Crowd4SDG
Citizen Science for the Sustainable Development Goals

Deliverable 1.2

Data Management Plan completed and approved

Deliverable identifier: D1.2

Due date: 31/10/2020

Justification for delay: due to exceptional activity for O17 Water Challenge as agreed with Project Officer

Document release date: 17/12/2020

Nature: Report

Dissemination Level: Public

Work Package: WP1

Lead Beneficiary: UNIGE

Contributing Beneficiaries: CERN, CSIC, POLIMI, UNITAR, UP

Document status: Final

Abstract:

This deliverable describes the data management life cycle of the data being collected, processed, and/or generated by the Crowd4SDG project. It aims at producing findable, accessible, interoperable, and reusable (FAIR) research data. This deliverable is produced by following the Guidelines on FAIR Data Management in the Horizon 2020 Framework Programme.

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

![Crowd4SDG logo]

Document history

	Name	Partner	Date	Version
Authored by	Jose Luis Fernandez Marquez, Giulia Brocco	UNIGE	20/10/20	1.0
Edited by	Jose Luis Fernandez Marquez, Giulia Brocco Marc Santolini Camille Masselot Oguz Mulayim Romain Muller Jesus Cerquides	All	05/11/20	1.0
Reviewed by	Marc Santolini Romain Muller Giulia Brocco Barbara Pernici Camille Masselot Oguz Mulayim Jesus Cerquides Elena Proden	All	08/12/20	1.0
Approved by	François Grey	UNIGE	17/12/20	1.0

D1.2 - Data Management Plan completed and approved

Table of Contents

D1.2 - Data Management Plan completed and approved

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 de Paris	UP	FR

D1.2 - Data Management Plan completed and approved

Crowd4SDG in brief

The 17 Sustainable Development Goals (SDGs), launched by the UN in 2015, are underpinned by 169 concrete targets and 232 measurable indicators. Some of these indicators have 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 will focus 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 will initiate 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 will use 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 will rigorously assess the quality of the scientific knowledge and usefulness of practical innovations occurring when teams develop new CS projects focusing on climate action. This will occur 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 will be 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.

5

D1.2 - Data Management Plan completed and approved

Grant Agreement description of the deliverable

Parts in bold describe the deliverable content.

"T1.3: IPR, DMP, Quality Management and Self-Assessment (UNIGE)

While the project is founded on the principles of Open Access and the use of Open Source, innovations in the form of technologies and tools may in some cases need to be protected for subsequent exploitation. UNIGE will lead the task of IPR coordination to reflect the interests of all the partners. Plans for use of project results will be prepared and included in the final report. **Crowd4SDG will take part in the open data initiative and prepare a Data Management Plan (DMP) to record, store and make available data sets that can benefit all interested stakeholders.** Quality assurance is the joint responsibility of all partners. A Quality Assurance Plan (incorporated with the Project Handbook) will include procedures to be applied to internal and external services and deliverables. The Plan will include, instructions, procedures, checklists and processes for reviewing deliverables (appointment of reviewers, checks for consistency, clarity, technical content, adherence to documentation standards, Open Access publication procedures etc). The Secretariat will set up and maintain the Risk Log."

D1.2 - Data Management Plan completed and approved

Purpose and scope of the deliverable

This deliverable describes the data management life cycle of the data being collected, processed, and/or generated by the Crowd4SDG project. It aims at producing Findable, Accessible, Interoperable, and Reusable (FAIR) research data. This deliverable is produced by following the "Guidelines on FAIR Data Management in Horizon 2020".

The Data Management Plan (DMP) template is designed to be applicable to any H2020 project that produces, collects, or processes research data. This deliverable is a living document in which information is made available on a finer level of granularity through updates as the implementation of the project progresses and whenever significant changes occur.

D1.2 - Data Management Plan completed and approved

1. Data Management Plan (DMP)

Crowd4SDG participates in the open data initiative providing a DMP to record, store and make available data sets that can benefit all interested stakeholders.

