communication.

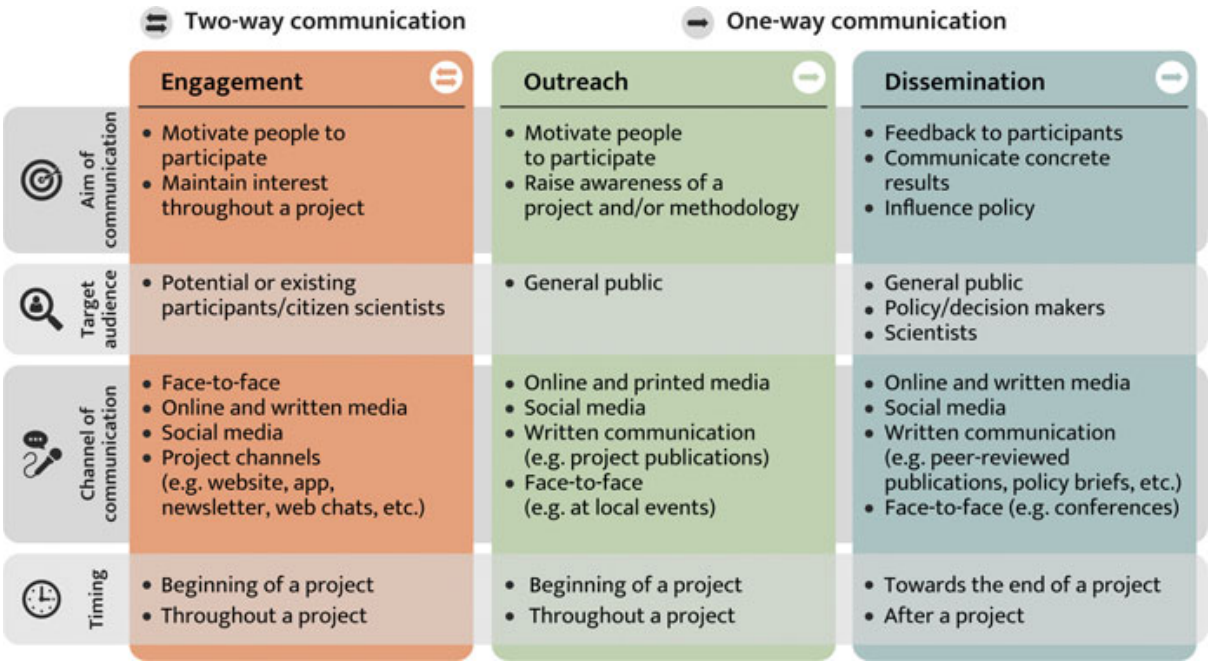


Figure 1. *Types of communication and their associated aims, target audiences, channels of communication, and most appropriate time point within a participatory project (Rüfenacht et al. 2021)*

The co-identified channels and strategies of science communication are listed below:

Channels

- Academic publications
- Other publications (magazines, online and/or offline, produced by associations, NGOs and diverse communities)
- Conferences at national and international level
- Events on CS (meetings carried out by participants and/or managers, etc.)
- Events on science communication (science fairs, European researchers' night, etc.)
- Events on the research topic (talks, debates, etc.)
- Local events organized by the communities (outside the research topic)
- Traditional media (TV, radio, press, etc.)
- Project channels (website, newsletter, etc.)
- Mailing and messaging platforms (telegram, whatsapp, etc.)
- Social networks (twitter, facebook, instagram, YouTube, linkedIn, etc.)

Strategies

- Contact with identified target groups
- Face-to-face interaction

- Online forum to discussion
- Storytelling
- Communication and dissemination plan
- Co-creation approaches

As for scientific publications related to CS, their exponential growth has been measured, as well as their increasing impact (Pelacho, Ruiz, Sanz et al. 2021).

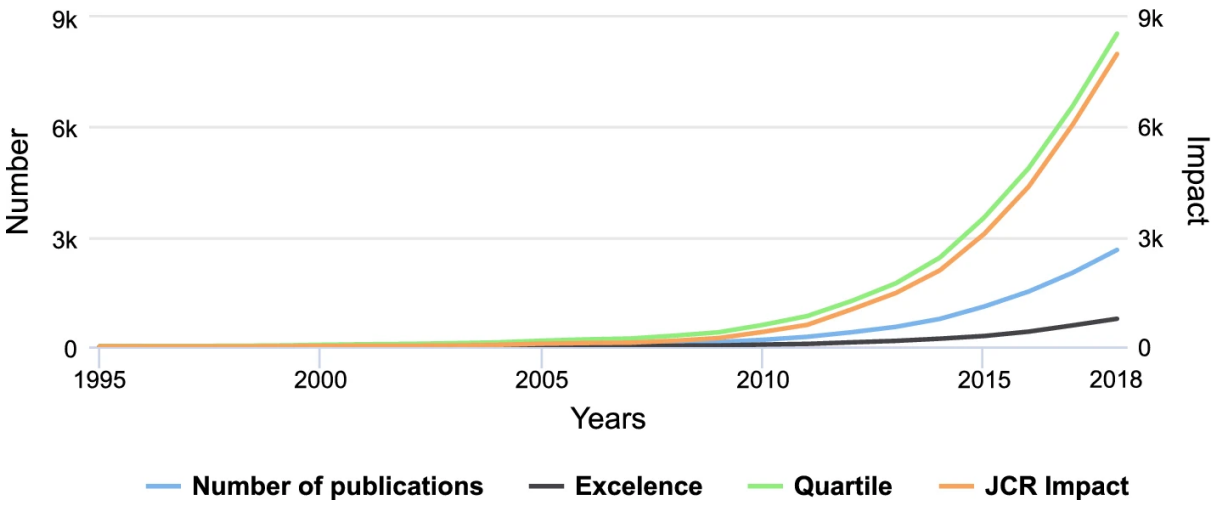


Figure 2. Evolution of number and impact of academic papers related to CS (Pelacho, Ruiz, Sanz et al. 2021)

The impact of academic publications in social networks has also begun to be measured, noting a significant growth mainly in twitter and blog posts, followed by facebook (Bautista-Puig et al. 2019).

Among the above mentioned channels and strategies, the one referring to **co-creation approaches** stands out. We can consider that this is the only element present in CS projects which is not in academic science. The co-creation approach - more or less present in the very different projects - implies, at a minimum, that there will be more people doing and communicating science. At the very least, the possibility of being part of a research project and generating knowledge is something conceptually very different from just receiving it. This is one of the principal reasons why CS has a special potential in science communication.

2.3 Barriers and challenges

Since CS mainly implies generation of scientific knowledge, there are barriers to effective communication shared with academic science. In addition, there are specific and challenging barriers in CS projects related to its relatively innovative methodologies. During the #CitSciComm Labs, we - all the NEWSERA participants - co-identified the principal obstacles regarding communication in the selected CS projects.

As for shared barriers, the following can be highlighted:

- **Lack of competences on communication and science communication**

Not all people have spontaneously good communication abilities. Of course there are examples of excellent science outreach workers without an academic or formal training in communication skills (just as there are excellent citizen scientists outside professional science). However, hundreds of academic programmes on science communication have been developed over the three past decades in order to assure the transferring of knowledge to the different audiences: society at large, policy makers, companies and industry, other scientists, and general journalists.

- **Lack of resources (e.g. time, funding) for developing communication strategies**

In any case, whether or not one has communication skills, time and resources are necessary for good science communication. A scientific project must include communication tasks as part of the research cycle, but if resources are insufficient, this last stage, especially science popularisation, is the most damaged without any doubt.

- **Disconnections between scientific jargon and stakeholders' languages**

This is one of the main barriers related to the previous one. The specificity of research concepts and terms, overall in scientific domains, makes especially difficult to properly transfer processes and results to diverse stakeholders. This requires considerable efforts to achieve matching the various languages.

- **Lack of wide communication strategies**

Overall in professional science, but in CS projects as well, there may be the unique interest of publishing in academic journals, also due to the current system of research assessment, mainly focused on the number of academic papers and its JCR impact factor. The publication of results of CS in academic publications is highly relevant in terms of validating methodologies and data, while at the same time these results must reach the population as a whole. A wide communication and dissemination plan is therefore relevant, *with the definition of concrete actions and the corresponding indicators for assessment.*

- **Inadequate acknowledgment of science communication relevance**

In recent decades there has been an important evolution regarding the need for greater communication of science to society at large, including legislative reforms that recognize this type of merits in the professional scientific career (e.g. Law of Science in Spain of 2011). Even so, there are difficulties of different kinds: (i) many

professional scientists have no interest - and / or resources - to carry out these tasks, (ii) many citizens have no interest in science or even distrust it, (iii) the cognitive deficit model continues to exist, according to which mistrust or lack of interest in citizens is always irrational or caused by ignorance (Irwin 1995).

The main specific barriers include:

- **Difficulties for CS to be recognised by academic science**

Whether due to lack of interest or lack of knowledge, there are still many CS results, methodologies and data that are not given credit by academic science. Although, as noted in section 2.2, there is a clear trend towards a greater presence - and recognition - of CS in academic science.

- **Lack of communication spaces between participants**

Physical meeting places, online platforms and forums are often difficult to get or maintain. However, this type of communication spaces favour the creation, not only of scientific encounters but also of affective links which led to strong communities of practice.

- **Lack of communication between participants and managers, in the case of top-down projects**

The lack of feedback between participants and managers, in the case of projects launched by academic institutions, is an important demotivating factor leading to the abandonment of projects. There is abundant literature reflecting this situation. Robust CS requires co-creative attitudes, even in projects where only relatively passive contributions are provided by participants.

- **Inadequate recognition of participants in scientific publications, in the case of top-down projects.**

Inadequate or absent recognition of participants in academic publications is also a major demotivating factor for continued participation. This is not only relevant for the sustainability of the projects but also from an ethical point of view because it could mean the instrumentalisation of the participants.

- **Difficulties related to open science**

The various relationships between open science and CS have been extensively studied, being CS one of the eight priorities for the Open Science Policy Platform (European Commission 2018). Claims about open data and open publishing are particularly necessary in CS projects, where citizen scientists contribute to the wider scientific community and society.

