



HAL
open science

CS tools final release - EU Crowd4SDG Project

Jesús Cerquides, Mehmet Oğuz Mülâyim, Barbara Pernici, Carlo A. Bono,
Rosy Mondardini, Marc Santolini, Jose Luis Fernandez-Marquez

► To cite this version:

Jesús Cerquides, Mehmet Oğuz Mülâyim, Barbara Pernici, Carlo A. Bono, Rosy Mondardini, et al.. CS tools final release - EU Crowd4SDG Project. European Union's Horizon 2020 research and innovation programme. 2023. hal-04706198

HAL Id: hal-04706198

<https://hal.science/hal-04706198v1>

Submitted on 23 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



Crowd4SDG

Citizen Science for the Sustainable Development Goals

Deliverable 2.3

CS tools final release

Deliverable identifier: D2.3

Due date: 30/04/2023

Document release date: 20/04/2023

Nature: Report

Dissemination Level: Public

Work Package: 2

Lead Beneficiary: CSIC

Contributing Beneficiaries: POLIMI, UNIGE, UNITAR

Document status: Final

Abstract:

The Citizen Science Solution Kit (CSSK) is a set of open-source software tools developed and maintained by the Crowd4SDG project partners to enable citizen scientists to conceive, design, launch, and manage their online citizen science projects and analyze the results and the performance of their community without any coding skills. This deliverable summarizes the development of CSSK tools throughout the project detailing their final releases and describing the real-world case studies that demonstrate the potential of the CSSK in citizen science.

For more information on Crowd4SDG, please check: <http://www.crowd4sdg.eu/>





Document history

	Name	Partner	Date
Authored by	Jesús Cerquides Bueno Oguz Mulayim Barbara Pernici Carlo Bono Rosy Mondardini Marc Santolini Jose Luis Fernandez-Marquez	CSIC CSIC POLIMI POLIMI CSCZ UPC UNIGE	28/02/2023
Edited by	Jesús Cerquides Bueno Oguz Mulayim	CSIC	01/03/2023
Revised by	Laura Wirtavuori Marc Santolini	CERN UPC	29/03/2023
Edited by	Jesús Cerquides Bueno Oguz Mulayim Barbara Pernici Carlo Bono	CSIC CSIC POLIMI POLIMI	12/04/2023
Reviewed by	Jose Luis Fernandez-Marquez	UNIGE	19/04/2023
Edited by	Jesús Cerquides Bueno Oguz Mulayim	CSIC	19/04/2023
Approved by	François Grey Jose Luis Fernandez-Marquez	UNIGE	20/04/2023



Table of Contents

Document history	2
Project Partners	4
Crowd4SDG in brief	5
Grant Agreement description of the deliverable	6
Purpose and scope of the deliverable	7
1. Introduction	8
1.1. The Crowd4SDG Citizen Science Solution Kit	8
1.2. Structure of the document	8
2. The Citizen Science Solution Kit	9
2.1. State of the art	9
2.2. Social media analysis pipeline	12
2.3. Conceptual framework	13
2.4. CSSK user roles	15
2.5. Selecting and filtering data: VisualCit	15
2.6. Building crowdsourcing projects with CSPB	16
2.7. Social media consensus analysis with crowdanalysis	17
2.8. Decidim4CS	17
2.9. CSLogger	18
2.10. CoSo (Collaborative Sonar)	18
3. Validation with case studies	20
3.1. Validation methods	20
3.2. Covid-19 social distancing	23
3.3. Covid-19 vaccines and menstruation	23
3.4. Albania earthquake	24
3.5. TriggerCit	24
3.6. Automatic geolocation of photos	25
3.7. Crowdsourced geolocation	25
4. Fulfillment of tasks	27
5. Interaction of WP2 with other Crowd4SDG work packages	30
5.1. Connection with WP3	30
5.2. Connection with WP4	30
5.3. Connection with WP5	31
5.4. Connection with WP6	31
6. Conclusions	32
References	33
Annex : List of abbreviations	37



Project Partners

	Partner name	Acronym	Country
1 (COO)	Université de Genève	UNIGE	CH
2	European Organization for Nuclear Research	CERN	CH
3	Agencia Estatal Consejo Superior de Investigaciones Científicas	CSIC	ES
4	Politecnico di Milano	POLIMI	IT
5	United Nations Institute for Training and Research	UNITAR	CH
6	Université Paris Cité	UPC	FR



Crowd4SDG in brief

The 17 Sustainable Development Goals (SDGs), launched by the UN in 2015, are underpinned by over 160 concrete targets and over 231 measurable indicators. Some of these indicators initially had no established measurement methodology. For others, many countries do not have the data collection capacity. Measuring progress towards the SDGs is thus a challenge for most national statistical offices.

The goal of the Crowd4SDG project is to research the extent to which Citizen Science (CS) can provide an essential source of non-traditional data for tracking progress towards the SDGs, as well as the ability of CS to generate social innovations that enable such progress. Based on shared expertise in crowdsourcing for disaster response, the transdisciplinary Crowd4SDG consortium of six partners is focusing on SDG 13, Climate Action, to explore new ways of applying CS for monitoring the impacts of extreme climate events and strengthening the resilience of communities to climate related disasters.

To achieve this goal, Crowd4SDG is initiating research on the applications of artificial intelligence and machine learning to enhance CS and explore the use of social media and other non-traditional data sources for more effective monitoring of SDGs by citizens. Crowd4SDG is using direct channels through consortium partner UNITAR to provide National Statistical Offices (NSOs) with recommendations on best practices for generating and exploiting CS data for tracking the SDGs.

To this end, Crowd4SDG rigorously assesses the quality of the scientific knowledge and usefulness of practical innovations occurring when teams develop new CS projects focusing on climate action. This occurs through three annual challenge based innovation events, involving online and in-person coaching. A wide range of stakeholders, from the UN, governments, the private sector, NGOs, academia, innovation incubators and maker spaces are involved in advising the project and exploiting the scientific knowledge and technical innovations that it generates.

Crowd4SDG has six work packages. Besides Project Management (UNIGE) and Dissemination & Outreach (CERN), the project features work packages on: Enhancing CS Tools (CSIC, POLIMI) with AI and social media analysis features, to improve data quality and deliberation processes in CS; New Metrics for CS (UP), to track and improve innovation in CS project coaching events; Impact Assessment of CS (UNITAR) with a focus on the requirements of NSOs as end-users of CS data for SDG monitoring. At the core of the project is Project Deployment (UNIGE) based on a novel innovation cycle called GEAR (Gather, Evaluate, Accelerate, Refine), which runs once a year.

The GEAR cycles involve online selection and coaching of citizen-generated ideas for climate action, using the UNIGE Open Seventeen Challenge (O17). The most promising projects are accelerated during a two-week in-person Challenge-Based Innovation (CBI) course. Top projects receive further support at annual SDG conferences hosted at partner sites. GEAR cycles focus on specific aspects of Climate Action connected with other SDGs like Gender Equality.

Grant Agreement description of the deliverable

Parts in bold describe the deliverable content.

WP2 - Enhancing Citizen science tools and methodologies [Months: 1-36]

CSIC, UNIGE, POLIMI, UNITAR

The work in each of the research fields will evolve through 3 different phases: (1) design and early prototyping, (2) beta release and (3) **final release**. The results from each phase will be demonstrated in a public event at the closing ceremony of each GEAR cycle, i.e., at M11, M21, and M30.

The focus will be on enhancing existing open-source Citizen Science tools (see Section 1.3.7.4) with AI components to work with large-scale datasets, improve data quality and speed up data analysis. The tools will be used during WP3 (specifically: Task 3.1) GEAR cycles in order to validate them, and support participants on the creation of CS projects.

D2.3 : CS tools final release [36]

CS tools final release.

Purpose and scope of the deliverable

This deliverable contains a description of the works carried out during the third and final year in Work Package 2 (WP2) “Enhancing Citizen science tools and methodologies” and a detailed description of the final releases of the Citizen Science Solution Kit (CSSK) tools which are the products of the work package.

1. Introduction

1.1. The Crowd4SDG Citizen Science Solution Kit

The Citizen Science Solution Kit (CSSK) is a set of open-source software tools maintained by the Work Package 2 (WP2) members of the Crowd4SDG project for developing and running crowdsourcing projects. The tools enable you, as a citizen scientist, to conceive, design, launch, and manage your online crowdsourcing project, and analyze the results and your community without any coding skills. CSSK also comprises powerful tools to extract relevant information from social media as it is becoming an increasingly important source in citizen science (CS).

Depending on your needs, you may want to use the whole set or a selection of tools in CSSK. In a full-fledged scenario, for example, you deliberate with your community on how to tackle a specific problem in your area on *Decidim4CS*. You may then decide to gather evidence related to the problem and you download messages and images shared on social media and filter relevant content with *VisualCit*. You create a mobile app with *CSLogger* and ask your community to input first-hand data from the area. Later, you upload all image and text content to *Citizen Science Project Builder* and the volunteers classify the images, also leveraging the text, e.g., by labeling the severity of the problem you are tackling as seen on the photo. Then, based on crowdsourced annotations, you compute the consensus on each image by *Crowdanalysis* and analyze the performance of your community to make more informed decisions.

Compared to state-of-the-art CS tools (Section 2.1), CSSK provides end-to-end services for a CS project that can be customized by citizen scientists, and not only by tool owners. After the third and final year of the Crowd4SDG project, in this deliverable, we report the final release of CSSK, describing the recent enhancements in the tools and the case studies which demonstrate CSSK's efficiency and impact on citizen science. The case studies include use cases generated by the GEAR innovation Cycle led by WP3

1.2. Structure of the document

This document is organized as follows. Section 2 starts with comparing CSSK to state-of-the-art CS tools, and describes in detail the enhancements in CSSK tools for their final releases. Section 3 presents the case studies in real-world scenarios that we tested CSSK in. Subsequently, Section 4 summarizes the three-year work to fulfill the WP2 tasks defined in the project's Grant Agreement. Then, Section 5 relates the work in WP2 with the remaining work packages of the project. Finally, we draw conclusions in Section 6 as we finalize the Crowd4SDG project.

2. The Citizen Science Solution Kit

In this section, we provide a detailed overview of the current state of the Citizen Science Solution Kit (CSSK). We start by presenting the literature on leveraging citizen science and social media to pursue the sustainable development goals (SDGs) defined by the United Nations (UN), and we compare CSSK with state-of-the-art in crowd-sourced social media analysis. Then, sections 2.2–2.10 provide the details of the improvements of each tool in the CSSK.