1.1. Identified datasets

Dataset Name and Outline of Data	Custodian	WP
DS1_Crowd4SDG_Coordination Consortium members, global subscribers to mailing lists, participants, meeting minutes and administrative documents.	UNIGE	WP1
DS2_Crowd4SDG_SDGSolutionKit Data resources needed to run the new AI tools for CS as well as the social media pipeline for CS.	CSIC/ POLIMI	WP2
DS3_Crowd4SDG_GEARPhasesEvaluation Statistics and evaluations (qualitative and quantitative) of the different GEAR Phases.	UNIGE/ CERN	WP3
DS4_Crowd4SDG_New_metrics_and_descriptors Surveys and digital traces from online collaboration tools used by participants	UP	WP4
DS5_Crowd4SDG_Impact_Recommendations	UNITAR	WP5
DS6_Crowd4SDG_Promotional_Material Event materials and outputs (such as Project Newsletters, blogs, tweets, social media posts, presentations, photographs, videos) which have been analysed/referred to by papers/research.	CERN	WP6

1.2. FAIR Data

1.2.1. Making data Findable, including provisions for metadata

Q: Is the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

D1.2 - Data Management Plan completed and approved

All data produced and/or used will have an associated metadata document (stored as a .txt file) which describes key aspects of the data, as defined in Annex 2 - Metadata to be recorded against each Dataset.

Q: What naming conventions do you follow?

Project deliverables are assigned a unique identifier Crowd4SDG-Dx.x-YYYYMMDD.

All files made publicly available should reference Crowd4SDG in their name, with the recommendation that the convention Crowd4SDG-xxxxxxxx where xxxxxxx is a meaningful short description.

Photographs and audio/visual recordings should be named Crowd4SDG-XXX-YYYYMMDDnnnnn where XXX-YYYYMMDD is the event identifier and nnnnn is a brief description of the event/photograph content.

Q: Will search keywords be provided that optimize possibilities for re-use?

Yes, the associated metadata document will contain the keywords associated with the dataset.

Q: Do you provide clear version numbers?

Version numbers are provided for both as a field in the associated metadata document and for deliverables in the deliverable description.

Q: What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Initially we will collect the metadata proposed in the associated metadata document. Depending on the nature of the data collected or processed, and following the submission rules of the repository chosen to make the data public, additional metadata can be added.

1.2.2. Making data openly Accessible

Q: Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.

D1.2 - Data Management Plan completed and approved

Relevant data generated or processed by WP4 and WP2 will be openly accessible following FAIR guidelines. In addition the data will be properly referenced from scientific publication to allow the replication of the experiments, and increasing the impact of the data sets.

Q: *How will the data be made accessible (e.g. by deposition in a repository)?*

Relevant research data will be referenced in scientific publications including journals and peer reviewed conference proceedings. In addition, the research data will be made publicly available through the ZENODO open-access repository developed under the European OpenAIRE program.

Q: *What methods or software tools are needed to access the data?*

Data will be shared using standard formats. So far, we do not identify the need of any specific software for accessing the data.

Q: *Is documentation about the software needed to access the data included?*

We do not foresee the need of any software to access the data.

Q: *Is it possible to include the relevant software (e.g. in open source code)?*

All code generated by the Crowd4SDG project will be released by using open source licenses.

Q: *Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.*

ZENODO open-access repository developed under the European OpenAIRE program. The code being developed is publicly available as open-source at the online GitHub repository (https://github.com/crowd4SDG) created for the Crowd4SDG project.

Q: *Have you explored appropriate arrangements with the identified repository?*

Github repository is currently used by the Crowd4SDG partners. The open-source code is publicly accessible at GitHub/GitLab and is provided with APGL-3.0 public license.
There are no arrangements yet with ZENODO.

D1.2 - Data Management Plan completed and approved

Q: *If there are restrictions on use, how will access be provided?*

At this stage we do not foresee any restriction on the use of the data.

Q: *Is there a need for a data access committee?*

At this stage we do not foresee the need for a data access committee.

Q: *Are there well described conditions for access (i.e. a machine readable license)?*

As recommended by the ZENODO dataset repository, we will add a machine readable Creative Commons License for each of the dataset published.

Q: *How will the identity of the person accessing the data be ascertained?*

Data published will be openly accessible without the need to identify the person accessing the data.

1.2.3. Making data Interoperable

Q: *Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?*

Yes, all data shared in the Crowd4SDG project will use standard formats.

Q: *What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?*

CSV and JSON will be the most common standard used to share datasets. In addition, we will follow the recommendations of the dataset repository adding the appropriate metadata to make the data interoperable.

Q: *Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?*

Yes, given the interdisciplinary nature of Citizen Science data, all data types will be using standard vocabularies.

D1.2 - Data Management Plan completed and approved

Q: *In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?*

We do not anticipate the use of any project specific ontology.

1.2.4. Increase data Re-use (through clarifying licenses)

Q: *How will the data be licensed to permit the widest re-use possible?*

All public data will be used creative common licenses. The most appropriate CC license will be decided for each case.

Q: *When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.*

At this stage we do not foresee the publication of any patents resulting from the Crowd4SDG's activities.