The main challenges can be summed up as follows:

- To ensure that all results are published through the most appropriate media (online and offline, inside or outside the academy, etc.) in accordance with the objectives of the projects.
- To communicate both the processes and the dissemination of results, and to do so by attending to the different representatives of quadruple helix and meeting their needs and/or requests.
- To establish a multidirectional communication, among all types of stakeholders.
- To ensure open science in all its aspects and stages, particularly in the publication of results.
- To have the necessary and diverse funding structures in place to support the above achievements, in particular by addressing technological and training gaps.
- To raise awareness among policy makers (scientific, educational, environmental, etc.) and academic scientists of the relevance of citizen science, both as a research methodology and as a powerful tool for science communication to bring science closer to society and to increase trust.

Our methodology for studying the effectiveness of science communication can be used in any type of ongoing research project. In CS, even more so, the analysis of social networks is fundamental, in particular because the exponential growth in the number of projects is largely due to the development of social networks and the technological means available. Kampal Social tool allows us to monitor many relevant parameters regarding science communication under the complex networks approach, revealing diverse relationships and their evolution, as well as properties useful as relevant indicators.

3. Methodology: data and tools to evaluate citizen science communication in citizen science projects

3.1 Collecting data on science communication in EU citizen science projects

NEWSERA conducted a survey on "A portrait of Science Communication Strategies by Citizen Science projects" (Task 2.1 within WP2), in order to collect data about communication tools, resources and targets as defined by CS projects. Respondents belong to different EU countries and develop research through citizen science in different domains. Results are similar and communication actions seem to be oriented towards: i) same target to be reached, mainly citizens; ii) social media as main channel for communication; iii) data access policy promoting open access (see *deliverable 2.1 Portrait of Citizen Science Communication Strategies in EU Citizen Science Projects*).

From the responses to the survey, the NEWSERA Consortium was able to contact with CS projects that explicitly indicated their interest for participating in #CitSciComm labs. A leader from each country (SfC, from Spain; FC.ID, from Portugal; UNIPD, from Italy) carried out accurate interviews with the first selected list of projects. During the individual interviews project leaders were asked about diverse aspects, several of them related to communication issues, such as:

- Main target audience (Children, Students of highschools, Teaching staff, Families, Amateurs, Skilled professionals, Organized people or Residents of an area)
- Can you describe your current communication strategy towards your target audience?
- Indicators for engagement of your target audience (how many people / stakeholders have you been able to engage)?
- What is your target for engagement?
- What is the most "hard-to-get" target audience from the quadruple helix in your project?
- What are your needs in terms of communication?

- Would you be willing to provide data on your current communication strategy (indicators) and on the changes derived during and after the implementation of the improved communication strategy to be co-created?

These important project-specific results still needed to be more concrete in order to assess the effectiveness of science communication in the projects. E.g. it is not enough knowing that the CS projects use twitter as a channel, but the monitoring requires knowing which accounts the project is linked to, which hashtags they use, etc., as well as how the different relationships evolve over time.

3.2 Collecting data of participant projects

In order to collect the necessary data for monitoring through the Kampal Social tool, the representatives of the 31 projects were asked to provide specific information, particularly related to the communication of their processes and results dissemination. The projects - 15 in Spain, 8 in Italy and 8 in Portugal - provided the information requested through a Google form, once the first round of Labs ended.

Figure 3. *Header of the form addressed to the projects participating in NEWSERA*

The corresponding datasheet with 26 fields (table 1) was filled with the information provided by the 31 projects. The goal was to have a starting point for each project which covered quadruple helix: citizen and society at large, academic

scientists, public sector and policy makers, industries and SMEs. Not all the fields were mandatory, since not all the projects had the same resources or status. The more information provided, the better the system will be able to measure the impact in different areas.

Field	Description, examples
ID	Project ID
Name of the project	Name of the project
Country	Main partner country
Language_1	Main language used in the project. Format ISO_639-1 https://es.wikipedia.org/wiki/ISO_639-1
Language_2	Secondary language used in the project. Format ISO_639-1 https://es.wikipedia.org/wiki/ISO_639-1
Domain	Examples: Biodiversity Ecology & Environment Biology Agriculture & Veterinary Science Insects and pollinators
Scale	Examples: European City National Regional Global Regional
Main_twitter_account	Main twitter account associated with the project. Example @CanSatSP
Twitter_academic_accounts_to_follow	Related twitter academic accounts separated by comma (researchers, researcher_groups,..): Example: @cosnet_bifi, @BarbaDelCid, MAX: 10 twitter accounts
Twitter_policy_makers_accounts_to_follow	Related twitter policy makers accounts separated by comma (governments, politics, ...): Example: @Ibercivis, @GobAragon, @EseroSp, @FECYT_Ciencia MAX: 10 twitter accounts
Twitter_journal_accounts_to_follow	Related journal twitter accounts separated by comma (newspapers, blogs,...): Example: @heraldoes, @elpais MAX: 10 twitter accounts
Main_facebook_account	URL Main facebook account associated to the project. Example https://www.facebook.com/Ibercivis/

	URL Main youtube account associated to the project
Main_youtube_account	https://www.youtube.com/channel/UCbkYEbbo8qMYbA3P5iLanGw
Keywords_to_follow	Keywords related to the project separated by comma. Try to use non generic keywords. Examples: CanSat 2018, satélite pequeño, MAX: 10
Hashtags_to_follow	Hashtags related to the project separated by comma. Try to use non generic keywords. Examples: #Cansat, #servetiv MAX: 10
Main_url	URL Main webpage associated to the project Example: https://ibercivis.es/project/cansat/
Publications_keywords_to_follow	Keywords used in scientific repositories to search for scientific publications related to the project separated by commas Example: CanSat 2018
Relevant_publications_1 (Title | year | DOI)	Specific publication associated with the research group to be monitored (number of citations along time) Example: Understanding the Citizen Science Landscape for European Environmental Policy: An Assessment and Recommendations | 2019 | http://doi.org/10.5334/cstp.239
Relevant_publications_2 (Title | year | DOI)	
Relevant_publications_3 (Title | year | DOI)	
Relevant_publications_4 (Title | year | DOI)	
Researcher_1_orcid	ORCID Researcher id (https://orcid.org) Example: 0000-0003-2772-3762
Researcher_2_orcid	0000-0023-2772-3762
Researcher_3_orcid	0000-0003-2773-3766
Researcher_4_orcid	0000-0003-2773-3766
Researcher_5_orcid	0000-0003-2773-3766

Table 1. *Projects data to be analyzed with Kampal Social tool*

Once the info requested was filled, Kampal started to collect data from multiple sources of information, using Kampal Social. For each project, a *kampal social project* was created with the information provided in the datasheet. Kampal Social deploys virtual machines with programs called collectors that are continuously scraping data from the internet and storing it in their databases. There are various

types of collectors and each one of them captures different metrics. Below is a brief summary of the type of data that is captured. Data collection depends on the source of information.

Twitter (continuous stream):

- Tweets
- Authors
- Mentions
- Retweets
- Links
- Hashtags
- Replies
- Tweet latitude and longitude if available
- Created_on, updated_on timestamps for tweets
- Number of likes

YouTube (collected daily for a given channel)

- Number of subscribers
- Number of visualizations
- Videos (channel list)
    - Title
    - Description
    - Date
    - Number of comments
    - Number of likes
    - Number of dislikes
    - Number of favourites

Google Scholar (collected weekly)

- Base researchers and related researchers
    - orcid
    - name
    - institution
    - total citations
    - publications
    - coauthors
- Base publications and related publications
    - title
    - publication year
    - number of citations
    - url
    - authors

Google Alert (collected daily)

- News and articles in blogs, forums, etc.

- title
- link
- published
- content

3.3 The Kampal Tool

Kampal Data Solutions S.L. is a spin-off of the University of Zaragoza (Spain) created in 2014. Kampal provides expertise and develops analytical tools to extract knowledge and value from data, supporting entities, research centers and companies in their management strategies and decision making activities.

Kampal has experience in the analysis of large and heterogeneous data sets, ranging from mobile communication data and characterization of social networks to the detection of collaboration patterns in R&D. Kampal applies methodologies taken from statistical physics, complex systems and computer science fields.

3.3.1 Kampal Social

Kampal social is an analytical tool for data visualization that offers a solution that allows the users to collect, filter, analyze, quantify and visualize the information they want from different internet sources.

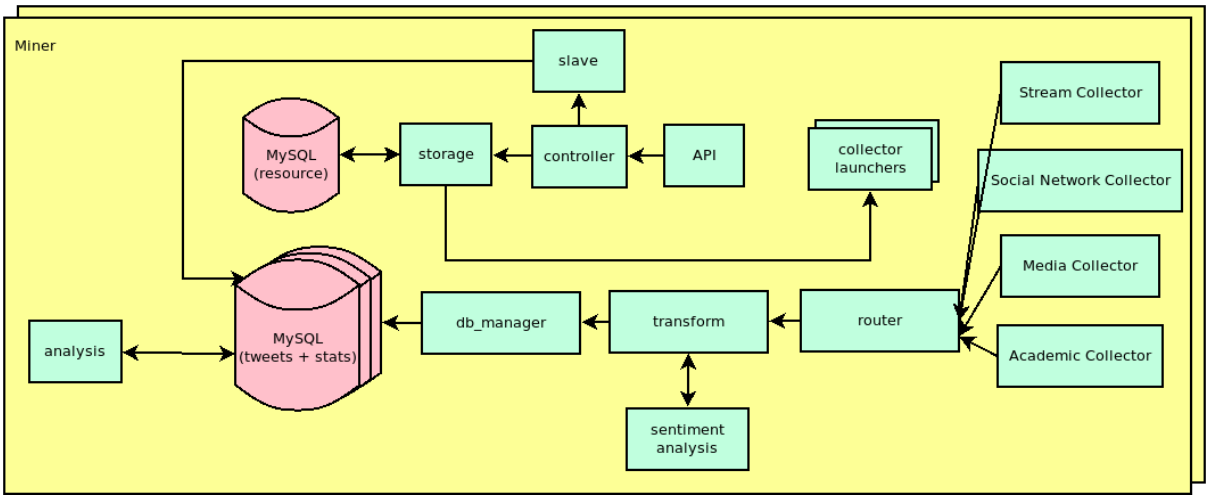


Figure 4. *Basic architecture diagram of a miner. A miner contains small programs called Collectors depending on source type that run endlessly crawling data from the web. The data crawled is transformed, enriched and routed to specific databases depending on project keywords.*

We can split the workflow in five steps:

1) **Project definition:** The user defines the type of information they want to analyze. The user provides keywords, accounts and other inputs through the Google form (figure 3), later transferred into a Kampal form (figure 5).