2.1. State of the art

CS has been proposed as a way to support the collection of data for reporting on the SDG indicators. In their paper “Mapping Citizen Science contributions to the UN sustainable development goals” [Fritz2019], the authors perform a detailed analysis of existing projects already covering some of the SDG indicators and suggest potential areas for further development of new initiatives, including those which address climate change. These are areas where collection of information is difficult with more traditional methods, such as household surveys and censuses. The United Nations Statistics Division (UNSD) also mentions this, which highlights how achieving and monitoring the SDGs require innovative ways to produce and apply data and statistics in addressing the multifaceted challenges of sustainable development (<https://unstats.un.org/sdgs/report/2017/harnessing>).

In the literature, social media is considered a form of crowdsourcing and not necessarily an activity of CS [Shanley2019]. However, the effect of combining the two approaches in a multi-dimensional framework has been studied in different research communities [Franzoni2022]. Also, the combination of automatic analysis of images contained in posts and the collaboration of human-based computing has been discussed by several authors in the literature [Havas2017, Alam2018, RaviShankar2019]. For automatic analysis of images, many recent approaches are based on AI and, in particular, neural networks [Imran2020, Pennington2022]. Hybrid deep-learning and crowdsourcing approaches have been analyzed as well [Anjum2021].

Social media presents a huge potential to provide key information for monitoring SDGs and supporting decision makers. Thanks to their ability to provide timely information about ongoing events [Imran2015, Adrot2022, Pennington2022], their role has been studied in several contexts, e.g., earthquakes [Sakaki2013], demonstrating how this timeliness can help prevent further losses. The EU Crowd4SDG project has explored the potential of social media to provide key statistics and local indicators [Deliverable D5.1] based on lessons learned from case studies for SDG monitoring.

Many research papers have been published on the topic of social media for emergencies, including a special track at the International Conference on Information Systems for Crisis Response and Management [Karray2022]. An analysis of the opportunities and challenges of such applications [Wiegmann2021] revealed how several challenges are related to the need of geolocating posts to leverage their contents, and the issues of analyzing posts in multilingual contexts [Lorini2019, Scalia2022]. Several approaches for geolocation and geocoding have been proposed in the literature [Middleton2014]. Precision in the location is important, but difficult to achieve as one can not rely on native locations of, for example, tweets. This is due both to their scarcity (most users keep location information disabled, as per the default setting), and because the location of the posting is often different from the location of the event occurring during an emergency. Many of these challenges have been addressed by the combination of automatic analysis and crowdsourcing techniques. Automatic analysis enables the processing of millions of tweets in an almost real-time

fashion, while crowdsourcing is used to curate the final dataset by validating and complementing the information, overcoming the limitations of automatic techniques (see for example [RaviShankar2019]).

For emergencies, some community platforms have been developed over the years. The first and most known example is Ushahidi [Okolloh2009], which allows crowdsourcing information about an ongoing crisis and displaying the information on maps. Other social media, e.g., Facebook with the Crisis Response Service (<https://www.facebook.com/about/crisisresponse/>), provide an open space to collect awareness information about emergency situations and notification services. Generic Citizen Science and Crowdsourcing tools such as Zooniverse (<https://www.zooniverse.org/>) and Amazon MTurk (<https://www.mturk.com/>) allow the design and implementation of interfaces to enable the crowd to perform social media data analysis. In these cases, the social media data needs to be updated and can not be mined from those platforms.

In this report, and in the Crowd4SDG project in general, the focus is on combining automatic analysis of images present in social media posts, with geolocation of the posts and crowdsourcing, inserting a human-in-the-loop approach in all phases. The goal is to refine the automatically extracted information and improve the selection of relevant posts. The CSSK presented in this study enables the implementation of such “analysis pipelines” and the easy combination and configuration of tools.

To illustrate the main differences and characteristics of the above mentioned approaches, Table 1 summarizes and compares some of them, representing the state of the art in the field [Bono2023]. They include tweet analysis, combining human and machine computing, emergency image extraction from Twitter (E2mC), community platforms, and generic CS and crowdsourcing platforms. The comparison is done by analyzing each approach along four dimensions:

- *Autonomous pipeline preparation*: as CS initiatives are often created bottom-up, organizers may want to develop their data collection and analysis autonomously. This dimension looks at the availability of configurable functionalities and services (e.g., filters) that allow the project organizers – ideally without the need for coding – to collect, filter, clean and analyze social media data.
- *Data collection from social media*: this dimension looks at the ability of the system to dynamically crawl social media with keywords selected by the user, automatically handling multimedia content, and visualizing the collected data, including spatial data.
- *Crowdsourcing*: examines the availability of crowdsourcing tools and methods to allow both the collection and evaluation of human contributions and feedback.
- *Flexibility and scalability*: refers to the ability of the platform to easily adapt to specific – and sometimes quickly changing - project requirements, including the ability to scale up to a large number of users. In the table, we focus on how flexible the service provisioning is with respect to its deployment on the computing infrastructure, provisioning as a service by the tool owner or a third party, and availability of the code as open source.

Also highlighted in the table is the possibility by Project Organizers (PO) to configure and execute their own project by acting directly on the available features and services, versus what only Tool Owners (TO), defined as producers or managers of tools and platforms, can do.



The conclusion from this overview of related work is that the CSSK approach brings the solution closer to citizens and communities, and provides a truly unique combination of automatic and crowdsourced data analysis that allows for easy creation of powerful projects, without the need for programming or technical expertise. For more details please refer to [Bono2023].

Approach	Description	Autonomous pipeline preparation		Data Collection from Social Media		Crowdsourcing		Flexibility and scalability
		T	P	T	P	O	O	
<i>Automatic Tweet Analysis Platforms</i>	Predefined services and tools for specific purposes, usually emergencies, e.g., [Adrot2022, Scalia2022]	Pipelines are predefined	✓	Crawling with keywords with either text or images considered	✓	✓	Sometimes used for quality evaluation	Depends on infrastructure
		Focus on pre-trained ML classifiers, focus on images and geolocation	✓					
<i>Combining human and machine computing</i>	Automatic classifiers incrementally trained for a specific context, e.g., in AIDR [Alam2018, Imran2020, Pennington2022]	Pipelines are predefined	✓	Crawling with keywords with focus on images	✓	✓	In some platforms used for classification evaluation	Depends on infrastructure
		Leverage on human feedback for improving ML classifiers	✓					
<i>E2mC</i>	Automatic collection of images from social media combined with crowdsourcing for emergencies [Havas2017, RaviShankar2019]	Pipelines are predefined	✓	Crawling with keywords with focus on images	✓		Used for classification by citizen scientists	Depends on service provider
		Focus on images and geolocation	✓					
<i>Community platforms</i>	Services provided for collecting information from involved citizens, e.g., Ushahidi [Okolloh2009], Facebook	Predefined services provided for general use	✓	Direct collection of posts within the platform	✓		Used for information collection	Depends on service provider



	Crisis Response								
<i>Generic Citizen Science and crowdsourcing platforms</i>	CS platforms for involving citizens in problem solving tasks, e.g., Zooniverse, Amazon MTurk	Pipelines are configurable		✓	Can be uploaded as dataset	✓	✓	Used for classification by citizen scientists	Depends on service provider
		Single service provided for general use	✓	✓					
<i>Crowd4SDG Citizen Science Solution Kit</i>	Services provided to enable citizen scientists configure their own project and collect data from social media	Pipelines are configurable	✓	✓	Crawling with keywords focus on text, images and geolocation	✓	✓	Powered by CS with advanced and tailorable consensus mechanism	Extensible, scalable, and reproducible (Open Source)
		Multiple services provided for general use	✓	✓					

Table 1. Comparison of approaches for crowd-sourced social media analysis. For the management of the tools/platforms, the role of Tool Owners (TO) and Project Organizers (PO) is highlighted.

2.2. Social media analysis pipeline

The primary focus of the CSSK pipeline is to rationalize the various aspects of social media analysis, combining them with a human-in-the-loop approach. Information extraction from social media can serve many use cases, since social media data are usually abundant, pervasive, ubiquitous, real-time and rich.

A critical challenge for the information extraction task is the ability to isolate relevant contents, up to some definition of relevance. Usually raw social media datasets contain a low percentage of relevant information [Negri2021, Scalia2022]. Discovering infrequent items in a high volume of data, usually with stringent constraints on timeliness and recall, calls for an automated data filtering approach. We utilized a data processing pipeline as a reference framework, where the pipeline instances are modeled after the specific use cases. These pipelines are split in the phases of data preparation and data analysis (see Figure 1).

In the data preparation, social media data is crawled using appropriate queries. Contents are then filtered in order to retain relevant information only. Filtering actions are organized in a library, and range from duplicated detection to automated image classification, usually performed with convolutional neural networks (e.g., [Asif2021]). Each use case requires a specific selection and configuration of filtering actions, as events have distinctive characteristics and occur in different locations around the world. In [Bono2022a], the configuration of this phase by CS is possible without technical assistance, using a library of available cleaning and filtering components. Filtered contents are then annotated with further information, such as inferred tags about the posts, and possibly its geolocation.

Prepared datasets are then ready for contributions from citizen scientists. For example, CS can refine the automatic filtering and annotations performed automatically, or perform

inherently difficult tasks that are unsuitable for machines. Then, in the Data Analysis phase, the data analysts explore hypotheses and distill information. Typical use cases include the retrieval of images from specific locations, thematic mapping in order to highlight the impact of ongoing events and the design and implementation of alerting systems.

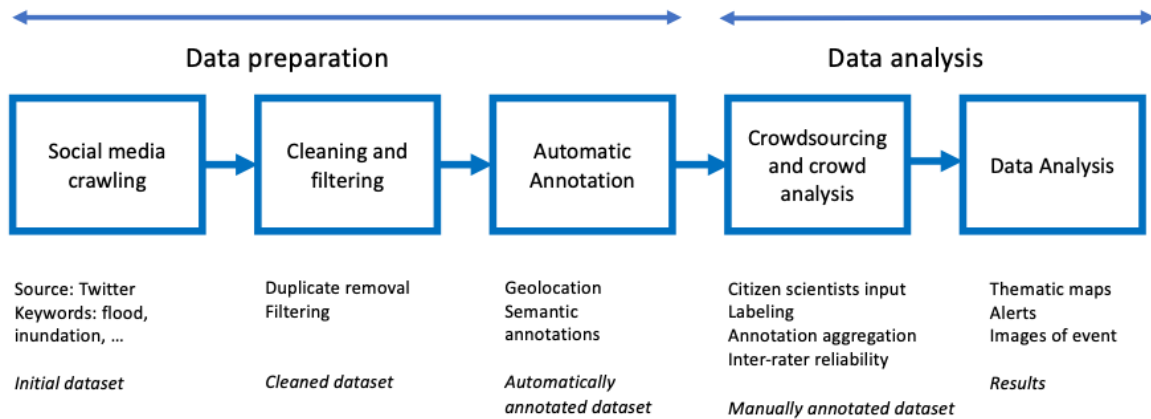


Figure 1. Process of social media analysis using the combined “human-computing” approach.