Q: *Are the data produced and/or used in the project usable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.*

Dataset published in the Crowd4SDG project will be made available through recommended dataset repositories to ensure the long time access of the data.

Q: *How long is it intended that the data remains re-usable?*

10+ years.

Q: *Are data quality assurance processes described?*

Different quality assurance processes will be applied depending on the nature of the data set. For example, public deliverables will be subject to internal peer review before the submission to the EC, dataset links to research activities will be made public after peer review process carried out by international conferences or journals. A more detailed description will be provided in the D1.1. Project handbook and quality plan prepared.

1.3. Allocation of resources

Each of the partners in the consortium commits to cover the cost of making open access both publications and datasets resulting from their research activities.

1.4. Data security

Q: *What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?*

Personal data such as individual questionnaire responses will be stored by partners in their private networks following the data privacy and data protection policies established by their institutions.

Section 2 identifies in detail:
- Partner responsible for securing data and the method of securing it;
- Partner responsible for backing up data.

For each dataset, a responsible partner has been identified (Custodian) – that partner will be responsible for ensuring everyone involved in the data is aware of their individual responsibilities.

Q: *Is the data safely stored in certified repositories for long term preservation and curation?*

Zenodo will be used as a certified repository for ensuring the long term preservation and curation of the data.

1.5. Ethical aspects

All data will be anonymised before being made public. Event participants will sign appropriate consent forms as identified in D7.1/D7.2/D7.3. Participants under 16 years old are not allowed to participate in the Crowd4SDG activities.
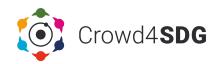
D1.2 - Data Management Plan completed and approved

2. Dataset details

2.1. DS1_Crowd4SDG_Coordination

Custodian: UNIGE
Work Package: WP1

Type of Data	Description	Access mechanism	Working Format	Primary Storage During Project	Access during project C=consortium P=responsible partner	Backup / security during project P=responsible partner	Preserved after project? Y=All data S=Sample N=No data	Preserved format	Preserved volume	Access after project U=unrestricted R=restricted
Partner information	Who is who	Google Drive	Spreadsheet	Google drive	C	P	Y	PDF	<1M	R
Internal communication	Meeting minutes and presentations	Google drive	Google Doc/ Slides/PPT	Google drive / UNIGE Intranet shared folder	C	P	Y	PDF/Word /PPT	<100M	R
Documents about the project	Proposal	Google drive	PDF/Word	Google drive / UNIGE Intranet shared folder	C	P	Y	PDF/Word	<20M	R
Documents about the project	Agreements (Grant Agreement, Consortium Agreement, UZH Linked Third Party Agreement, Covid-related Agreement)	Google drive	PDF/Word	Google drive / UNIGE Intranet shared folder	C	P	Y	PDF/Word	<50M	R
Project Deliverables	Final deliverables uploaded on the	Google drive/EC	Google Doc/PDF	Google drive / UNIGE	C	P	Y	PDF	<200M	R/U

D1.2 - Data Management Plan completed and approved

		EC Portal	Portal		Intranet shared folder						
GEAR phases information	Templates of communications to participants, presentations, handbooks and other documents relevant to each phase	Google drive	Google Docs, Google Sheets, Google Slides, PDF, Word, Excel, PPT	Google drive	C	P	Y	PDF/Word /Excel/PPT	1Gb	R	
GEAR phases information	Participants personal data, consent forms, certificates, etc.	UNIGE Intranet shared folder	PDF, Word, Excel	UNIGE Intranet shared folder	P	P	N	PDF, Word, Excel	>1Gb	R	
Timesheets	Timesheets of people working on the project	Internal drives of the partners	Word, Excel	Internal drives of the partners/UNIGE Intranet shared folder	P	P	Y	Word, Excel	<100M	R	

2.2. DS2_Crowd4SDG_SDGSolutionKit

Custodian: CSIC/POLIMI
Work Package: WP2

Type of Data	Description	Access mechanism	Working Format	Primary Storage During Project	Access during project C=consortium P=responsible partner	Backup / security during project P=responsible partner	Preserved after project? Y=all data S=Sample N=no data	Preserved format	Preserved volume	Access after project U=unrestricted P=Partially restricted
Code	Source code of the tools developed	Git + ssh Git + https	Git	Github GitLab	C	P	Y	Git	~ 10Gb	P
Scientific Papers	Working versions of the scientific papers that are jointly created by the project participants	web	LaTeX, MS Word	Overleaf, Google docs	P	P	N	-	-	-
Tweets	Tweets ids and links crawled for the pipeline tool	CSV	CSV	Local filesystem	P	P	Y	CSV	~10Gb	U
CS project data	Data regarding CS projects stored in the decidim4cs tool	Postgres DB	Postgres DB	Dedicated server	P	P	Y	Postgres DB	~1Gb	P
Crowdsourcing data	Data obtained in crowdsourcing activities	Postgres DB	Postgres DB	Pybossa server	P	P	S	CSV	~1Gb	P