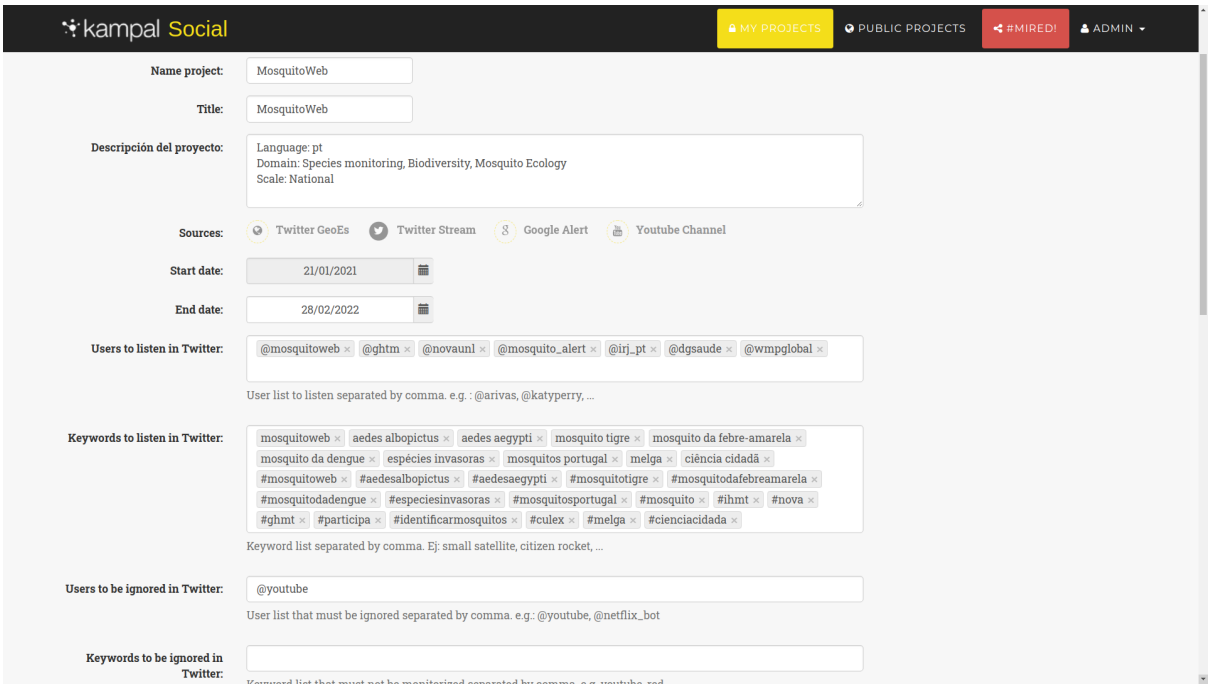


Figure 5. *Web form for project creation. The user selects which sources of information are going to be used and the initial set of keywords, users, publications, channels, etc. that are going to be monitored.*

2) **Data extraction:** Once the user has defined which information is going to be analyzed, the system monitors different internet sources like social networks, mass media, academic publications, forums, etc. and gathers raw data periodically.

3) **Transformation and loading:** All the information gathered from different sources needs to be processed. The data is filtered, cleaned, structured and stored into different databases.

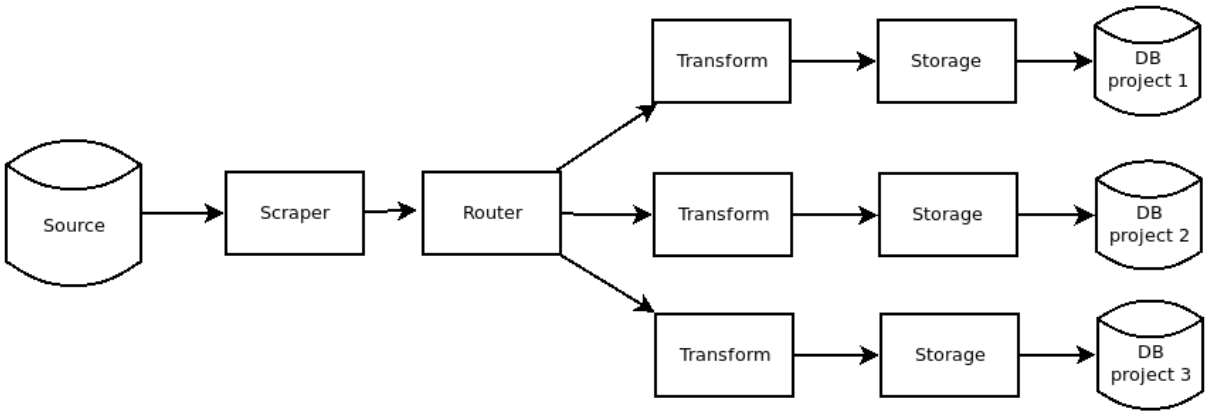


Figure 6. *Transformation flow. The data crawled from the web is usually unstructured and messy. It needs to be cleaned, transformed, homogenized and enriched before loading into a database.*

4) **Analysis:** Once the data has been homogenized and structured, different methods are used in order to identify, among other things:

- **Influencers:** the system uses classical methods based on "popularity" with metrics like mentions, followers or retweets and methods based on complex networks like the betweenness, pagerank, assortativity and others, to detect users that are hubs or have an influence in the discussion topic.

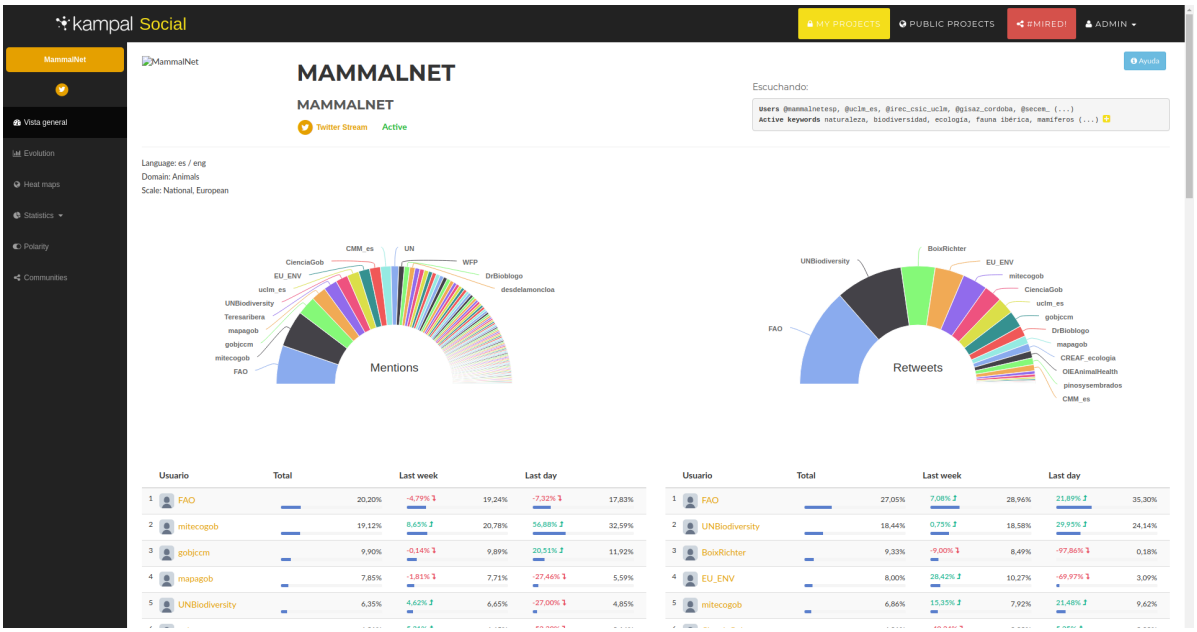


Figure 7. *Relevant users screen. This graph shows the most relevant users in Twitter categorized by Mentions and Retweets and their influence.*

- **Topics and keywords:** Natural Language Processing methods are used in order to detect keywords and topics that are being used within the scope of the project. The project can be edited to alter the keywords to follow if some new interesting topics appear during the course of the project. Besides this, other keywords that were initially introduced could be discarded if we discovered that they are too generic or introduce noise.