2.3. Conceptual framework

The CSSK has been developed as a set of flexible tools able to support the collection, curation, and analysis of data for different goals and applications. The kit supports both automatic computation and citizen participation in the analysis process. In particular, Crowd4SDG has been using the CSSK as an integrated framework for social media analysis based on three of its components: VisualCit, Citizen Science Project Builder (CSPB), and Crowdanalysis. The framework supports analysis pipelines that progressively transform a raw dataset of information collected on social media into a curated—i.e., cleaned, processed and annotated—set ready for data analysis, following the approach illustrated in the previous section. In the curated dataset, posts have been cleaned from irrelevant information, marked with automatically extracted geographical information, and further localized and annotated by citizen scientists via crowdsourcing.

The three components of the CSSK have different roles and functionalities. The goal of VisualCit (short for Visual Citizen Scientist) is to enable citizen scientists and data analysts to query social media in order to obtain data and to interactively define automated filters and annotations to clean and refine the selected data. This allows reducing the number of irrelevant posts that will then feed into the manual analysis. CSPB allows the creation of crowdsourced data analysis projects where participants perform various analysis tasks on existing digital data (text, images, etc.). This data may originate from citizen-generated content on social media.

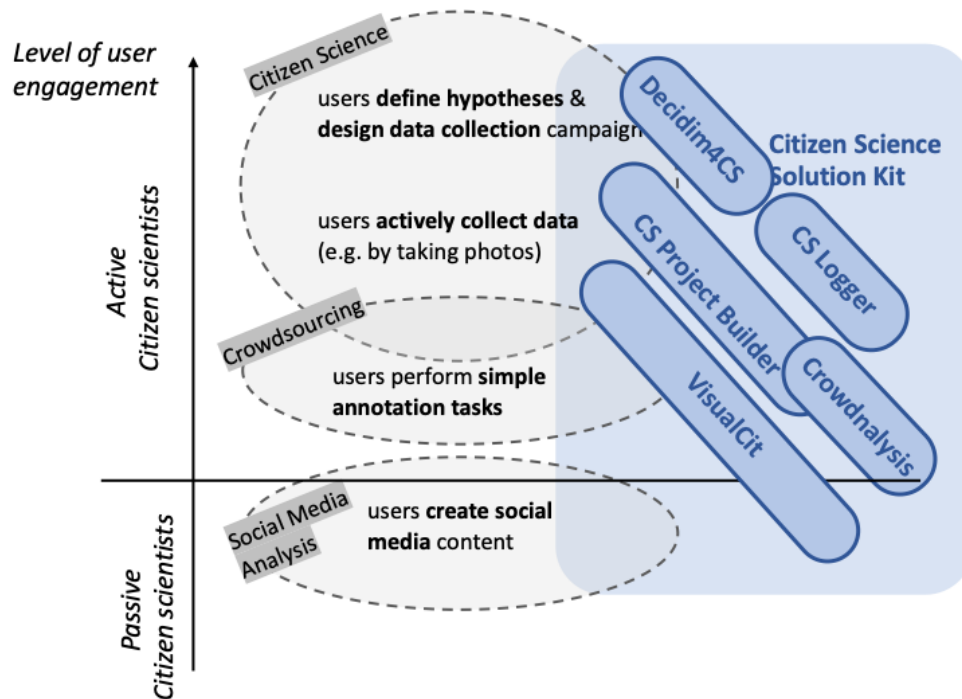


Figure 2. Citizen Science Solution Kit levels of engagement.

Tasks performed can include confirming the relevance of selected images/texts, identifying information in images/texts and contributing new information by assigning predefined labels or annotating them with free text, and so on. As crowdsourced information is known to be subject to possible biases or errors by the contributors, an overall analysis of the crowd behavior is necessary to identify possible problems and to aggregate and evaluate citizens' contributions. The analysis of crowdsourced information is performed by the Crowdanalysis tool. Each of the CSSK components is described in detail in the following subsections.

One of the key elements to consider when designing tools based on citizen participation is their desired "level of engagement", i.e., how actively citizens are supposed to provide support, process information, and suggest solutions. For geographical applications, Craglia and Shanley propose an analysis framework [Craglia2015] in which candidate tools and initiatives are classified according to the level of required engagement of citizens and the implicit/explicit nature of the geographical information. Figure 2 uses the level of engagement, as defined in [Craglia2015], to analyze the CSSK components and their associated activities. In Crowd4SDG, we consider both active and passive citizen participation. Purely passive participation consists in providing information about ongoing events by posting on social media without being aware or interested in the possible use of the information. In this case, the CSSK can process the information without further citizen involvement. In other scenarios, we consider a more active participation, where citizen scientists can contribute by analyzing, validating, and annotating information, or contributing new data. In VisualCit for example, the focus is on collecting information from passive, citizen-generated content. However, citizens can also actively participate in defining the type of information to be collected. VisualCit can also be used as an active collection tool, in cases where specific hashtags focusing on a given event or situation are used [Grasso2016].

We note that we carried out Crowdanalysis' beta-release integration into CSPB as an on-demand service as reported in D2.2 (pp.12-14). Crowdanalysis-Service is activated by a

button on the CSPB frontend and automatically computes and exports the consensus for tasks in a project CSPB via Crowdanalysis. Citizen Science Center Zurich's technical staff will integrate Crowdanalysis-Service in CSPB production environment, with CSIC's guidance when needed.

2.4. CSSK user roles

To clarify the possible ways to interact with the tools during the pipeline process, we distinguish two main roles in the use of the CSSK: CS project manager and “citizen scientist”. CS Project Managers could be single individuals (e.g., researcher or citizen scientist) or teams putting forward an idea and implementing its corresponding data analysis protocol. The CS Project Manager is responsible for defining and refining the goals of the project, the type of analysis to be performed and the desired outcome.

CS Project Managers can play three different specific roles:

- *Organizer*: When CS Project Managers have set their goals, they must configure the tools and data preparation pipeline, select the analysis scenario, the volume of data to be analyzed and the frequency of the analysis. If crowdsourcing is part of the scenario, the manager must identify citizens' specific tasks.
- *Executor*: Once the data preparation pipeline has been configured, the CS Project Manager can execute it starting from an initial set of (raw) data. This is an iterative process: the assessment of intermediate results might initiate a reconfiguration of some of the tools to improve quality, or even a revision of the initial goal.
- *Analyst*: In case of crowdsourced tasks, the CS Project Managers should analyze the results of the crowd. This includes the identification of the most appropriate model for computing consensus, and the analysis of the behavior of the crowd, looking for anomalous or opportunistic behavior that might invalidate the results.

With the term “citizen scientist”, we indicate anybody contributing to the preparation or analysis phases, i.e., anybody involved in performing tasks or providing information useful for the analysis goals. A visual interface is provided to support configuration operations and quality-improving actions on the VisualCit and CSPB tools. In VisualCit, execution can be performed via an interactive interface for the initial samples, whereas execution on large volumes of data is performed using a service-based architecture. In CSPB, an interactive interface is provided both for the setup and for the execution of tasks by the crowd, and to extract data for analysis by the CS Project Manager. Both components provide a set of web services that can be invoked via Application Programming Interfaces (APIs) from both interactive and programmatic interfaces. Once the data are prepared with VisualCit, crowdsourcing can be performed. A common data exchange format is defined for the three tools based on CSV files.

2.5. Selecting and filtering data: VisualCit

VisualCit is a social media processing tool aimed at exploring and building data processing pipelines. The goal of these pipelines is to extract data that is relevant for a given objective from social media streams. The relevance is estimated on media attachments of the posts, using a configurable combination of decisions, based on machine learning classifiers. Data processing actions currently supported by VisualCit are i) *ingestion*, which collects data from social media with configurable criteria; ii) *cleaning*, mainly to reduce the number of duplicate posts and images; iii) *selection*, to select posts that are relevant to the goals of the social media data analysis to be performed (e.g. selecting all images containing a particular type of

object); and iv) *augmentation*, to automatically derive further information (such as the location) for the posts being analyzed.

VisualCit is general purpose, pluggable across different data sources, easily extensible, and available both with a graphical interface and as a web service. The graphical interface supports the configuration phase in a visual manner. Pipeline configurations with satisfactory outputs can then be applied to large amounts of data through the web service.

The requirements and the implementation of VisualCit are described in D2.1 [Deliverable 2.1, pp. 17-20]. The first year of the project was devoted to designing and testing the VisualCit framework described therein. The implementation consisted of developing a set of components that can be combined in a flexible way and can be orchestrated as web services. The implementation was focused on developing the backend components of the system.

We extended the functionalities of VisualCit in year two, including both the backend and frontend components, as described in D2.2 [Deliverable 2.2, pp. 15-17]. VisualCit was harnessed with a visual interface. New actions for filtering images and performing geolocation of the posts were added. Input and output facilities, both for data and pipeline configurations, were added.

During the third year, several enhancements have been performed in order to get a better experience and performance of the system. The user interface has been redesigned. A graphical tool to help the extraction of queries from the crawled data was added. On the backend, some of the filters were updated with more modern versions. For efficiency, the Graphics Processing Unit (GPU) processing was also introduced in the backend.

VisualCit was utilized in prototypes from GEAR cycle projects, in order to explore the potential of collecting social media data for specific goals. The Donate Water project (Cycle 2) evaluated the ability to collect evidence about the status of water sources in Nigeria. The Let's Clean Up project (Cycle 3) plans to integrate evidence collected from social media in their pilot using VisualCit.

2.6. Building crowdsourcing projects with CSPB

The Citizen Science Project Builder (CSPB) is a web-based tool that allows a CS project leader to enroll a volunteer community to perform complex data classification tasks. These may take different forms, from classifying images to transcribing documents, from collecting samples to taking pictures or recording other forms of digital data (e.g. audio and video clips). Project creators can set up their project via a web interface that requires limited technical knowledge, and little or no coding skills.

In year 1 of the Crowd4SDG the CSPB evolved by including new functionalities and enhancing the user experience to support better community engagement. In 2021, two major functionalities were added: the automation of the computation of consensus (see next section) and the integration of geolocation. Geolocation of social media information (ie. finding the geographical coordinates of the image or post extracted from social media) is a key process to make the information actionable, and a template to set up projects to perform such tasks was implemented.

In year 3 of the Crowd4SDG project, the CSPB has been used extensively both by students testing their projects and by teams of students testing their projects and by researchers. Notably, the EC Joint Research Council (JRC) has used the platform to experiment with new

crowdsourced methodologies for Copernicus' images extraction and analysis during natural disasters.