D1.2 - Data Management Plan completed and approved

Technological Infrastructure document	Document describing the overall technological infrastructure of the project	Google docs	Google docs	Google docs	C	P	N	-	-	-

2.3. DS3_Crowd4SDG_GEARPhasesEvaluation

Custodian: UNIGE/CERN
Work Package: WP3

Type of Data	Description	Access mechanism	Working Format	Primary Storage During Project	Access during project C=consortium P=responsible partner	Backup / security during project P=responsible partner	Preserved after project? Y=all data S=Sample N=no data	Preserved format	Preserved volume	Access after project U=Unrestricted R=Restricted
Handbook	Cookbooks and other documents produced to aid event participants	Google Slide	Google Slide	Google drive	C	P	Y	Google Slides	<20Mb	U
Event consent forms	Photo and Video Recording & sharing consent forms	MSWord/PDF	MS Word/PDF	UNIGE Intranet shared folder	P	P	Y	PDF	55 Forms	R
Participant Information	Raw Data from Goodwall with initial list of applicants	Google Spreadsheet	Google Spreadsheet	Google Drive	C	P	Y	Google Spreadsheet		R
Participant Information &	Information including	Google Form,	Google Spreadsh	Google Drive	P	P	Y	Google Spreadsheet		R

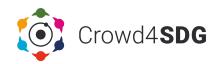| Surveys	personal data	Google Spread Sheet	eet							
Session Recordings	Event Recordings including participant presentations and expert presentations	Media Player	Media Player	Google Drive	C	P	Y	Link to the shared drive		R
Project Documenta tion	Documentation in SDGinProgress, Including text, images, slides and related links	Digital Platform	Digital Platfor m	Digital Platform	C	P	Y	Link to the collection of project url		U
Project Slides	Project work summarised over weekly presentation	Google Slides / PDF	Google Slides or other appropri ate softwar e	Google Drive	C	P	Y	Link to the Google shared drive		R
Project Summary	Project summary submitted by the participants	MSWord / PDF	Appropri ate Softwar e	Google Drive	C	P	Y	Link to the shared drive		R
Project Evaluation Forms	Mentor Evaluation forms collected over the 4 weeks	Google Google Form	Google Spreads heet	Google Drive	P	P	Y	Google Spreadsheet		R

2.4. DS4_Crowd4SDG_New_metrics_and_descriptors

Custodian: UP
Work Package: WP4

Type of Data	Description	Access mechanism	Working Format	Primary Storage During Project	Access during project C=consortium P=responsible partner	Backup / security during project P=responsible partner	Preserved after project? Y=all data S=Sample N=no data	Preserved format	Preserved volume	Access after project U=unrestricted P=Partially restricted R = restricted
Survey - answers	New metrics and descriptors identified by experts and citizen science projects participants	Google Form	Google spreadsheet	Google Drive	C	P	Y	Zenedoo	<20Mb	U
Transcripts	Transcripts of interviews or interview notes, anonymized	Google Drive	Google docs	Google Drive	C	P	Y	PDF/Word	<20Mb	P - on demand
Protocols	Description of the research methods and tools	Google Drive	Google docs	Google Drive	C	P	Y	PDF/Word	-	U
De-identified participant online activity	Online activity of Crowd4SDG participants	Digital Platform	csv	Google Drive	P	P	Y	Zenedo	~1Gb	U
Participants sensitive information	Include personal information: email, age,	Digital Platform	csv	CRI OVH cloud server (The data is stored in	P	P	Y	-	-	-

D1.2 - Data Management Plan completed and approved

	location, and other information as displayed online in participants profiles			France (Gravelines), secured with SSH protocol)						
Identifier matching table	Document allowing to retrieve participants from anonymized data	NA	txt	CRI intranet	P	P	N	-	-	-
Code	Source code of data analysis	Github	Github	Github.com	C	P	Y	Git	~ 1Gb	U
Reports	Results of the work conducted by the WP4	Google Drive	Google docs	Google Drive	C	P	Y	PDF pre-print	<20Mb	U
Offline activities (anonymized)	Offline collaborations of Crowd4SDG participant	Digital Platform	JSON objects	Heroku	P	P	Y	Zenodo	.~1Gb	U