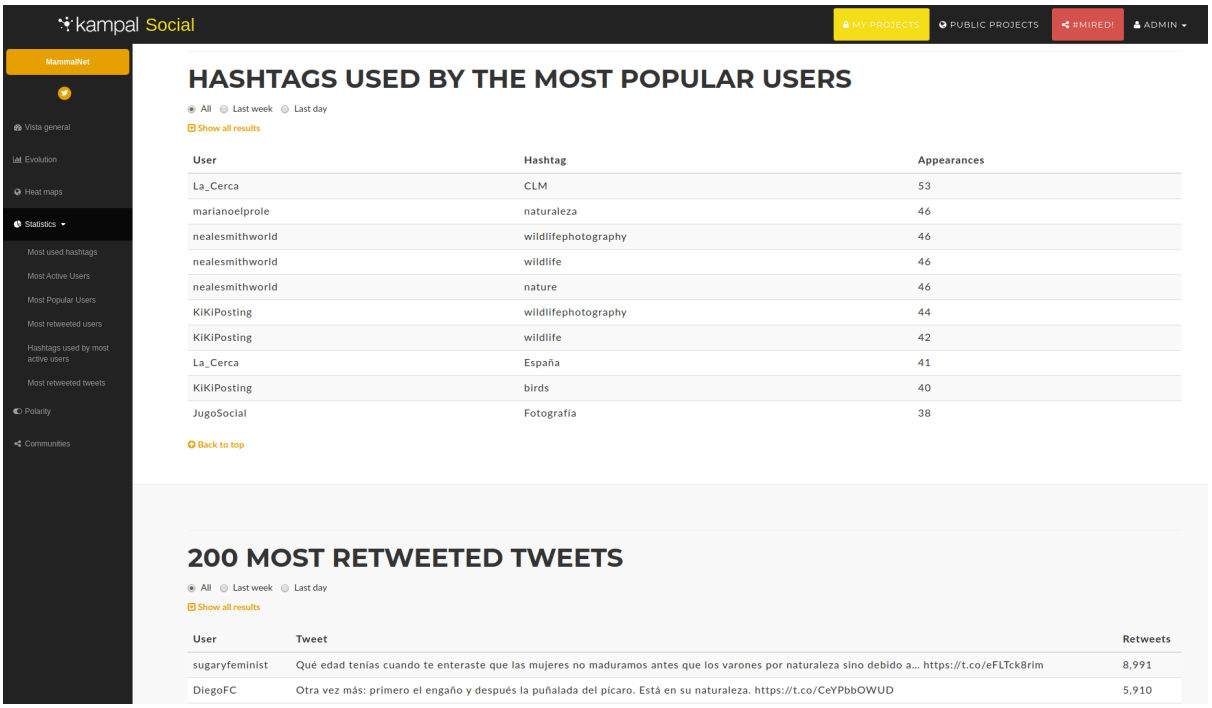
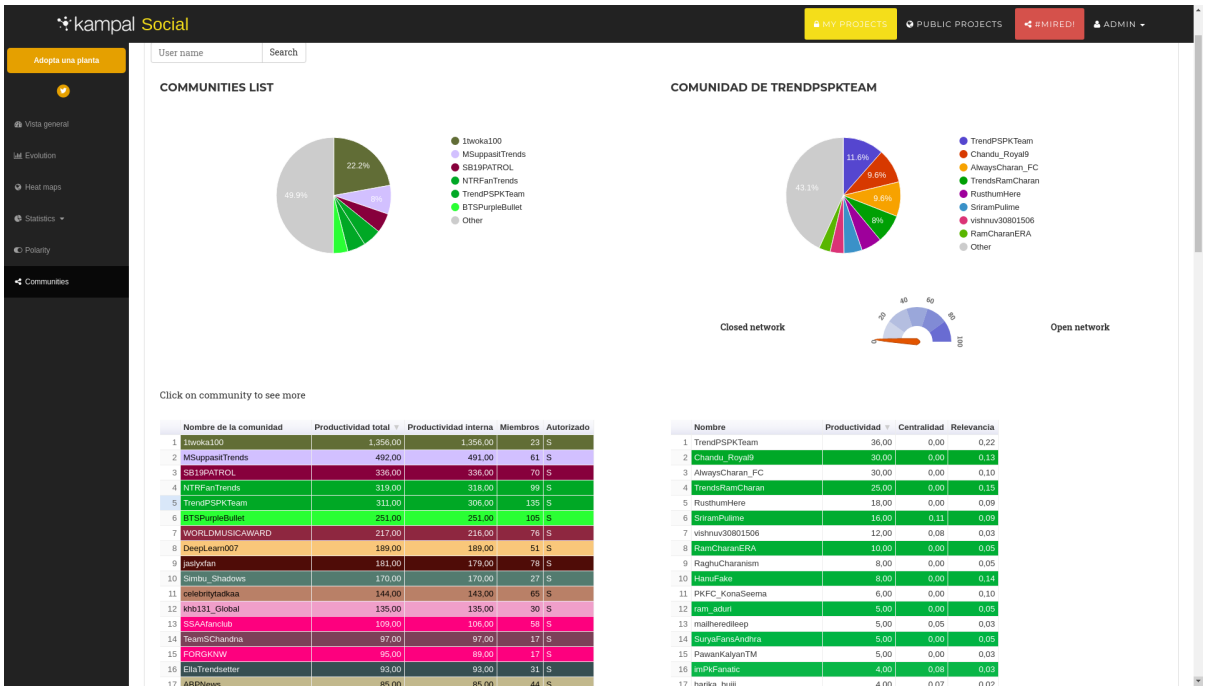


Figure 8. *Hashtags and topics list used in Twitter for a given project.*

- **Communities with common interests/topics:** From some sources of information the system can evaluate the networks that naturally are created from the interaction between the different users. Networks based on mentions, retweets, papers coauthors, comments, etc, are analyzed. The system is able to find communities of interaction between users. These communities can offer information about confronted groups, haters, bots, etc.



	Nombre de la comunidad	Productividad total	Productividad interna	Miembros	Autorizado
1	1twoka100	1,356,00	1,356,00	23	S
2	MSuppasitTrends	492,00	491,00	61	S
3	SB19PATROL	336,00	336,00	70	S
4	NTRFanTrends	319,00	318,00	99	S
5	TrendPSPKTeam	311,00	306,00	135	S
6	BTSPurpleBullet	251,00	251,00	105	S
7	WORLDMUSICAWARD	217,00	216,00	76	S
8	DeepLearn007	189,00	189,00	51	S
9	jaslyxfan	181,00	179,00	78	S
10	Simbu_Shadows	170,00	170,00	27	S
11	celebritytadkaa	144,00	143,00	65	S
12	khb131_Global	135,00	135,00	30	S
13	SSAAfanclub	109,00	106,00	58	S
14	TeamSChandna	97,00	97,00	17	S
15	FORGKNW	95,00	89,00	17	S
16	EllaTrendsetter	93,00	93,00	31	S
17	ABPNews	85,00	85,00	44	S

	Nombre	Productividad	Centralidad	Relevancia
1	TrendPSPKTeam	36,00	0,00	0,22
2	Chandu_Royal9	30,00	0,00	0,13
3	AlwaysCharan_FC	30,00	0,00	0,10
4	TrendsRamCharan	25,00	0,00	0,15
5	RusthumHere	18,00	0,00	0,09
6	SriramPulime	16,00	0,11	0,09
7	vishnuv30801506	12,00	0,08	0,03
8	RamCharanERA	10,00	0,00	0,05
9	RaghuCharanism	8,00	0,00	0,05
10	HanuFake	8,00	0,00	0,14
11	PKFC_KonaSeema	6,00	0,00	0,10
12	ram_aduri	5,00	0,00	0,05
13	mailheredileep	5,00	0,05	0,03
14	SuryaFansAndhra	5,00	0,00	0,05
15	PawanKalyanTM	5,00	0,00	0,03
16	imPkFanatic	4,00	0,08	0,03
17	harika_hyd	4,00	0,07	0,02

Figure 9. *Communities detected in Twitter ordered by relevance. The tool allows us to explore how communities are evolving, their members and their properties.*

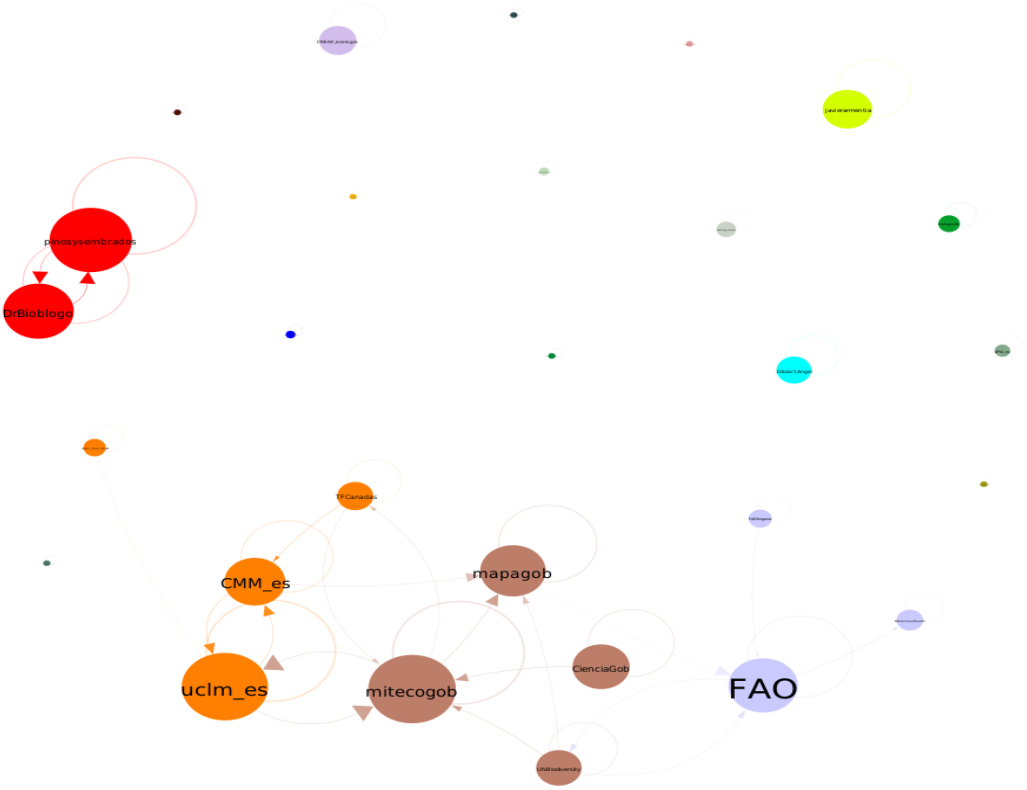


Figure 10. *Community clustering. Each node (circle) represents a community of multiple users.* A label for the community has been chosen automatically. Each link between nodes represents one or more interaction between communities. Communities with great interaction are represented close to each other. Color indicates similarity too.

- **Sentiment analysis:** Kampal Social evaluates the content of the data collected using Natural Language Processing techniques. The goal behind sentiment analysis is the process of determining the emotional tone behind a series of words, and is used to try to understand the attitudes, opinions and emotions expressed in an online text.

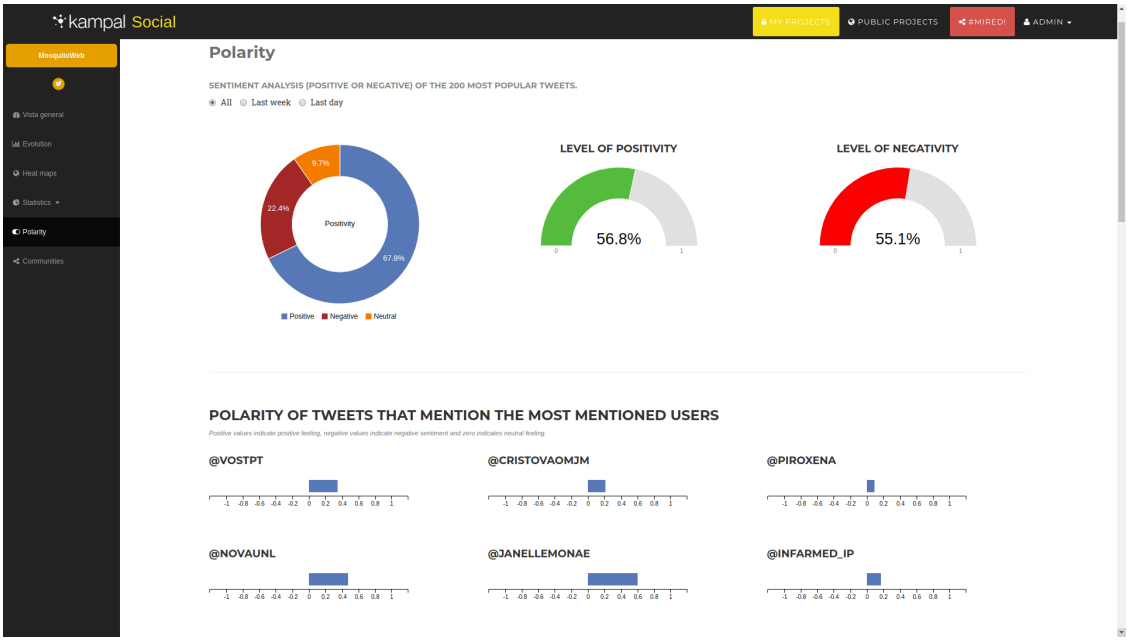


Figure 11. *Polarity graph. The tool analyzes the most relevant users to determine they are viewed as positive, negative or neutral.*

- **Geographical impact:** Extraction of the location by several methods: coordinates field, center of the bounding box of the place field, and user position if provided.