2.7. Social media consensus analysis with crowdanalysis

When the volunteers contributing to a CS project in CSPB complete the classification tasks (e.g., labeling the severity of damage in photos) on the platform, the individual annotations of the crowd members need to be aggregated to achieve a consensus on each task. Majority Vote is the most common aggregation method where the consensus on a task is the most voted class, and each annotator's vote has an equal contribution. This method counts on the wisdom of the crowd, but it cannot tackle situations such as when i) there is uncertainty of how many annotators are required to obtain reliable results, and ii) some annotators perform better on some tasks than others.

We developed the Crowdanalysis software library [Crowdanalysis2022] to address the critical needs that arise in the planning of a crowdsourcing project and the analysis of crowd-sourced data in CSPB. Crowdanalysis incorporates probabilistic consensus models that estimate individual annotator error rates for a given set of tasks even when the ground truth is unavailable. These models enable Crowdanalysis to weigh the contribution of each annotator and thus, yield a more reliable consensus.

The details of the requirement analysis and the implementation of Crowdanalysis were reported in [Deliverable 2.1, pp. 15-17]. In year two, we integrated Crowdanalysis into CSPB as an on-demand service to automate the annotation aggregation process in CSPB by leveraging the advanced consensus methods provided by Crowdanalysis. Owing to this integration, CSPB users can export the consensus on tasks annotated by the crowd in their project, upon a single click. The details of the integration of Crowdanalysis-Service into CSPB were reported in [Deliverable 2.2, pp. 12-14].

Since year two, Crowdanalysis is being distributed as an open-source software library at The Python Package Index (PyPI) repository of software: <https://pypi.org/project/crowdanalysis/>. The dissemination at PyPI eases the discovery and installation of our tool by the software development and scientific communities. The details of the distribution are reported in [Deliverable 2.2, p. 12]. We will be publishing new versions of Crowdanalysis as we add new features to the tool. Currently, we are working on methods for the aggregation of geolocation annotations and efficient strategies to assign annotation tasks to annotators based on their expertise in specific geographic regions. More details of this recent work are given in Section 3.7.

2.8. Decidim4CS

Co-creation of a CS project requires citizens and scientists to self-organize, propose and discuss ideas, schedule meetings, conduct surveys, and much more. Decidim4CS is a digital platform for participatory citizen science and constitutes the deliberation technology used in the Crowd4SDG project (<https://decidim4cs.ml/>). It allows citizen scientists to take an active role in CS project design and execution by making proposals, participating in ongoing debates or starting new ones, attending online meetings, making collective and democratic decisions through voting, and monitoring the implementations of these decisions.

Decidim4CS is based on Decidim ('we decide' in Catalan), a free open-source software originally created by the Barcelona City Hall as a participatory democracy platform for cities and organizations. The analysis of requirements and design details are reported in

[Deliverable 2.1, pp. 20-24]. The enhancements to Decidim4CS in the second year were reported in [Deliverable 2.2, p. 17].

D2.2 marked the end of WP2's Task 2.1 –deliberation technologies for citizen science– that Decidim4CS is related to. However, in year three, we maintained the platform up-and-running and continued to promote it in the GEAR-3 cycle. The team Let's Clean Up, one of the two finalists, adopted Decidim4CS as the community deliberation platform to address the waste management problem in Rome. In their prototype (<https://decidim4cs.ml/assemblies/lcu>), they created a debate for each of the 15 municipalities and another one for discussions to improve the project itself.

In year three, we also initiated a proposal to integrate the “debate aggregation” algorithms developed within this task into Decidim [MetaDecidim2022a]. At the date of writing, our proposal is under consideration by the Decidim community.

2.9. CSLogger

Citizen Science collection projects engage people in the acquisition of data. In the case of digital data, this happens very often via the use of smartphone applications. CS Logger is an open-source data collection platform that makes it easy for anyone to build and configure customized mobile applications for their CS projects.

It features all traditional and domain-generic Citizen Science activities, such as the collection of geo-located images, video, and audio, together with surveys, quizzes, and other tasks to engage the crowd in the collection and description of all sorts of digital data. Via a web-based dashboard, project creators can choose from a menu of functionalities and build a tailored app without prior programming or design experience.

Several enhancements were made to the original code in year 2 of the Crowd4SDG project, including the possibility to create “public” projects (i.e. projects where everybody can contribute and there is no need for a personal invitation), the possibility to also create projects accessible via a web browser, and an enhanced interface for a more detailed description of the process to contribute. Following these modifications, project creators can use CS Project Builder and CS Logger in a pipeline, where data collected with a smartphone app can be imported and made available for analysis in a web interface.

In the third year of Crowd4SDG, student projects have used the tool to test and pilot ideas and processes of data collection, for example the projects related to mapping presence and functioning of water sources in rural areas. Other CS projects, beyond the ones generated by students, and including new EU projects, are exploring the tool and including its use in their proposals and planning (i.e. EU HealthFerm).

2.10. CoSo (Collaborative Sonar)

In the GEAR cycle 1 report (D4.3), we introduced the CoSo platform [CoSo2021, Tackx2021] for collecting self-reported data on collaborations and task allocation. CoSo was developed to overcome the setback introduced by the shift to a fully online program due to the COVID pandemic, and the subsequent impossibility to implement a sensor-based contact-tracing app to monitor *in-situ* team interactions using smartphone sensors (Bluetooth, Wi-Fi and GPS). Following these setbacks, we adapted our strategy by developing a standalone platform for the active monitoring and contextualization of team interactions using self-reports and surveys.

CoSo is composed of three main applications. First, a backend gathers all user data into a structured database and provides an API, allowing for a simplified management and data extraction from the database of GEAR participants. Second, a mobile application for Android and iOS and its supplementary web interface allow users to journal tasks, receive notifications, fill surveys, and gather immediate insights about their logged data. This encompasses the main desired features from MindLogger (a development of the Child Mind Institute in New York, USA), and provides an alternative pathway for contact-tracing, replacing passive bluetooth interaction information, limited to on-site interactions and yielding quantitative yet non-contextualized insights, by active self-reports on participant interactions, generalized to both on-site and online interactions and contextualized around shared tasks. Finally, a frontend web application allows for data visualization, team management, survey and communication creation by the research team as well as survey filling by users. This last part was finalized during GEAR 2 and introduced in [Deliverable 4.6].

During GEAR 3, we generalized the CoSo platform to adapt to any future data collection by allowing administrators to create new programs (beyond GEAR cycles), add participants, and monitor the interactions. The CoSo app was made as agile as possible, in order to adapt to different needs from different participatory programs. For example, the visualization dashboard has been designed to allow the representation of various types of survey questions rather than specific, hard-coded questions, so that they can fit the particular inquiry of the survey designed.

3. Validation with case studies

This section presents the case studies we used to validate the performance of the Citizen Science Solution Kit (CSSK) in real-world scenarios. We start by introducing the methods and quality dimensions employed in the validation. Then, we give Covid-19-related two case studies that are of particular importance as the pandemic coincided with the start of the project. Afterward, we give four more case studies related to natural disasters.

3.1. Validation methods

The quality of the results obtained by crowd-based data analysis depends on different factors: the quality of the dataset obtained from the previous preparation and analysis phases, how the task has been managed and described, and the skills and expertise of the citizen scientists. In our case studies, we based our validation on the quality of the resulting data. Adopting this end-to-end approach, we can evaluate the quality of the result by considering the effect of other components: the filtering components in VisualCit, which can also present quality issues, as they are based on models trained with machine learning methods since configuration parameters of these models can have an impact on the final results, as well as the quality of the crowdsourcing tasks assigned to the crowd and the quality of the results of crowdsourcing activities as described in Section 2.7.

In order to assess this quality several methods can be applied:

- *Individual*: The role of humans in crowdsourcing can be extended to the quality assessment phase. The accuracy of a given output is evaluated by individuals, e.g., by citizen scientists or external experts.
- *Group*: The assessment is performed by a group of people (typically citizen scientists), e.g., through voting.
- *Computation-based*: This includes assessment methods that can be performed by a machine without the involvement of humans.
- *Validation datasets*: Validation can be performed by comparing the obtained results with external sources focusing on the same problem, e.g., surveys in the field.

The computation-based methods can support the evaluation of the quality of results from different perspectives and consider different quality dimensions:

- *Accuracy*: This can be computed considering a given ground truth and suitable comparison operators to evaluate the similarity between the obtained results and the desired values. The assessment of accuracy can be precise depending on the reference source considered.
- *Timeliness*: Some data analyses have stringent time constraints. Timeliness measures the temporal validity of the input data; only not outdated values should be considered. This means that the results should be produced within a given time after the event occurs.
- *Provenance*: The reliability of the input data increases if information about acquisition methods, processing steps, and the way in which crowdsourcing has been performed (e.g., crowd characteristics, redundancy, and aggregation methods) is available. It is important in the evaluation of the results to obtain information on the data

provenance: the more input data is reliable, the more trustworthy the analysis results are.

A quality control methodology for datasets derived from social media has been studied in Crowd4SDG to assess whether extracted data can be used as non-traditional data sources by national statistical offices (NSOs) [Deliverable 5.2]. A scoring approach was proposed to evaluate the quality of the dataset. In addition to the validation criteria mentioned above, a few new criteria were added for validation focusing on the data production process, confidentiality, and impartiality. The results of the validation for case studies are discussed in the next subsections. For a comparative overview of the validation methods used, three of the case studies are listed in Table 2. In the table, the main validation criteria are discussed for the results of each of the CSSK components used in the case studies. The results derived from the quality control methodology are reported in the Data quality assessment column. Finally, a general end-to-end evaluation is given under the General evaluation column and the main issues mentioned under Problems.