2.5. DS5_Crowd4SDG_Impact_Recommendations

Custodian: UNITAR
Work Package: WP5

Type of Data	Description	Access mechanism	Working Format	Primary Storage During Project	Access during project C=consortium P=responsible partner	Backup / security during project P=responsible partner	Preserved after project? Y=all data S=Sample N=no data	Preserved format	Preserved volume	Access after project U=unrestricted
Study design document	Outlines the research methodology to identify quality requirements and produce recommendations for CS data for SDGs from the perspective of NSOs and decision-makers	MsWord/PDF	MsWord/PDF	PDF	C	P	Y	MsWord/PDF	1 document	R
Survey questionnaires and responses	Survey questionnaires and responses for the first research described above and survey questionnaires and responses as part of the impact evaluation at the end of the project	UNITAR Survey Monkey account	PDF/Excel	UNITAR Survey Monkey account/PDF	P	P	Y	PDF/Excel	6 PDF/6 excel documents	R
Interview protocols	Interviews conducted as part	Word / PDF/audio	Word/PDF	Word/PDF	P	P	Y	Word/PDF	2 compilations	R

D1.2 - Data Management Plan completed and approved

	of the initial research and as part of impact evaluation at the end	/video record if individual's agreement given								
Policy briefs	2 policy briefs: 1 on fitness for purpose od CS data for NSOs/policy makers and 1 on impact of the project	PDF	Word	PDF	C	P	Y	PDF	2 documents	U
Recommendations report	1 report with recommendations on quality of data	PDF	PDF	PDF	C	P	Y	PDF	1 document	R
3 analysis report	3 reports providing a summary of quality assessments of GEAR cycle/data related projects and recommendations	Word/PDF	PDF	PDF	C	P	Y	PDF	3 documents	R
Impact assessment report	1 report containing impact assessment from the viewpoint of data and SDG relevant innovations	Word/PDF	PDF	PDF	C	P	Y	PDF	1 document	R

D1.2 - Data Management Plan completed and approved

2.6. DS6_Crowd4SDG_Promotional_Material

Custodian: CERN
Work Package: WP6

Type of Data	Description	Access mechanism	Working Format	Primary Storage During Project	Access during project C=consortium P=responsible partner	Backup / security during project P=responsible partner	Preserved after project? Y=all data S=Sample N=no data	Preserved format	Preserved volume	Access after project U=unrestricted
Communication record	Record of communication activities	Google Drive	Google form	Google Drive	P	P	Y	PDF	1000 rows	U
Dissemination record	Record of dissemination events about Crowd4SDG	Google Drive	Google form	Google Drive	C	P	Y	PDF	100 rows	U
Event promotional material	Flyers, Event programmes. promotional videos, digital banners etc.	Google Drive	Pdf, jpg, gif, png, mv3/4 etc	Google Drive	C	P	Y	original format	10Gb	U
Branding source Data	Logo designs, printed material designs	Google Drive	Pdf, jpg, gif, png, etc.	Google Drive	C	P	Y	original format	1Gb	U

D1.2 - Data Management Plan completed and approved

Annex 1: List of abbreviations

Abbreviation	Description
CC	Creative Commons
CS	Citizen Science
DoA	Description of the Action
DMP	Data Management Plan
EAB	External Advisory Board
EC	European Commission
ECGA	European Commission Grant Agreement
GEAR	Gather, Evaluate, Accelerate, Refine
NSO	National Statistical Office
O17	Open Seventeen Challenge (online coaching)
PO	Project Officer
QA	Quality Assurance
SB	Supervisory Board
SDG	Sustainable Development Goal
TMB	Technical Management Board
WP	Work Package
WPL	Work Package Leader

D1.2 - Data Management Plan completed and approved

Annex 2: Metadata to be recorded against each dataset

Each WP folder in the Crowd4SDG Google drive should contain a file with the following information:

Field	Value
Project Name	Crowd4SDG
Principal Investigator	François Grey
Project Description	
Funding Sources	Horizon 2020
Responsible Organisation	<Partner Name>
Contact Person	<Full name>
Contact Email	<Email address>
Data Overview	Dataset description as given by this document
Technical Information of Files	For qualitative data, description and format of each element
Coding Instrument	
Data Collection Start Date	
Data Collection End Date	
Confidentiality Classification	Confidential/Public
Data Subject	Events/Methodology/Technology/People
Keywords	List of keywords
Version	

D1.2 - Data Management Plan completed and approved