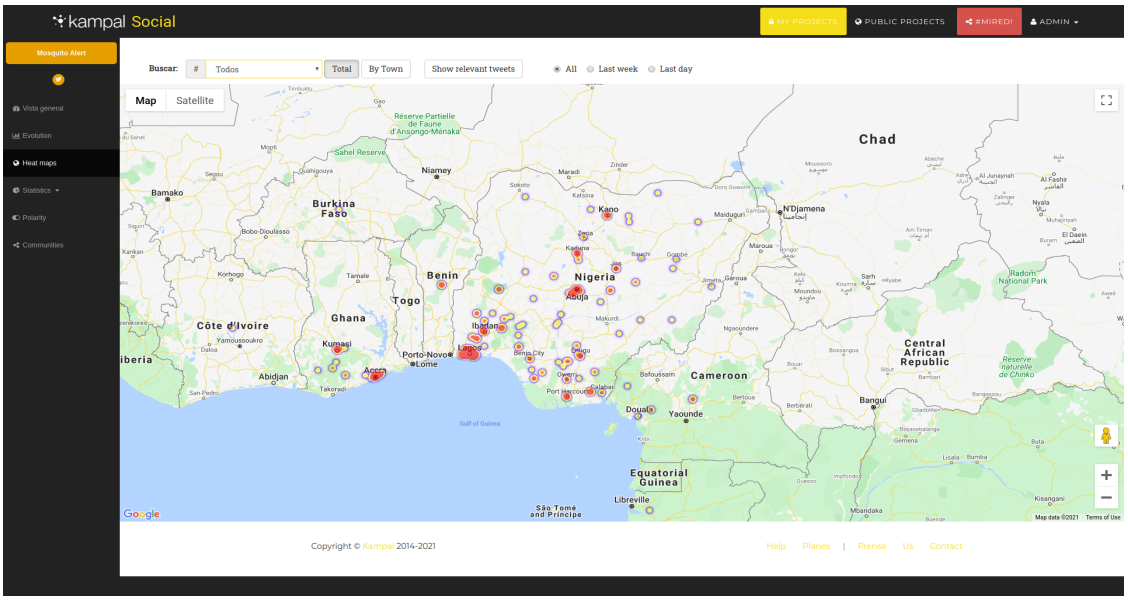


Figure 12. *Heatmap geolocation. Red spots indicate hubs with greater activity.*

- **Temporal evolution:** All the metadata captured is enriched with info about when data was created, updated or deleted. Multiple charts with time series depending on topics/users are shown in order to detect fluctuations in the activity for the global project or some specific topics.

Figure 13. *Activity charts. Each line represents a relevant user/keyword. X axis are days and Y axis are the number of occurrences.*

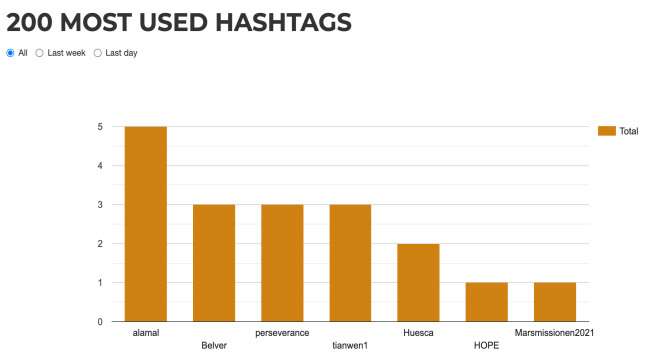
5) **Visualization**: Analysis is presented to the user using multiple visualizations, including community networks, geomapping, temporal evolutions, rankings and sentiment analysis.

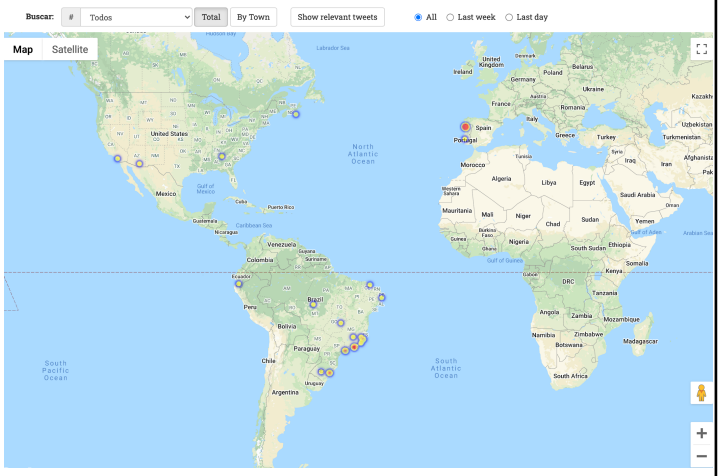
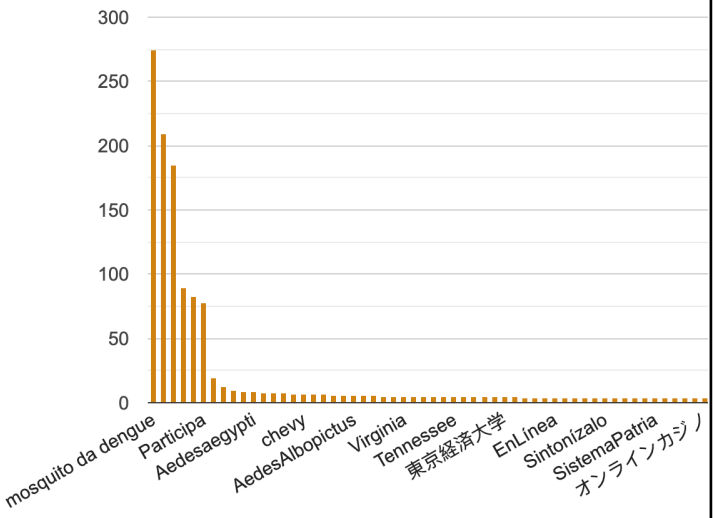
4. Results: Evaluation of citizen science communication of NEWSERA PILOTS

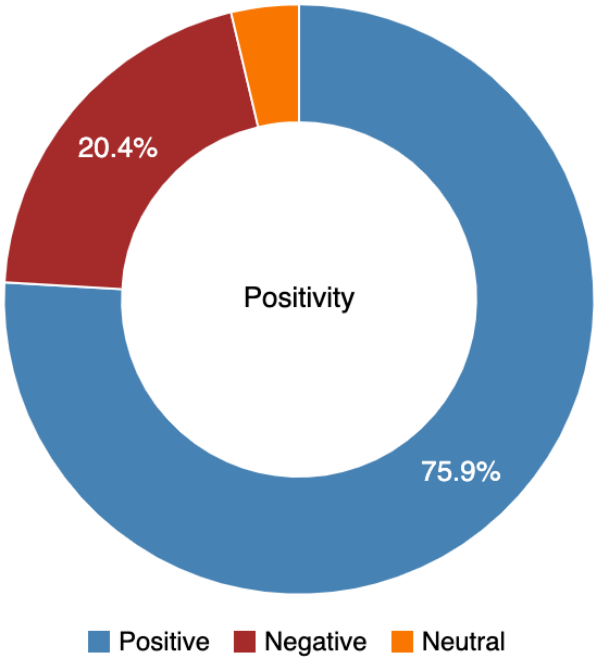
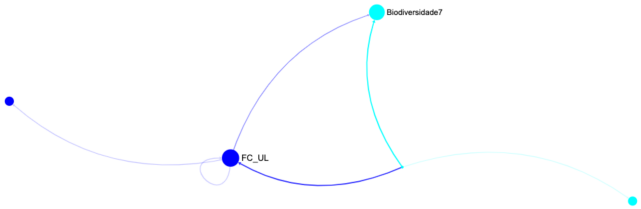
4.1 Results for each NEWSERA Pilot

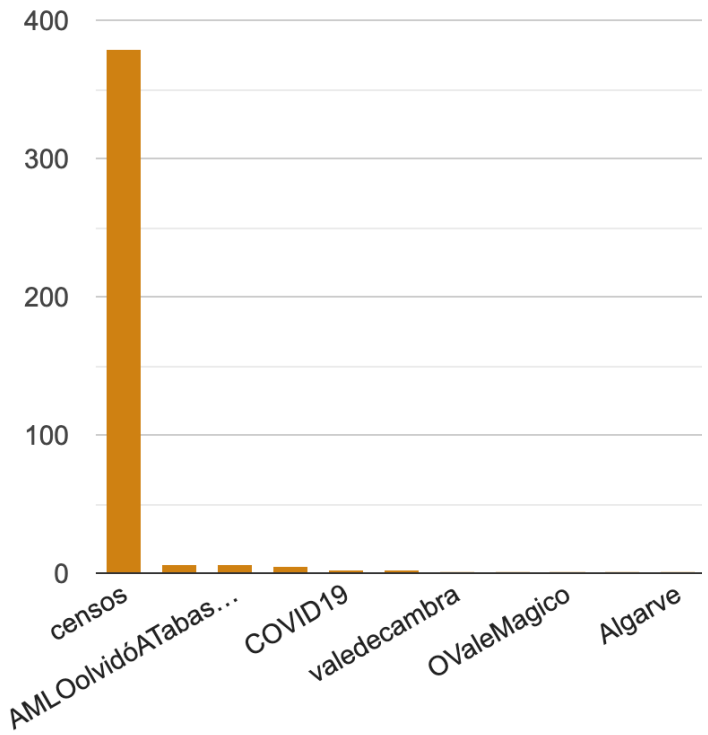
Here is presented a very brief analysis of each of the participating projects using Kampal Social. In this deliverable we present a preliminary analysis for each project, and related to twitter (although we are currently also collecting statistics from Youtube and Google Alerts). Each project will have real-time access to their analysis and will be able to evaluate the efficiency of the measures taken to improve their communication in a continuous manner. A deeper impact evaluation analysis will be presented in *D5.1 and D5.2 Iteration Cycle: Impact assessment of new communication strategies for each stakeholder group*.