Case study	Goal general/specific	Source	VisualCit task	VisualCit configuration and validation dimension	Crowdsourcing tasks	Crowdsourcing configuration criteria	Crowd analysis	Crowd validation	Data quality assessment of resulting dataset	General evaluation	Problems
Albania earthquake	general goal: collect images with damage assessment of buildings use cases: A, B specific goal: relevant images	cleaned images dataset from Twitter 907 images	eliminate non relevant images	64% non relevant discarded, 83% precision, 90% recall and 82% overall accuracy	grading tasks tool: Amazon MTurk redundancy: 10	time to complete a task difficulty of task	applied (all methods)	TPR: 98%, TNR: 52%	metadata are needed	annotator error rates remained fairly the same; suggesting we can trade crowd time for higher recall	a qualified crowd would be preferable
COVID-10 social distancing	general goal: derive indicators about social behaviors in COVID-19 by country use case: C specific goal: annotated geolocalized images	Twitter posts with images 500K/week (three weeks)	eliminate non relevant images geolocate	0.5% classified as relevant and geolocated accuracy of geolocation: 84% accuracy of filtering: 92-99% (depending on applied filter)	question answering tool: Project Builder redundancy: 3	high volume: sampling needed non expert crowd sufficient difficult of finding motivated crowd	Majority Vote; Dawid-Skene [Dawid1979]	crowd accuracy for social distancing: 71% (w.r.t. crowd consensus by Dawid-Skene method without ground truth)	comparison with independent survey proved satisfactory; dataset evaluation positive, more multilingual approach requested	sampling reduced significantly the amount of analyzed data	volume crowd recruitment
Thailand floods (see subsection TriggerCit)	general goal: timely alerts (<24h) of floods monitoring social media use case: E specific goal: spatial representation of the event	Twitter posts, 4 million in two days	extract posts with images, filter out non relevant images, geolocate	completeness of the results improved with text analysis (45% increase)	not applied due to required time constraints	-	-	-	comparison with independent survey proved satisfactory; dataset evaluation positive, general multilingual approach required	coverage of regions not uniform compared with Tweet native geolocation only, accuracy improved and less bias on capital city	volume, timeliness, relevance

Table 2. CSSK Case studies

3.2. Covid-19 social distancing

This case study, developed in 2020 during the initial phases of the Covid-19 pandemic, aimed to derive indicators of social distancing behavior and face mask usage in different countries. This case study employed a complete data preparation pipeline. Selected raw data were captured from social media from May to August 2020, crawling Twitter with generic COVID-related keywords, and then filtering posts with VisualCit web services, selecting only images that are classified as photos (excluding memes, maps, drawings, etc.), in outdoor spaces, and containing at least two persons. When a native geolocation was not present in the tweet, the posts were geolocated using the CIME algorithm [Scalia2022]. Crowdsourcing performed with the CS Project Builder was used to confirm the relevance according to the above-mentioned criteria and the automatic geolocation of the images, and to assess whether face masks were correctly used and social distancing rules followed. The majority vote was used to compute the consensus on the crowd annotations. The resulting images were aggregated to build behavioral indicators on a national basis. The case study results are published in [Negri2021] and reported in [Deliverable 2.1, pp. 25-28].

3.3. Covid-19 vaccines and menstruation

Stability of the menstrual cycle is a key indicator of health, and its alteration can affect the physical, emotional, sexual, and social aspects of menstruating individuals' lives. This case study built on top of the work of Edelman et al. [Edelman2022] who showed a statistically significant increase in cycle length after vaccination against COVID-19. Based on that, we decided to study the potential association between vaccination time and change in cycle length, using citizen science to obtain scientific evidence to answer the question: Is the phase of the menstrual cycle relevant when getting the Covid-19 vaccine?

The work was developed as a cooperation between (i) Lunar App, an Argentinian startup which has its roots in an open innovation methodology which influenced the GEAR cycle and took place prior to Crowd4SDG, (ii) Hospital de la Santa Creu i Sant Pau in Spain, who contributed the gynecological knowledge, (iii) Aquas, the quality agency of the Catalan Health System, who provided clinical data analysis knowledge, and finally (iv) University of Geneva and (v) CSIC who were coordinating the efforts and contributing with artificial intelligence knowledge. We analyzed data collected by the menstrual cycle tracking smartphone application Lunar App. This application allows users to track their menstrual cycle and menses, recording the beginning and end dates, pain intensity, blood loss quantity during menses (more, equal, or less than usual), and their COVID-19 vaccination status.

The database contained 28,876 users and 162,529 cycles. The distribution of the percentages of the users' age ranges (years) was as follows: 18 to 24, 11.85%; 25–34, 49.15%; 35 to 44, 28.56%; 45 to 54, 8.31%; other, 2.13%. We filtered the database, keeping only users who had reported their vaccination status and at least 5 consecutive cycles. We considered the first doses or monodoses of the vaccine for the analysis and removed incomplete or wrong data. After this filtering process, we ended up with 371 users and 1,855 cycles registered between September 2020 and February 2022. The relatively small size of the final sample is caused by the imposed restrictive inclusion and exclusion criteria to ensure the maximum attainable data quality.

For analysis, we employed the self-controlled case series method [Petersen2016]. We observed an increase in the median cycle length of 0.5 (0.0–1.0) days (P value <0.005) for all individuals, with 8.08% of the individuals having an increase of 8 or more days, which is considered clinically significant [Male2022]. In addition, we observed no significant variations in the percentages of cycles with abnormal blood loss or pain intensity.

Furthermore, the stratified analysis showed an association between the phase of the menstrual cycle of the individual at vaccination time and cycle length change. Thus, individuals vaccinated during the follicular phase showed a median cycle length increase of 1 (0.0–1.0) day (P value <0.005), with 11.82% of the users having an increase of 8 or more days. The presented results suggest considering the phase of the menstrual cycle for the design of future COVID-19 vaccination policies and recommend vaccination during the luteal phase (i.e. days 14-28 of the menstrual cycle).

This case study shows the potential of obtaining top-level scientific results by analyzing citizen contributed data and has been published in the most relevant scientific journal in the discipline of Gynecology [Velasco-Regulez2022].

3.4. Albania earthquake

In this case study, we aimed to provide a working example of how we fit the conceptual and mathematical framework behind Crowdanalysis to real-world data, and demonstrate how we could benefit from our software library in an emergency scenario, where we refer to crowdsourcing. The imagery dataset we worked with was the courtesy of the Qatar Computing Research Institute, which contains social media images posted for the 2019 Albanian earthquake. The framework and case study results are published in [Cerquides2021a] and reported in [Deliverable 2.1, pp. 28-30]. We demonstrated in this case study that the probabilistic framework behind Crowdanalysis enables a methodology that

- can be applied to scenarios where the hypothesis of infallible experts does not hold;
- can be used to characterize and study the different behaviors of different communities (e.g., experts, volunteers and paid workers);
- can be actioned to perform prospective analysis, allowing the manager of a citizen science experiment to make informed decisions such as, deciding which crowd to work with.

The dataset of this case study is published at the Zenodo Open-Science platform [RaviShankar2021].

3.5. TriggerCit

This case study, developed after the United Nations Satellite Centre (UNOSAT) requirements in 2022, aimed to automatically trigger data collection on social media when a natural disaster started, in order to understand when and where major events are unfolding. This machinery was intended to trigger satellite data collection and processing even before the activation requests to UNOSAT, and was developed as an autonomous system for detecting and characterizing large scale events. The studied cases were multiple flooding events in Thailand and Nepal. First, an adaptive system is configured, mostly automatically, in order to query the appropriate social media data and monitor it. The challenge of multilingualism is addressed by building this system language-wise. A neural network regressor is then trained, based on historical time series, in order to estimate when a particular kind of event is ongoing. If this condition holds true, social media data and media attachments are crawled and filtered using a VisualCit pipeline. Data are selected based on their image and text contents, and then enriched with geographical locations and projected on a real-time map, in order to get a representation of the spatial impact of the event. The case study results are published in [Bono2022b, Bono2022c] and also partially reported in [Deliverable 2.1, pp. 22-23].

3.6. Automatic geolocation of photos

Social media data has proved to be extremely relevant to assess damage and improve the understanding after a natural disaster occurs [e.g., Havas2017, Barozzi2019], especially, in the first 48 hours which are crucial to allow emergency responders to coordinate their actions. Precise geolocation of outdoor photos crawled from social media (e.g., by VisualCit) can have a critical impact in Disaster Management as it would facilitate faster first response. However, automatic image geolocation is a highly challenging task since photos taken even from the same location exhibit immense variations due to different camera settings, seasons, daylight conditions, and present objects. Also, images are often ambiguous and could provide very few cues about their location. In the absence of discriminative landmarks, humans can leverage their world knowledge to infer the location of a photo, using hints like the language of street signs or the driving direction of cars. Most previous work on image geolocation focused on identifying and geolocating landmark buildings (e.g., [Avrithis2010, Quack2008]) whereas very few approaches tried to geolocate images just by using pixels (e.g., [Li2012, Sattler2012]).

Due to the complexity of the problem, in [Murgese2022a], we restricted the area of geolocation to a single city, treating geolocation as a classification problem where the districts of a city are the classes to be distinguished. To this end, we created a geolocation software pipeline, where we first crawl two major online image sources for cities, namely Flickr and Mapillary. The latter source provided us with “robotic” image datasets, typically created by recording videos from static cameras mounted on a vehicle, resulting in images available in sequences and from a consistent point of view (usually the street). As the second step, to classify the photos into districts, we exploited the novel Focal Modulation Networks (FocalNets) [Yang2022] that are proven to perform effectively and efficiently in visual modeling for real-world applications.

Our experiments with Barcelona images yielded promising results and indicated us future work directions to explore. Our methodology and findings are published as a book chapter [Murgese2022a], and formed the basis of two MSc theses [Alcaina2022, Murgese2022b].

3.7. Crowdsourced geolocation

As mentioned in Section 3.6, photo geolocation is a hard task that has not been efficiently solved at the present time despite many efforts and plenty of different techniques and strategies devised to address the problem. A current state-of-the-art approach to tackle this issue is to ask a certain number of annotators to report the location of a given photo and then try to build up a consensus by computing, for instance, the reports' mean. In our recent work [Rocco2023], we aimed to build a paradigm to learn important characteristics about the annotators regarding their geolocation skills and use this information to get a much more efficient and precise consensus. To this end, we built mathematical models that can accommodate annotator behavior based on the location of the entity to be annotated. We also created computational models that can effectively learn the annotator behavior given an annotation dataset. These learned models of behavior can be leveraged to distribute geolocation tasks efficiently to a group of annotators. For example, depending on the location of a given photo, we may prefer an annotator who we learned to be more accurate in that region than other annotators to geolocate the image more precisely.

Following the research directions stated in the Geolocation Workshop that we organized in March 2022 at the Artificial Intelligence Research Institute (IIIA-CSIC) premises (reported in D2.2, p.12), we collaborated with the Joint Research Center (JRC-EU) in their research activity that aimed at reducing the overall time to identify locations impacted by a natural disaster.

Our collaboration demonstrated that it is possible to automatically predict the difficulty of geolocating an image, thus allowing us to filter out images impossible to geolocate and giving priority to those images which can be geolocated within a few minutes with high precision. The training dataset for JRC's Machine Learning classifier was curated by using CSPB to collect volunteer annotations for the difficulty of geolocating a dataset of disaster images and Crowdanalysis to compute the consensus on the annotations. Figure 3 shows the CSPB screen used for geolocating a photo. The outcome of this collaboration is due to be disseminated in a research article soon after the publication of this report.

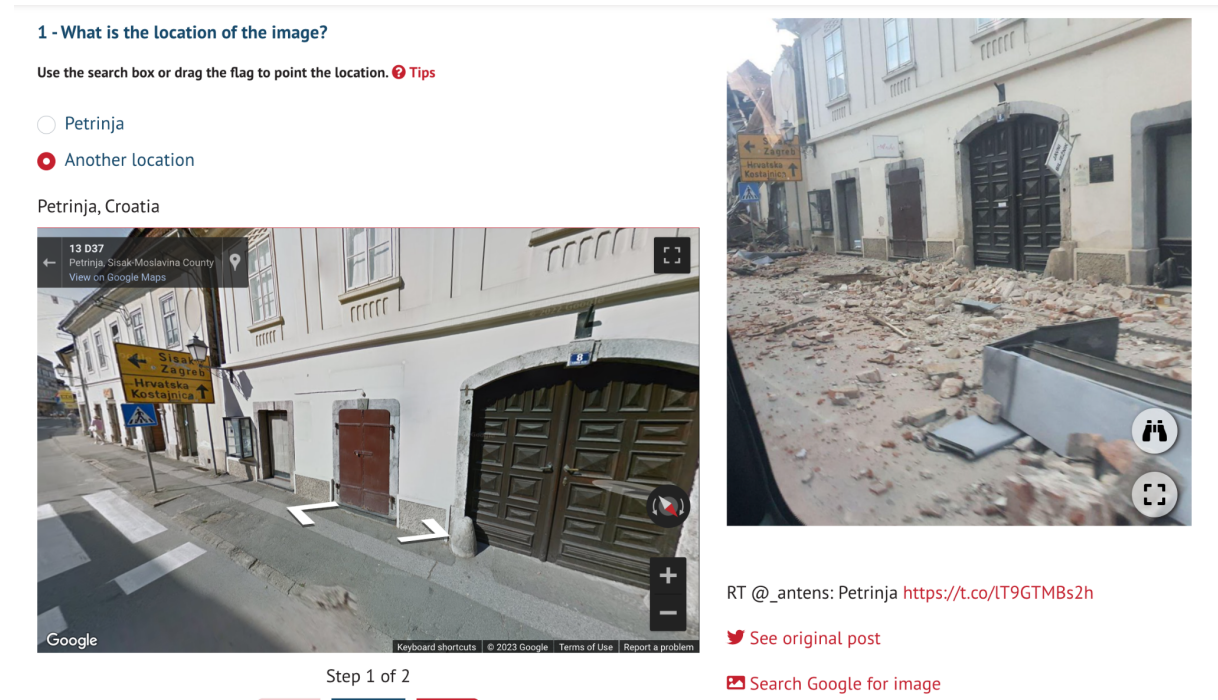


Figure 3: Screenshot from the CSPB for geolocating social media content using Google Street View (photo from Croatia earthquake).

4. Fulfillment of tasks

Table 3 summarizes the actions carried out by WP2 beneficiaries over the 3 years of the Crowd4SDG project to fulfill the tasks described in the Grant Agreement (pp. 12-13). The references column presents either the scientific publication resulting from the execution of the action, or the deliverable where the action is described.

Task	Actions	Year	References
T2.1: Deliberation technologies for citizen science	Initiated a proposal to integrate the “debate aggregation” algorithms developed within this task into Decidim. Decidim (decidim.org) is the base platform on which Decidim4CS –the public deliberation platform developed in Crowd4SDG– is built on. The proposal is under consideration by the Decidim community.	3	[MetaDecidim2022a]
	Participated in the 17 th session of the LAB MetaDecidim on Deliberative Democracy at Barcelona.		[MetaDecidim2022b]
	Completed this task by introducing models of human-machine debates and extending our work in Year 1.	2	[Deliverable 2.2, pp. 27-30]
	Developed a formal model for large-scale human debates, and algorithms to compute collective decisions.	1	[Deliverable 2.1, pp. 34-38]
T2.2: Human-machine collaborative learning	Addressed the problem of automatically estimating a photo’s geographical location with the objective of geolocating images taken from disaster areas.		[Murgese2022a]
	Developed machine learning models that learn annotator behavior for geolocation annotations. These models can be leveraged to assign geolocation tasks to annotators efficiently to get a more reliable and precise consensus.	3	[Rocco2023]
	Published Crowdanalysis as a software package to simplify its use in CS applications.	2	[Crowdanalysis2022]



	Developed Crowdanalysis-Service and carried out its beta-release integration into CS Project Builder for automatic computation of consensus on annotations.		[Deliverable 2.2, pp. 12-14]
	Set the foundations of a mathematical engineering tool to build hierarchical models of annotation.		[Deliverable 2.2, pp. 31-32]
	Built a conceptual and mathematical foundation for human-machine collaborative learning in Citizen Science. Developed the Crowdanalysis software library based on this framework for computation of intelligent consensus that can represent individual annotator behavior.	1	[Deliverable 2.1, pp. 15-17]
T2.3: Agreement and data quality analysis	The work on pipeline configurability was continued to allow the designer, in a human-in-the-loop fashion, to identify and configure the pipeline components to achieve the best quality. We focused on defining and improving pipelines for the analysis of social media by enhancing the back-end functionalities for the human-in-the-loop approach. We refined the methodology and developed tools for improving data pipelines for social media analysis, to make them more effective and efficient.	3	[Bono2022a] [Geisler2022]
	We built a configurable pipeline to iteratively process a crowdsourced dataset removing the noisy data and improving the overall quality of the dataset. The model and main components were defined.	2	[Deliverable 2.2, pp. 32-33]
	We focused on i) automatically estimating the accuracy of citizen scientists, their responses, and how to best aggregate them and ii) we have provided an evaluation of the quality of data collected from social media sources, evaluating its accuracy and precision.	1	[Deliverable 2.1, pp. 40-44]
T2.4: Self-composition . Adaptive services	We focused on defining and improving social media data analysis pipelines, enhancing the back end functionalities for human-in-the-loop approach to social data analysis and on extracting large scale spatio-temporal	3	[Bono2023]

	<p>descriptions from social media, focusing on the estimation of the data quality evaluation and uncertainty of the results.</p> <p>A methodology and tool environment has been developed to reconfigure the pipeline adapting it to the case being considered.</p>		
	The components and metrics for self-adaptation of an adaptive pipeline have been studied and tested.	2	[Deliverable 2.2, pp. 35]
	We started to develop a methodology to exploit the results of the above-mentioned data quality analysis as feedback to further optimize the composition of the data analysis pipeline. The main phases have been identified.	1	[Deliverable 2.1, pp. 45-47]
T2.5: Enriching Social Media content by Citizen scientist	<p>We aimed to support data discovery and enrichment by CS with techniques leveraging both image and textual contents of posts.</p> <ul style="list-style-type: none"> - We studied how a query dictionary for obtaining event-related data from social media can be automatically derived starting from a minimal dictionary, with minimal supervision from CS - The VisualCit interface was enhanced with a tool that is able to derive relevant keywords and hashtags from the crawled data, on order to refine the search in an interactive way http://visualcit.polimi.it:20003/ 	3	[Bono2022c]
	<p>We endowed VisualCit with a publicly accessible new interactive interface, additional services to use alternative social media (e.g., Flickr, disaster streaming data) and multilingual support. http://visualcit.polimi.it:7778/</p>	2	[Deliverable 2.2, pp. 36-37]
	<p>We completed some AI components for geolocation, published in [Scalia22]. Web services have been created for image classifiers for emergency contexts.</p>	1	[Deliverable 2.1, pp. 48-50]

Table 3. Task accomplishment summary. Bold years are the years the tasks are completed.

5. Interaction of WP2 with other Crowd4SDG work packages

CSSK is created by WP2 in order to provide citizen scientists with a set of advanced tools that facilitates the management and execution of their projects. CSSK has been developed throughout the Crowd4SDG Research Triangle illustrated in Figure 4. Crowd4SDG's research methodology is an iterative process of refinement thanks to the continuous interactions among different work packages. In the following subsections, we highlight the interactions that shaped the final release of CSSK.

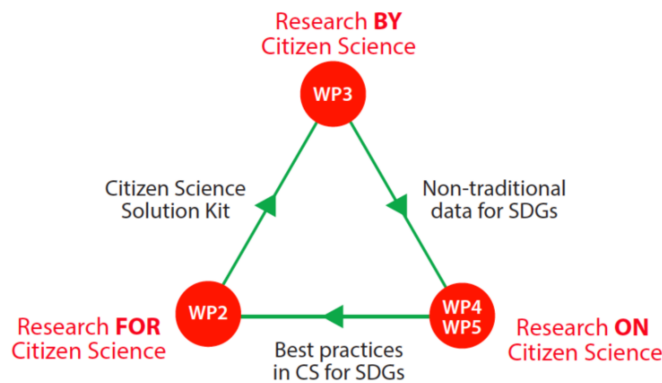


Figure 4. Crowd4SDG Research Triangle

5.1. Connection with WP3

Throughout the three GEAR cycles the WP2 members were supporting the teams joining the GEAR cycles as part of WP3. For the third and final GEAR cycle, the WP2 members joined the members of other WPs to refine the plans for interaction with GEAR-3 participants. A more direct interaction was planned, starting from the Gather phase. (The Gather phase of GEAR is a call for CS projects through the O17 Challenge, on which a specific SDG theme is launched and widely publicized, notably through related EU support actions such as EU-Citizen Science.)

The CSSK tools detailed in Section 2 were presented to the participating teams in the GEAR-3 cycle kickoff meeting overseen by WP3. Subsequently, WP2 provided on-demand and proactive coaching throughout the Evaluate and Accelerate phases via Crowd4SDG's Slack channels (<https://crowd4sdg.slack.com>). Most notably a tool owner representing each of the CSSK tools was present during the Accelerate phase each time the teams presented their progress so far. This enabled timely suggestions on which tools the teams could benefit from, and feedback on the way the teams were planning to use the tools. Additionally, we organized several hands-on online sessions to explore and facilitate the adoption of the tools by the teams, specifically for CS Project Builder, CS Logger and VisualCit.

5.2. Connection with WP4

WP4 calculates citizen collaboration analytics, such as team diversity and structure, by exploiting digital traces of usage recorded in the CSSK tools, e.g. a citizen scientist's contributions to a public debate on Decidim4CS. Since year two, the CSSK usage data has been made available to WP4 via the tools' APIs for third-party software. Beside these built-in records, CS Project Builder, VisualCit, and Decidim4CS also bear Google Analytics integrations that provide WP4 with a better understanding of the CSSK's user profiles and

user engagement, such as country of a connection, average time spent on a tool, referring site, etc.