4.1.1 Portugal

Name	Accounts & hashtags	Captures
FRISK	**Active keywords:** #friskproject, #silurus, #belver, #alamal	
FRISK analysis: The project does not have strong presence in social media		
Vacaloura.pt	**Active keywords:** vacaloura.pt, lucanus cervus, vaca-loura, lucanus, barbarossa, vaca-ruiva, **Hashtags:** #vacalourapt, #lucanuscervusportugal, #vacaruiva	**Not enough data collected yet**

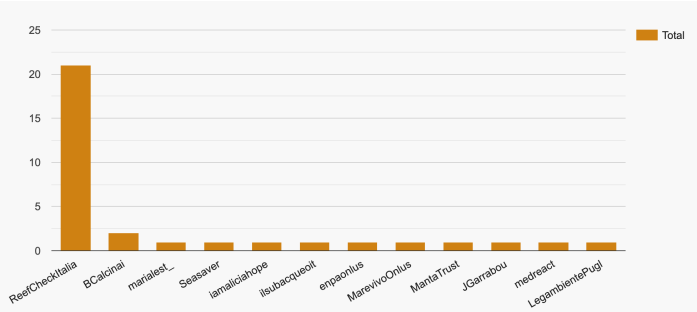
Mosquito Web	**Users:** @mosquitoweb, @ghtm, @novaunl, @mosquito_alert, @irj_pt, @dgsaude, @wmpglobal **Active keywords:** mosquitoweb, aedes albopictus, aedes aegypti, mosquito tigre, mosquito da febre-amarela, mosquito da dengue, espécies invasoras, mosquitos portugal, melga, ciência cidadã, **Hashtags:** #mosquitoweb, #aedesalbopictus, #aedesaegypti, #mosquitotigre, #mosquitodafebrea marela, #mosquitodadengue, #especiesinvasoras, #mosquitosportugal, #mosquito, #ihmt, #nova, #ghmt, #participa, #identificarmosquitos, #culex, #melga, #cienciacidada	 ***Figure 14.*** *Heatmap geolocation. Red spots indicate hubs with greater activity.*  ***Figure 15.*** *Most used keywords. Refinement is needed.*
MosquitoWeb Analysis: The project potentially has a strong community in Brazil (as well as Portugal). There is some refinement to be done on the active keywords.		

Biodiversity4All	**Users:** @biodiversity4a, @biodiversidade7 **Active keywords:** #biodiversity4all	 ***Figure 16.*** *Tweets have on average strong positive sentiment*  ***Figure 17.*** *The network is not strong*
Biodiversity4All analysis: Work is needed to improve network of mentions		
Memória para Todos	**Users:** @nova_fcsh **Active keywords:** memória para todos **Hashtags:** #memoriasdetrazerporcasa, #memoriaparatodos, #memoriasdasavenid	**Not enough data collected yet**

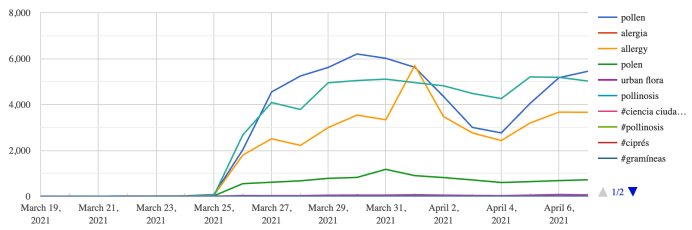
	as, #diasdamemoria, #memoriasdaopera	
Censos de Bolboletas	**Active keywords:** censos borboletas, censos, tagis **Hashtags:** #censosborboletas, #censosborboletasdeportugal	 ***Figure 18.*** *Most used hashtags.*
Censos de Bolboletas analysis: censos, a *common* word, should be removed from the set of words to follow. Work is needed to improve presence on twitter.		
LixoMarino.app	**Active keywords:** #lixomarinhoapp, #lixomarinho.app, lixomarinhoapp, lixomarinho.app	**Not enough data collected yet**

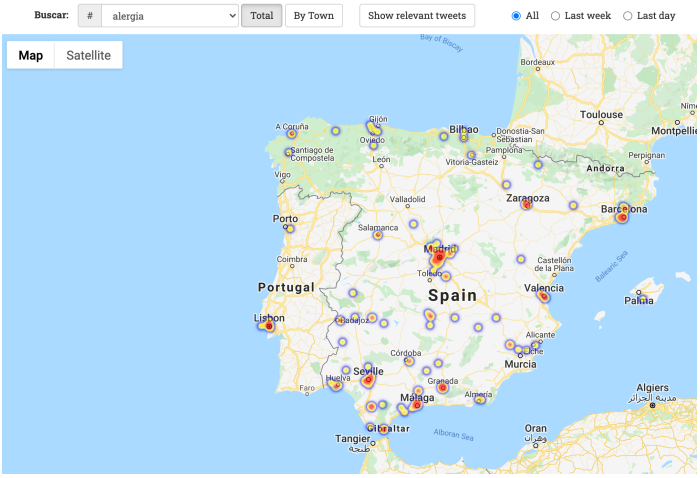
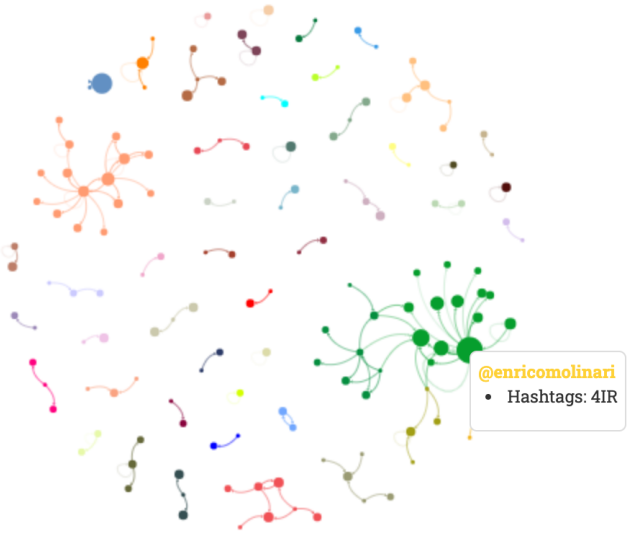
4.1.2 Italy

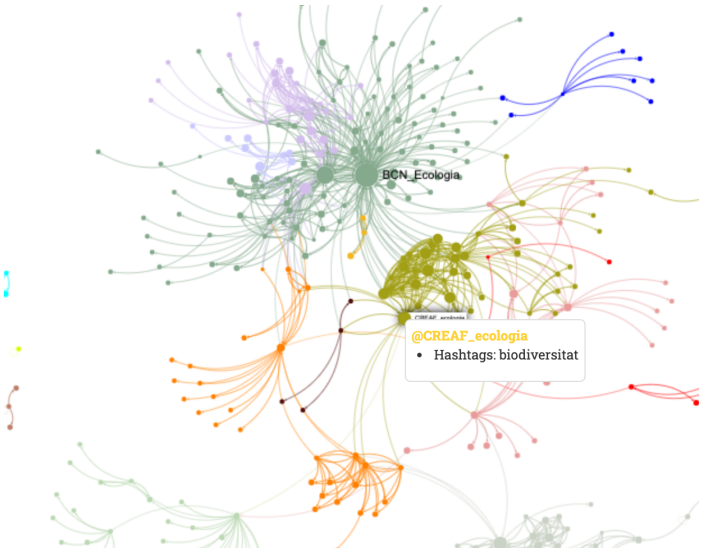
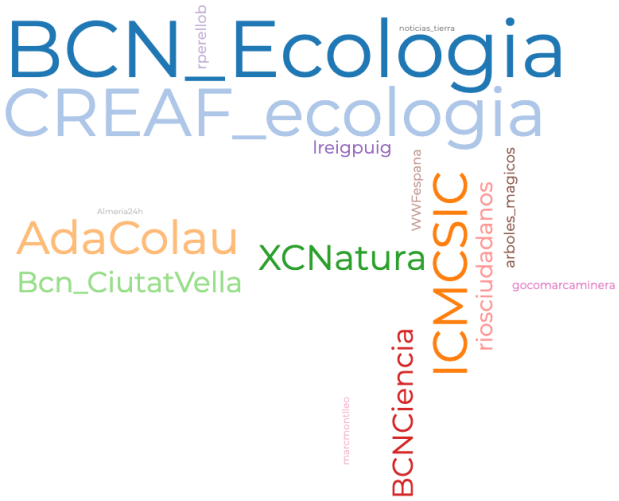
RomaUp	**Users:** @asudonlus	**Not enough data collected yet**

Easin	**Users:** @eucitsci	**Not enough data collected yet**
Easin analysis: The selected user to follow is the *ECSA* account. It's needed to improve users and keywords to follow.		
The school of Ants	No keywords provided	**Not enough data collected yet**
Reef Check Italia	**Users:** @reefcheckitalia	 ***Figure 19.*** *Active users*
Reef Check Italia analysis: It's needed to provide other users and keywords to follow.		
Onlus	No keywords provided	**Not enough data collected yet**
Sea cleaner	No keywords provided	**Not enough data collected yet**

4.1.3 Spain

Planttes	**Users:** @planttes, @aerobiologia, @uabbarcelona, @scaic_cat, @palinologos, @ae_aerobiologia **Active keywords:** alergia, allergy,	 ***Figure 20.*** *Active keywords evolution*

	polinosis, pollinosis, polen, pollen, flora urbana, urban flora, **Hashtags:** #citizenscience, #ciencia ciudadana, #alergy, #alergia, #pollinosis, #polinosis, #platanus, #ciprés, #parietaria, #olivo, #gramíneas	 ***Figure 21.*** *Alergia is one of the most used keywords, geolocating around cities*
Planttes analysis: Although some of the selected keywords might introduce some noise in the analysis, It is clear that the project tackles an issue of concern to the general public.		
COMPASS	**Active keywords:** cardiovascular diseases, mobile applications, ehealth, citizen science, empowerment, collective intelligence, smart city, mhealth, healthy aging, **Hashtags:** #citizenscience, #educacio360, #smartcity, #calidaddevida, #tejidocomunitario	 ***Figure 22.*** *@enricomolinari is one of the biggest nodes*
COMPASS: Some of the keywords (like empowerment) introduce noise in the analysis, as they are being used out of context.		