To extract further information on citizen scientists' interactions throughout the GEAR cycles, WP4 conducted surveys with the participants and tracked the communication on Crowd4SDG's Slack channels between the participants and mentors and coaches from Crowd4SDG partners, in general, and from WP2 members, in particular.

Altogether, the digital traces and survey answers help WP4 compute the metrics that assess the quality of citizen science data.

5.3. Connection with WP5

WP5 analyzes the datasets used in the case studies by CSSK tools with respect to their potential value for monitoring extreme climate events and SDGs in particular. In the third year, following our research detailed above, we studied the problems in collecting citizens-generated data and we participated in events organized by UNITAR. In particular, we presented our work in the event of the Expert Group Meeting On Harnessing Data By Citizens For Public Policy And SDG Monitoring: A Conceptual Framework (10-11 November 2022 in Thailand) and we are planning to co-organize a Crowd4SDG event in the UN World Data Forum 2023 in China, in a learning event entitled "The Five 'Ws' of citizen science or citizen generated data", co-organized by UNITAR; United Nations Statistics Division (UNSD/DESA); Spanish Council for Scientific Research (CSIC); Politecnico di Milano; Instituto Nacional de Estadística y Geografía (INEGI), Mexico.

In year three, WP5 also assessed the quality of the LunarApp dataset mentioned in Section 3.3, with respect to a set of predefined criteria including Accessibility, Timeliness, Coverage and Privacy. The LunarApp dataset has received a quality score of 1.875 within the range of [0..2].

In year two, as WP2 partners we had participated in several meetings organized by UNITAR with participants from NSO Maldives and local Civil Society Organizations. The idea of these meetings was to see if CSSK tools could help them in their projects which target SDG indicators 14.1.1b (Marine plastic debris) and 14.5.1 (Coverage of protected areas). To our knowledge, UNITAR is still working with the Maldives team and it might be the case that the CSSK could serve them even after the end of Crowd4SDG project as the tools will be available for use.

5.4. Connection with WP6

WP2 partners joined the outreach activities coordinated by WP6 to promote the GEAR-3 cycle before its start.

Furthermore, to contribute to the dissemination and outreach activities organized by WP6, we shared with the WP6 members our Crowd4SDG-related and peer-reviewed publications, and information on events where we presented our research and the Crowd4SDG project during the third year.

Finally, all WP2 partners participated in the Geneva Triologue that was held on March 16, 2023 and the Crowd4SDG Final Conference on the following day. Members of the WP2 represented their work to the audience at both events. Both events took place at the Globe of Science and Innovation at CERN and IdeaSquare.

6. Conclusions

In this deliverable we presented the final release of Citizen Science Solution Kit (CSSK) developed and maintained by the Work Package 2 (WP2) partners in the Crowd4SDG project. We concisely summarized the development of CSSK throughout the previous two years and detailed the improvements in the last year of the project. We described the case studies with real-world scenarios and the collected citizen science data that demonstrate the competence and potential of CSSK in citizen science.

One of the key motivations in the development of CSSK was to aid the adoption of CSSK tools by the participants of the GEAR cycle in Crowd4SDG. The two finalists of GEAR-3 successfully incorporated our tools in their prototypes. Team Acquatech created a mobile app via *CSLogger* to gather the status and location of the hand pumps in Sub-Saharan Africa that need repairs to address the lack of access to clean water. Team Let's Clean Up tackles the waste management problem in Rome. They used *Decidim4CS* as the deliberation platform for citizens to pinpoint the problems they are facing. The team gathered waste-related Twitter messages from the area via *VisualCit*. *Citizen Science Project Builder* was used to further annotate and filter the collected images.

In this report, we also described in detail the fulfillment of each task assigned to WP2 in the Grant Agreement. The extensive work by WP2 has been disseminated in more than 20 peer-reviewed scientific publications, all reported in D2.1, D2.2, and this document. As a fitting postlude, a paper dedicated to CSSK is published in a top-level journal [Bono2023] involving all WP2 partners.

In the third year, we refurbished the tools page of the Crowd4SDG web site with tech-sheets, introductory videos, and tutorials on the tools (<https://crowd4sdg.eu/about-2/tools/>). WP2 partners will maintain the accessibility of CSSK tools after the Crowd4SDG project. Particularly, the Politecnico di Milano (POLIMI) will continue with the university research on social media data preparation pipelines, information extraction, and geolocation. The research will explore other social media platforms, such as Reddit, and the integration of information from different sources to analyze its precision both in space and time. In a new project on cybersecurity awareness (<https://www.cs-aware-next.eu/>), the research on data preparation pipelines which started in Crowd4SDG will be continued as well. Spanish National Research Council (CSIC) and the University of Barcelona are currently leading a proposal to integrate the "debate aggregation" algorithms developed during the Crowd4SDG project into Decidim, the democratic citizen participation platform that *Decidim4CS* is based on. CSIC continues to work on advanced models of aggregation for geolocation tasks. The outcome of this research will be integrated into *CrowdNalysis*. Citizen Science Centre Zurich carries on with *Citizen Science Project Builder* and *CSLogger* in their CS projects and regularly enhances these two tools.

References

- [Adrot2022]** A. Adrot, S. Auclair, J. Coche, A. Fertier, C. Gracianne, and A. Montarnal, "Using social media data in emergency management: a proposal for a socio-technical framework and a systematic literature review," in Proc. ISCRAM, 2022, May 2022.
- [Alam2018]** F. Alam, F. Ofli, and M. Imran, "Processing social media images by combining human and machine computing during crises," *Int. J. Hum. Comput. Interaction*, vol. 34, no. 4, pp. 311–327, 2018.
- [Alcaina2022]** Gerard Alcaina, Image geolocation with Focal Modulation Networks, MSc Thesis in Modelling for Science and Engineering, Facultat de Ciències, Universitat Autònoma de Barcelona, 2022.
- [Anjum2021]** S. Anjum, A. Verma, B. Dang, and D. Gurari, "Exploring the use of deep learning with crowdsourcing to annotate images," *Human Computation*, vol. 8, no. 2, pp. 76–106, 2021.
- [Asif2021]** A. Asif, S. Khatoon, M. M. Hasan, M. A. Alshamari, S. Abdou, K. M. Elsayed, and M. Rashwan, "Automatic analysis of social media images to identify disaster type and infer appropriate emergency response," *Journal of Big Data*, vol. 8, no. 1, pp. 1–28, 2021.
- [Avrithis2010]** Yannis Avrithis, Yannis Kalantidis, Giorgos Tolia, and Evaggelos Spyrou. Retrieving landmark and non-landmark images from community photo collections. In *Proceedings of the 18th ACM international conference on Multimedia*, pp. 153–162, 2010.
- [Barozzi2019]** Sara Barozzi, Jose Luis Fernandez-Marquez, Amudha Ravi Shankar, Barbara Pernici, et al. Filtering images extracted from social media in the response phase of emergency events. In *16th Conference on Information Systems for Crisis Response and Management*, pp. 1–12, 2019.
- [Bono2022a]** Carlo A. Bono, Cinzia Cappiello, Barbara Pernici, Edoardo Ramalli, Monica Vitali, A Conceptual Model for Data Analysis Pipelines: Supporting the Designer in Datasets Construction, 2022, accepted for publication on *ACM Journal on Data and Information Quality (JDIQ)*, pp. 24
- [Bono2022b]** Carlo Bono, Barbara Pernici, Jose Luis Fernandez-Marquez, Amudha Ravi Shankar, Mehmet Oğuz Mülâyim, Edoardo Nemni, TriggerCit: Early Flood Alerting using Twitter and Geolocation - a comparison with alternative sources, *ISCRAM 2022 Conference Proceedings – 19th International Conference on Information Systems for Crisis Response and Management*, R. Grace and H. Baharmand, Eds. Tarbes, France, pp. 674–686. 2022. <http://idl.iscram.org/show.php?record=2447>
- [Bono2022c]** C. A. Bono, M. O. Mülâyim, and B. Pernici, "Learning early detection of emergencies from word usage patterns on social media," in *Proceedings of the 2022 ITDRR Information Technology in Disaster Risk Reduction*, Kristiansand, Norway, 2022.
- [Bono2023]** Carlo Bono, Mehmet Oğuz Mülâyim, Cinzia Cappiello, Mark Carman, Jesus Cerquides, Jose Luis Fernandez-Marquez, Rosy Mondardini, Edoardo Ramalli, Barbara Pernici, A Citizen Science Approach for Analyzing Social Media With Crowdsourcing, in *IEEE Access*, vol. 11, pp. 15329-15347, 2023, <https://doi.org/10.1109/ACCESS.2023.3243791>
- [Cerquides2021a]** Cerquides, J.; Mülâyim, M.O.; Hernández-González, J.; Ravi Shankar, A.; Fernandez-Marquez, J.L. A Conceptual Probabilistic Framework for Annotation Aggregation of Citizen Science Data. *Mathematics* 2021, 9, 875. <https://doi.org/10.3390/math9080875>