Urbamar	**Users:** @anellides, @icmcsic, @cos4cloud, @bluenetcat, @jaume_piera, @caterinons, @centreplatjabcn, @bcn_ecologia, @cciudadanaes, @maitepelacho, @volambiental, @bcnciencia, @xarxanet, @elheraldo **Active keywords:** urbmarbio, ciencia ciudadana, biodiversidad marina playas de barcelona, biodiversidad marina urbana, anellides, **Hashtags:** #urbamar, #urbamarbio, #biodiversidadmarinaurbana, #biodiversitatmarinalesplatgedebarcelona, #elmarmésaprop, #elmarmáscerca, #proyectosanellides, #anellides	  ***Figure 23 & 24.*** *Two big communities appear in the analysis. The first one is leaded by BCN_Ecologia, the second one by CREAF_ecologia*
Urbamar analysis: Maybe the users and keywords selected are out of scope. Further refinement is needed.		
Mosquito Alert	**Users:** @mosquito_alert, @creaf_ecologia,	

@ceabscic, @upfbarcelona, @mamirandatweet, @jmartinezpuente, @admartinou, @m_ferraguti, @tmontal, @europeveo, @onehealthpact, @aedescost, @eu_aliens, @infravec2, @salutpublicabcn, @cciutadana, @ibercivis, @bcnciencia, @rubueno, @eu_h2020, @diariara, @lavanguardia, @elperiodico, @cronicaglobal

Active keywords:

mosquito, aedes, picaduras, modelo, notificaciones, culex, mosquitoes

Hashtags:

#mosquito, #citizenscience, #aedesalbopictus, #aedes, #dengue, #zika

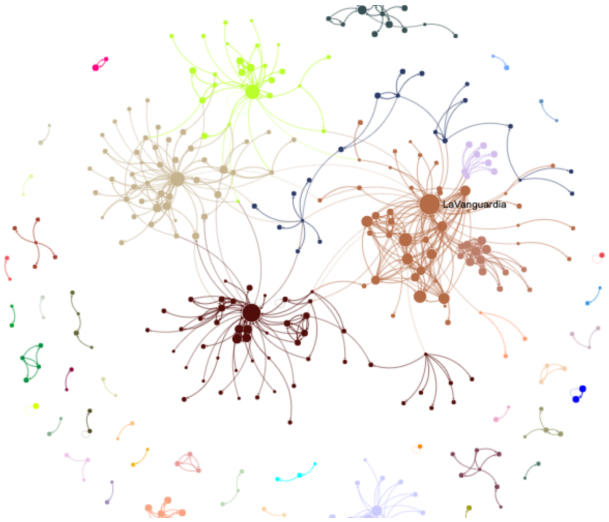


Figure 25: *Since the beginning of the analysis, the mosquito_alert account is not yet connected with two main networks (leaded by @LaVanguardia and @elpediodico)*

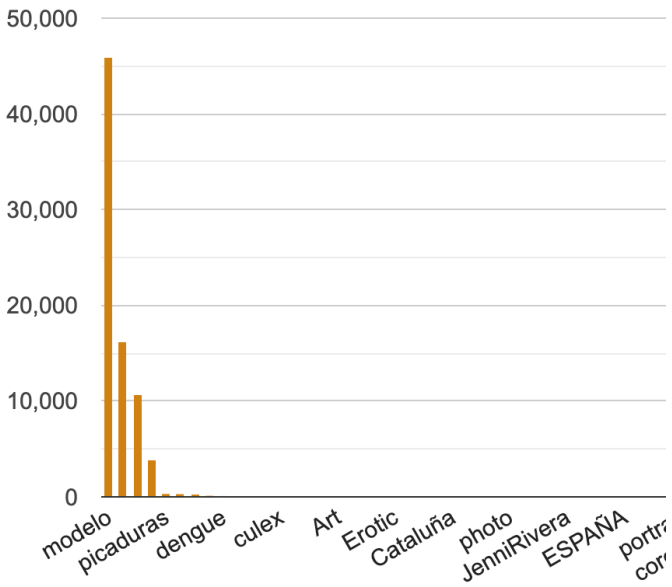
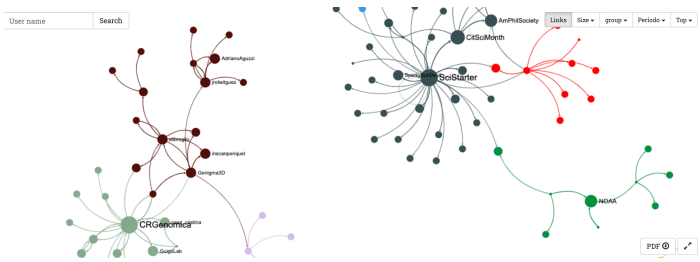
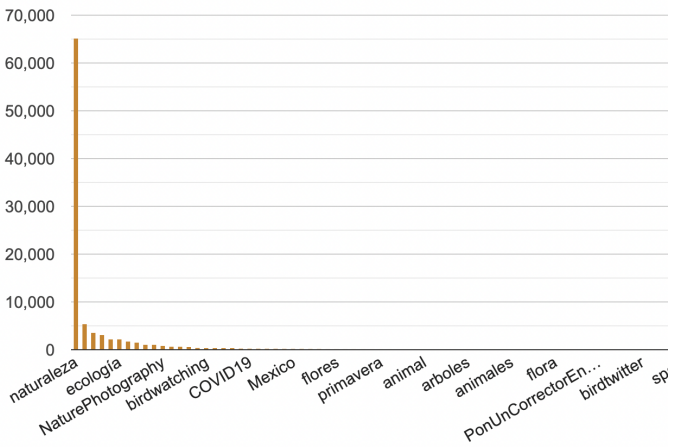


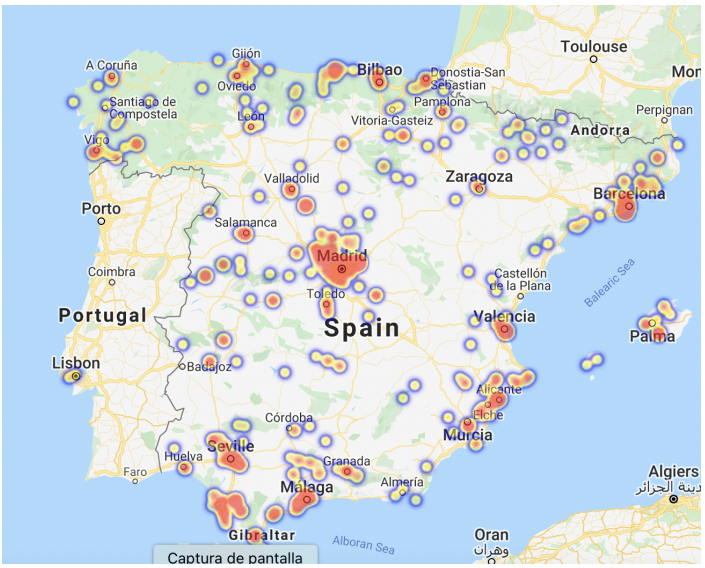
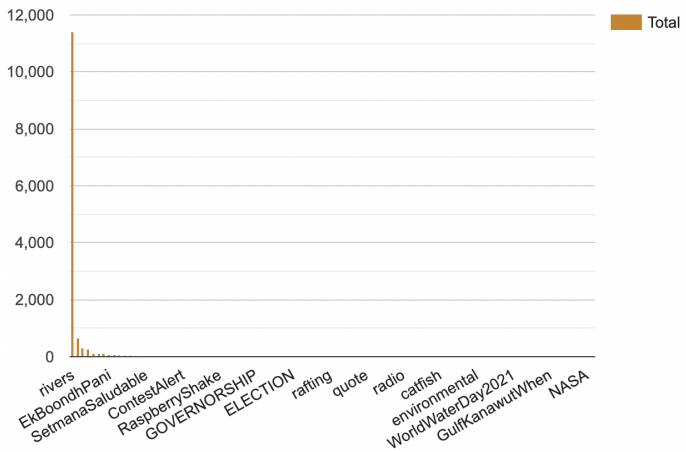
Figure 26: *Most used hashtags. Maybe "modelo" keyword is used out of context*

Mosquito Alert: Some active keywords and users are overshadowing main users related to the projects. A review of keywords, hashtags and users is needed to match the project.

Biodiversidad Virtual	**Users:** @bvwerbrs, @rjbotanico, @mncncomunica, @museuterresebre, @botanicoucm, @botanicsoller, @jardbotanicoclm, @eeamadrid, @equo_acoruna, @theooberhuber, @ecologistasen, @ecologistasclm, @ecoaccionhuesca **Active keywords:** biodiversidad, fotografía, testing, testing power, insectarium, taxofoto, bv, punto bv, observatorio ciudadano de la biodiversidad, ocb, **Hashtags:** #fotografia biodiversidad, #biodiversidadvirtual, #taxofoto, #taxovideo, #laminasbv, #testing	 ***Figure 27:*** *Geolocation of tweets with #testing hashtag.*
Biodiversidad Virtual Analysis: Some of the analyzed hashtags (like #testing) introduce noise in the analysis.		

Genigma	**Users:** @genigma3d, @marciuslab, @cnag_eu, @mamartirenom, @crgenomica, @marcdiesse, @elibroglio, @jrotwitguez, @lifetimeini **Active keywords:** 3dgenomics **Hashtags:** #citizenscience, #cienciaciudadana, #ciènciaciutadana, #3dgenomics	 ***Figure 28:*** *CitizenScience is the most used keyword*  ***Figure 29:*** *Communities*
Genigma analysis: @genima3d community is not yet connected with @SciStarter community.		

DENIS	**Users:** @denisproject_en, @_usj_, @insilicoheart, @investigausj, @boincnetwork, @boincprojects, @gobaragon, @zaragoza_es **Active keywords:** volunteer computing, cardiac modeling, electrophysiology, **Hasthags:** #citizenscience, #health, #electrophysiology, #volunteercomputing, #boinc	 ***Figure 30:*** *Most active keywords*
Denis analysis: Common keywords such as *health* and users such as @gobaragon insert noise into the analysis.		
MammalNet	**Users:** @mammalnetesp, @uclm_es, @irec_csic_uclm, @gisaz_cordoba, @secem_, @ecologiaumh, @christiangortaz, @uclmdivulga, @creaf_ecologia, @efsa_eu, @animals_efsa, @fao, @mapagob, @cienciagob, @mitecogob, @gobjccm, @oieanimalhealth, @eu_env, @unbiodiversity, @maresdeevelyn, @jg_cantero, @escuderoroman,	 ***Figure 31:*** *Most used hashtags*

	@cmm_es, @boixrichter, @pinosysembrados, @drbioblogo, @javierarmentia, @apia_es **Active keywords:** naturaleza, biodiversidad, ecología, fauna ibérica, mamíferos, fototrampeo, conservación, cámaras trampa, **Hashtags:** #cienciaciudadana, #wildlifephotography, #naturaleza, #biodiversidad, #fauna, #mamíferos, #conservación, #ecología, #fotografía, #wildlifediseases	 ***Figure 32:*** *Heatmap for "naturaleza" keyword*
Mammalnet analysis: It is needed to remove from the analysis keywords generating noise, such as *naturaleza (nature).*		
RiuNet	**Users:** @fehmlab, @riunet, @unibarcelona, @icrawater, @idaea_csic, @irbioub, @irtacat, @unirioja, @ecocrealab, @aigua_cat, @ch_ebro, @agbar, @ub_divulga, @ecomandanga **Active keywords:** rivers, ecological status, temporary rivers, macroinvertebrates, water quality, environmental	 ***Figure 33:*** *Most used hashtags*

<table>
<tr>
<td></td>
<td>education, hydrology, citizen science, biomonitoring,

Hashtags:

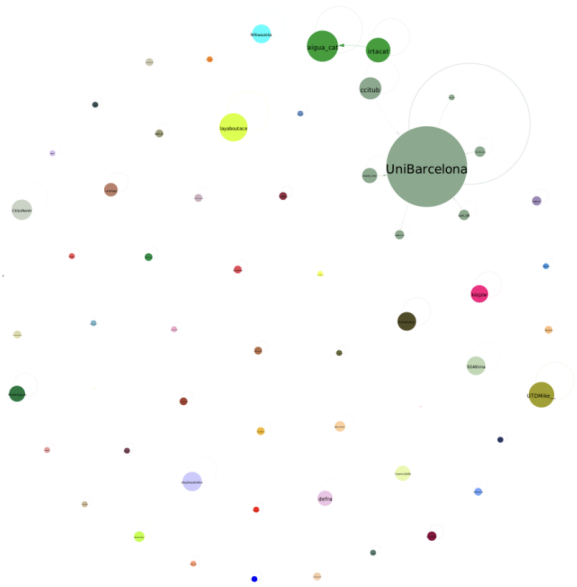
#ecologicalstatus, #temporaryrivers, #macroinvertebrats, #waterquality, #environmentaleducation, #hydrology, #citizenscience, #biomonitoring, #macroinvertebrados, #biomonitoreo</td>
<td> 

Figure 34: @UniBarcelona is the biggest community in this analysis</td>
</tr>
<tr>
<td colspan="3">Riunet analysis: It is needed to remove from the analysis keywords generating noise, such as rivers.</td>
</tr>
<tr>
<td>Adopt a plant</td>
<td>Active keywords:

long-term monitoring, plants, population abundance, trends, extinction risk,

Hashtags:

#citizenscience, #plantmonitoring</td>
<td>
<table>
<tr><th>Hashtag</th><th>Appearances</th></tr>
<tr><td>trends</td><td>48,215</td></tr>
<tr><td>plants</td><td>38,155</td></tr>
<tr><td>DarkSkyAPI</td><td>454</td></tr>
<tr><td>StanWorld</td><td>375</td></tr>
<tr><td>WeatherbitAPI</td><td>288</td></tr>
<tr><td>citizenscience</td><td>270</td></tr>
<tr><td>SB19</td><td>249</td></tr>
</table>
Figure 35: Hashtag appearances
</td>
</tr>
<tr>
<td colspan="3">Adopt a plant: Active keywords are too generic and do not show the actual status of the project</td>
</tr>
</table>

Vivencia Dehesa	**Users:** @vivencia_dehesa, @indehesa, @seepastos, @agroecologiaumh, @secs_spain, @cei_a3, @ias_csic, @mitecogob, @diputacioncc, @mapagob, @euagri, @eu_env, @paconadal, @escarabajotve, @lvnatural, @elobjetivoverde, @lacronicaverde, @ecogallego, @maruizlv **Active keywords:** vivenciadehesa, dehesa, agriculturaregenerativa, biodiversity, swales, aves, wildlife, **Hashtags:** #dehesadeextremadura, #espaciosprotegidos, #wildlife, #wetland	 ***Figure 36:*** *#Dehesa (one of the active keywords) appears in some locations in spain*
Vivencia Dehesa analysis: There are some users in the analysis such as @minecogov and @eu_env that are overshadowing other users.		

Biook biohacking	**Users:** @biook_org **Active keywords:** ciencia para niños, talleres de ciencia, ciencia abierta, ciencia divertida, experimentos de ciencia, biología, biohacking, innovación social, **Hashtags:** #cienciaparaniños, #zientzia, #diybio, #cienciacomunitaria, #cienciaparatodos, #egizuzientzia, #hazciencia, #sciencecommunity, #socialinnovation	 ***Figure 36:*** *Biggest communities*
Biook analysis: Biook_org does not appear between biggest communities		

4.2 Next steps

There are two main tasks in improving the analysis of individual projects. It is worth remembering that the objective of Kampal is to obtain a set of quantitative indicators that can measure the efficiency of improvements in the communication strategy of each of the projects.

Through the analysis carried out (briefly shown in the previous section), we have seen how, on the one hand, there are projects that are carrying out an analysis of hashtags or keywords that introduce noise into the system. For these, a review of users, keywords and hashtags is necessary. On the other hand, there are projects that do not have a twitter account. Although the creation of one is recommended, we understand that there may not be sufficient resources available. Even so, the analysis of related keywords and hashtags can indicate the extent to which they are being used, showing us possible target audiences.

The other task is to extend the analysis to other platforms. To this end, Kampal has developed tools that will allow similar analyses to be carried out on Youtube, Google Alerts and Google Scholar. This will make it possible to extend the analysis by introducing new audiences. These new tools will also be available in real time for each of the projects.

5. Conclusions

A first result has to do with the difficulties encountered in the application of the methodology. In order to harness the power of the Kampal tool, more data is needed. It is to be hoped that the actions to be introduced in the projects after this first evaluation will lead to a greater amount of available data, which in turn will reflect better communication strategies.

Precisely the fact that there is still little data available for many of the projects reveals that these projects do not have an elaborated communication strategy, at least on twitter. This lack of strategy is also evident when, in many cases, very generic keywords (e.g. model, nature, rivers) are used, which introduce a great deal of noise and therefore lose analytical capacity.

Of course, there exists a big variety among projects regarding their twitter impacts. Some projects are carried out by small groups or they are of very recent creation, others are long-term projects, others are developed by well-established research centres and/or funded by relevant entities.

The continuous contact between NEWSERA, the pilots and stakeholders of the quadruple helix, particularly through the #CitSciComm Labs, will allow us to gather more and new data one in order to refine and complete the analyses. This first assessment of actions and co-identified indicators will be useful for the iterative cycles of our impact evaluation framework, and eventually to consolidate and/or improve science communication through citizen science.

6. References

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). http://dx.doi.org/10.3390/publications7010012

BAYSICS (n.d). Bavarian Synthesis Information Citizen Science Portal for climate research and science communication. Retrieved from: https://www.baysics.de/

Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K.V., & Shirk, J. (2009). Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. *BioScience*, 59(11), pp. 977–984. https://doi.org/10.1525/bio.2009.59.11.9

Felt, U., Wynne, B., Callon, M., Gonçalves, M.E., Jasanoff, S. et al. (2007). Taking European knowledge society seriously. Report of the expert group on Science and governance to the science, economy and society directorate, Directorate-General for research, European Commission. Sixth framework programme. ⟨hal-02823887⟩

Göbel, C., Nold, C., Berditchevskaia, A., & Haklay, M. (2019). How does citizen science 'do' governance? Reflections from the DITOs project. *Citizen Science: Theory and Practice*, 4(1), 31. https://doi.org/10.5334/cstp.204

Jasanoff, S. (2003) Technologies of Humility: Citizen Participation in Governing Science. *Minerva,* 41, pp.223–244. https://doi.org/10.1023/A:1025557512320

Ley de la Ciencia en España (2011). https://www.recolecta.fecyt.es/sites/default/files/contenido/documentos/Ciencia_Libro_XMF.pdf

Pelacho, M., Ruiz, G., Sanz, F., Tarancón, A., & Clemente-Gallardo, J. (2021). Analysis of the evolution and collaboration networks of citizen science scientific publications. *Scientometrics* 126, pp.225–257. https://doi.org/10.1007/s11192-020-03724-x

Red4C (n.d). Red4C: Ciencia Ciudadana y Cambio Climático. Retrieved from: https://red4c.es/

Rüfenacht, S. Woods, T., Agnello, G., Gold, M., Hummer, P., Land-Zandstra, A., & Sieber, A. (2021) Communication and Dissemination in Citizen Science. In: Vohland

K. et al. (eds) *The Science of Citizen Science*. Springer, Cham. https://doi.org/10.1007/978-3-030-58278-4_24

Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R. & Wagenknecht, K. (2021) Editorial: The Science of Citizen Science Evolves. In: Vohland K. et al. (eds) *The Science of Citizen Science*. Springer, Cham. https://doi.org/10.1007/978-3-030-58278-4_1

Von Schomberg R., Keesenberg R. (2018). Open Science Policy Platform Recommendations (OSPP-REC). European Commission. Directorate-General for Research and Innovation.