- [CoSo2021]** CoSo Platform Temporary Website <https://coso-wl-frontend.herokuapp.com/>, Last accessed 2023-04-06.
- [Craglia2015]** M. Craglia and L. Shanley, "Data democracy–increased supply of geospatial information and expanded participatory processes in the production of data," *International Journal of Digital Earth*, vol. 8, no. 9, pp. 679–693, 2015.
- [Crowdanalysis2022]** J. Cerquides, M.O. Mülâyim, crowdanalysis: A software library to help analyze crowdsourcing results, (2022). Zenodo. <https://doi.org/10.5281/zenodo.5898579>.
- [Dawid1979]** Dawid AP, Skene AM. Maximum Likelihood Estimation of Observer Error-Rates Using the EM Algorithm. *Appl Stat.* 1979; 28(1):20. <https://doi.org/10.2307/2346806>
- [Deliverable 2.1]** Crowd4SDG Deliverable 2.1 – CS tools design and early prototype available (2021)
- [Deliverable 2.2]** Crowd4SDG Deliverable 2.2 – CS tools beta release (2022)
- [Deliverable 4.6]** Crowd4SDG Deliverable 4.6 – Interface for visualization of team analytics from the platform and from the in situ collected data (2023)
- [Deliverable 5.1]** E. Proden, "Crowd4SDG Deliverable 5.1 - Initial report on relevance and quality-related considerations of citizen-science generated data," September, 2022, <https://crowd4sdg.eu/wp-content/uploads/2021/10/D5.1-Initial-report-on-relevance-and-quality.pdf>.
- [Deliverable 5.2]** E. Proden, "Crowd4SDG Deliverable 5.2 - Data usability assessment and recommendations for SDGs GEAR cycle 1", 2021, <https://crowd4sdg.eu/wpcontent/uploads/2021/10/D5.2-Data-usability-assessment.pdf>.
- [Dianmu2021]** Dianmu Tropical Storm <https://floodlist.com/asia/thailand-tropical-storm-dianmu-floods-september-2021>
- [Edelman2022]** Edelman A, Boniface ER, Benhar E, Han L, Matteson KA, Favaro C, Pearson JT, Darney BG. Association Between Menstrual Cycle Length and Coronavirus Disease 2019 (COVID-19) Vaccination: A U.S. Cohort. *Obstet Gynecol.* 2022 Apr 1;139(4):481-489. doi: 10.1097/AOG.0000000000004695. Epub 2022 Jan 5. PMID: 34991109; PMCID: PMC8936155.
- [Franzoni2022]** C. Franzoni, M. Poetz, and H. Sauermann, "Crowds, citizens, and science: a multi-dimensional framework and agenda for future research," *Industry and Innovation*, vol. 29, no. 2, pp. 251–284, 2022.
- [Fritz2019]** S. Fritz, L. See, T. Carlson, M. Haklay, J. L. Oliver, D. Fraisl, R. Mondardini, M. Brocklehurst, L. A. Shanley, S. Schade, and U. Wehn, "Citizen science and the united nations sustainable development goals," *Nature Sustainability*, vol. 2, no. 10, pp. 922–930, 2019.
- [Geisler2022]** S Geisler, ME Vidal, C Cappiello, BF Lóscio, A Gal, M Jarke, M Lenzerini, B. Pernici et al., Knowledge-Driven Data Ecosystems Toward Data Transparency, *ACM Journal of Data and Information Quality (JDIQ)* 14 (1), 1-12, March 2022
- [Grasso2016]** V. Grasso and A. Crisci, "Codified hashtags for weather warning on twitter: an italian case study," *PLoS currents*, vol. 8, 2016.
- [Havas2017]** Clemens Havas, Bernd Resch, Chiara Francalanci, Barbara Pernici, Gabriele Scalia, Jose Luis Fernandez-Marquez, Tim Van Achte, Gunter Zeug, Maria Rosa (Rosy) Mondardini, Domenico Grandoni, Birgit Kirsch, Milan Kalas, Valerio Lorini, and Stefan Ruping. E2mC: Improving emergency management service practice through social media and crowdsourcing analysis in near real time. *Sensors*, 17(12):2766, 2017.



- [Imran2015]** M. Imran, C. Castillo, F. Diaz, and S. Vieweg, "Processing social media messages in mass emergency: A survey," *ACM Comput. Surv.*, vol. 47, no. 4, pp. 67:1–67:38, 2015.
- [Imran2020]** M. Imran, F. Ofli, D. Caragea, and A. Torralba, "Using AI and Social Media Multimodal Content for Disaster Response and Management: Opportunities, Challenges, and Future Directions," *Information Processing & Management*, vol. 57, no. 5, p. 102261, 2020.
- [Karray2022]** H. Karray, A. D. Nicola, N. Matta, and H. Purohit, Eds., *ISCRAM 2022 Conference Proceedings - 19th International Conference on Information Systems for Crisis Response and Management*. ISCRAM Digital Library, 2022.
- [Li2012]** Yunpeng Li, Noah Snavely, Dan Huttenlocher, and Pascal Fua. Worldwide pose estimation using 3d point clouds. In *European conference on computer vision*, pp. 15–29. Springer, 2012.
- [Lorini2019]** V. Lorini, C. Castillo, F. Dottori, M. Kalas, D. Nappo, and P. Salamon, "Integrating social media into a pan-european flood awareness system: a multilingual approach," *arXiv preprint arXiv:1904.10876*, 2019.
- [Male2022]** Male, Victoria. "Menstruation and covid-19 vaccination." *BMJ*, 376 (2022).
- [MetaDecidim2022a]** MetaDecidim - Designing Decidim Together: Proposal to aggregate the opinions issued in a debate, <https://meta.decidim.org/processes/roadmap/f/122/proposals/16956> [Last accessed 20230113]
- [MetaDecidim2022b]** 17th session of the LAB Metadecidim - decidim.delibera: Decidim Deliberative Democracy. Canòdrom, Barcelona, (June 3, 2022). <https://meta.decidim.org/assemblies/eix-lab/f/87/meetings/1671>
- [Middleton2014]** S. E. Middleton, L. Middleton, and S. Modafferi, "Real-time crisis mapping of natural disasters using social media," *IEEE Intelligent Systems*, vol. 29, no. 2, pp. 9–17, 2014
- [Murgese2022a]** Fabio Murgese, Gerard Alcaina, Oguz Mulayim, Jesus Cerquides, & Jose Luis Fernandez Marquez (2022). Automatic Outdoor Image Geolocation with Focal Modulation Networks. Atia Cortés, Francisco Grimaldo, & Tommaso Flaminio (Eds.), *Frontiers in Artificial Intelligence and Applications* (pp 279–288). IOS Press. <https://doi.org/10.3233/FAIA220349>
- [Murgese2022b]** Fabio Murgese, Automatic Outdoor Image Geolocation with Focal Modulation Networks. MSc Thesis in Artificial Intelligence, Department of Computer Science, Università Di Pisa, 2022.
- [Negri2021]** V. Negri, D. Scuratti, S. Agresti, D. Rooein, G. Scalia, A. R. Shankar, J. L. Fernandez-Marquez, M. J. Carman, and B. Pernici, "Image-based social sensing: Combining AI and the crowd to mine policy-adherence indicators from twitter," in *43rd IEEE/ACM International Conference on Software Engineering: Software Engineering in Society, ICSE (SEIS) 2021, Madrid, Spain, May 25-28, 2021*. IEEE, 2021, pp. 92–101.
- [Okolloh2009]** O. Okolloh, "Ushahidi, or 'testimony': Web 2.0 tools for crowdsourcing crisis information," *Participatory learning and action*, vol. 59, no. 1, pp. 65–70, 2009.
- [Pennington2022]** C. V. Pennington, R. Bossu, F. Ofli, M. Imran, U. W. Qazi, J. Roch, and V. J. Banks, "A near-real-time global landslide incident reporting tool demonstrator using social media and artificial intelligence," *International Journal of Disaster Risk Reduction*, p. 103089, 2022.
- [Petersen2016]** Petersen, I., Douglas, I., & Whitaker, H. (2016). Self controlled case series methods: an alternative to standard epidemiological study designs. *BMJ*, 354.

- [Quack2008]** Till Quack, Bastian Leibe, and Luc Van Gool. World-scale mining of objects and events from community photo collections. In Proceedings of the 2008 international conference on Content-based image and video retrieval, pp. 47–56, 2008.
- [RaviShankar2019]** A. Ravi Shankar, J. L. Fernandez-Marquez, B. Pernici, G. Scalia, M. R. Mondardini, and G. Di Marzo Serugendo, “Crowd4Ems: A crowdsourcing platform for gathering and geolocating social media content in disaster response,” *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 331–340, 2019.
- [RaviShankar2021]** Ravi Shankar, Amudha, Fernandez-Marquez, Jose Luis, Mülâyim, Mehmet Oğuz, Cerquides Bueno, Jesus, Firmansyah, Hafiz Budi, & Pernici, Barbara. (2021). Crowd4SDG - Crowdsourced image classification and damage assessment (1.0) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.5535744>
- [Rocco2023]** Rocco Ballester, Yanis Labeyrie, Mehmet Oğuz Mülâyim, Jesus Cerquides, Mathematical and Computational Models for Crowdsourced Geolocalization (paper under revision for the Conference of the Catalan Association for Artificial Intelligence (CCIA 2023)).
- [Sakaki2013]** T. Sakaki, M. Okazaki, and Y. Matsuo, “Tweet analysis for real-time event detection and earthquake reporting system development,” *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 4, pp. 919–931, 2013.
- [Sattler2012]** Torsten Sattler, Bastian Leibe, and Leif Kobbelt. Improving image-based localization by active correspondence search. In *European conference on computer vision*, pp. 752–765. Springer, 2012.
- [Scalia2022]** G. Scalia, C. Francalanci, B. Pernici, CIME: Context-aware geolocation of emergency-related posts, *Geoinformatica*, 26 (1), 125-157
- [Shanley2019]** L. Shanley, A. Parker, S. Schade, and A. Bonn, “Policy perspectives on citizen science and crowdsourcing,” *Citizen Science: Theory and Practice*, 4(1), p.30., 2019.
- [Tackx2021]** Raphael Tackx, Leo Blondel, and Marc Santolini. 2021. Quantified Us: a group-in-the-loop approach to team network reconstruction. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers (UbiComp-ISWC '21 Adjunct)*, September 21–26, 2021, Virtual, USA. ACM, New York, NY, USA 6 Pages. <https://doi.org/10.1145/3460418.3479363>
- [Velasco-Regulez2022]** Velasco-Regulez, B., Fernandez-Marquez, J. L., Luqui, N., Cerquides, J., Arcos, J. L., Fukelman, A., & Perelló, J. (2022). Is the phase of the menstrual cycle relevant when getting the covid-19 vaccine?. *American Journal of Obstetrics & Gynecology*, 227(6), 913-915.
- [Wiegmann2021]** M.Wiegmann, J. Kersten, H. Senaratne, M. Potthast, F. Klan, and B. Stein, “Opportunities and risks of disaster data from social media: a systematic review of incident information,” *Natural Hazards and Earth System Sciences*, vol. 21, no. 5, pp. 1431–1444, 2021.
- [Yang2022]** Jianwei Yang, Chunyuan Li, and Jianfeng Gao. Focal modulation networks. <https://arxiv.org/abs/2203.11926>, 2022.



Annex : List of abbreviations

Abbreviation	Description
AI	Artificial Intelligence
API	Application Programming Interface
CBI	Challenge-based Innovation (in-person coaching)
CIME	Context-based IMage Extraction
CoSo	COllaborative SOnar
CS	Citizen Science
CSIC	Spanish National Research Council
CSCZ	Citizen Science Centre Zurich
CSSK	Citizen Science Solution Kit
D2.1	Deliverable D2.1 - CS tools design and early prototype available
D2.2	Deliverable D2.2 - CS tools beta release
GEAR	Gather, Evaluate, Accelerate, Refine
IIIA	Artificial Intelligence Research Institute of CSIC
ML	Machine Learning
NSO	National Statistical Office
O17	Open Seventeen Challenge (online coaching)
SDG	Sustainable Development Goal
UNOSAT	United Nations Satellite Centre
UNSD/DESA	United Nations Statistics Division (Department for Economic and Social Affairs)
WP	Work Package