

SMARTLAGOON

DELIVERABLE 6.9 DATA MANAGEMENT PLAN Report 1.0



This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101017861.



Innovative modelling approaches for predicting Socio-environmental evolution in highly anthropized coastal LAGOONS

Deliverable 6.9

Deliverable No.:	6.9
Project Acronym:	DMP
Full Title:	Data Management Plan
Grant Agreement No.:	101017861
Workpackage/Measure No.:	6
Workpackage/ Measure Title:	Data Management Plan
Responsible Beneficiary:	Universitat Politècnica de València (UPV)
Responsible Author:	José M. Cecilia (UPV)
Responsible Co-Author(s):	All
Date:	21/06/2021
Type:	Open Research Data Plan (ORDP)
Dissemination level:	Public

HISTORY OF CHANGES		
Date	Content	Author(s)

The content of this deliverable represents the views of the authors only and sole responsibility; it cannot be considered to reflect the views of the European Commission or any other body of the European Union. The European Commission and the Agency do not accept any responsibility for use that may be made of the information it contains SmartLagoon-Marchio-Orizzontale

PUBLISHABLE SUMMARY

The SMARTLAGOON Data Management Plan (DMP) describes the procedures used in the project for the handling of data during and after the end of the project. The plan discusses what kind of data will be collected, processed, and synthesized, which methodology and standards will be applied during data collection and handling and elaborates procedures for sharing and open access to the SMARTLAGOON data and for curation and preservation of the data. Furthermore, procedures in relation to the General Data Protection Regulation (GDPR) are defined and how SMARTLAGOON ensures the protection of the involved companies' data, information, and privacy rights.

As part of Horizon 2020, the SMARTLAGOON project participates in a pilot action on open research data. The aim is to provide indications as to what kind of data the project will collect, how the data will be preserved and which sharing policies will be adopted towards making these data readily available to the research community. The project's efforts in open research data are outlined giving particular attention to the following issues:

- The types of open and non-open data that will be generated, collected or processed by the consortium, via in-situ monitoring, stakeholders interactions, social-media crawling and research, during the project's lifespan;
- The technologies and infrastructures that will be used to securely preserve the data long-term (including after the project has ended);
- The standards used to encode the data;
- The data exploitation plans;
- The sharing/access policies applied to each dataset.

The plan can be considered as a checklist for the future and as a reference for the resource and budget allocations related to data management. The content of this document builds upon the input of the project partners. Several interviews and discussions in different project meetings with all partners have been carried out outlining the DMP's objectives and stating the required information in a structured manner. This has been edited by UPV and will be disseminated to the partners. The compiled answers have been integrated into a coherent plan (i.e., this document), which will be updated throughout the project lifecycle. This DMP will be a living document, which will evolve as the project progresses in accord with the project's efforts in this area. At any time, the DMP will reflect the current state of the consortium's agreements regarding data management, exploitation and protection of rights and results. The considered storage facilities are outlined and tutorials are provided for their use.

Abbreviations

API: Application Programming Interface

DMP: Data Management Plan

DOI: Digital Object Identifier

FAIR: Findable, Accessible, Interoperable and Re-usable

GDPR: General Data Protection Regulation

OA: Open Access

ORDP: Open Research Data Pilot

Partners short names

UCAM: Fundación Universitaria San Antonio (Spain, Research Organization)

UPV: Universitat Politècnica de València (Spain, Research Organization)

WIT: WaterITech ApS (Denmark, High-tech SME)

UU: Uppsala University (Sweden, Research Organization)

NIVA: Norwegian Institute for Water Research (Norway, Research Organization)

UNIBO: Università di Bologna (Italy, Research Organization)

PHO: Photrack, AG (Switzerland, High-tech SME)

VIELCA: Vielca Ingenieros, S.A. (Spain, SME)

INDEX

<i>PUBLISHABLE SUMMARY</i>	2
<i>Abbreviations</i>	3
<i>Partners short names</i>	3
<i>INDEX</i>	4
1. <i>Introduction</i>	5
2. <i>SMARTLAGOON Data Summary</i>	5
2.1 Data set description, reference, and name	6
3. <i>Fair Data</i>	9
3.1 Open Data	9
3.2 Accessible Data	10
3.3 Interoperable Data	11
3.4 Reusable Data	11
4. <i>Allocation of resources</i>	12
5. <i>Data Security</i>	12
6. <i>Ethical Aspects</i>	13
7. <i>Other Issues</i>	14
<i>Annex I: Tutorial on Zenodo – Open digital repository</i>	16
<i>Annex II: Tutorial GitHub</i>	19

1. Introduction

The SMARTLAGOON Data Management Plan (DMP) gives an overview of the data and information collected throughout the project and shows the interaction and interrelation of the data collecting activities within and between the work packages. The DMP will also link these activities to the SMARTLAGOON partners and discuss their responsibilities with respect to all aspects of data handling.

Furthermore, the SMARTLAGOON DMP will lay out the procedure for data collection, consent procedure, storage, protection, retention and destruction of data, and confirmation that they comply with national and EU legislation. The DMP will ensure that the exchange of data of companies and industries is in full compliance with the participating companies and industries internal data protection strategies. This DMP aims at providing an effective framework to ensure comprehensive collecting and handling of the data used in the project. Thereby and wherever trade secrets of the participating companies and industries are not violated, SMARTLAGOON strives to comply with the open access policy of Horizon 2020.

The DMP is intended to be a living document which will be adjusted to the specific needs of SMARTLAGOON throughout the project's runtime and will be adapted whenever appropriate. The DMP will be revised during the project within Task 6.4 Data Management Plan, including new data, changes in consortium policies regarding innovation potential or decision to file a patent, and changes in the consortium composition and external factors.

This plan will establish the measures for promoting the findings during SMARTLAGOON's lifecycle and will set the procedures for the sharing of data of the project. Addressing FAIR principle for research data (Findable, Accessible, Interoperable and Re-usable) SMARTLAGOON DMP will consider:

- Data set reference and name
- Data set description
- Standards and metadata
- Data sharing and handling during and after the end of the project
- Archiving and preservation (including after the end of the project)

The following document made use of the HORIZON 2020 FAIR DATA MANAGEMENT PLAN TEMPLATE and was written with reference to the Guidelines to FAIR data management in Horizon 2020 [1] and the GDPR (Regulation (EU) 2016/679).

2. SMARTLAGOON Data Summary

Being in line with the EU's guidelines regarding the DMP, this document addresses for each dataset collected, processed and/or generated in the project the following characteristics: dataset description, reference and name, standards and metadata, data sharing, archiving and preservation. At this point in time, an estimation of the size of the data cannot be given. To this end, the consortium develops several strategies that will be followed to address the above elements.

This section shall be provided a detailed description of these elements to ensure their understanding by the partners of the consortium. For each element, we also describe the strategy that will be used to address it.

2.1 Data set description, reference, and name

Each dataset will be assigned a unique name that will serve as a reference for easy identification by all members of the consortium (see Section 3.3 for details). This name can also be used as the identifier or primary key of the dataset. All data files produced, including emails, must include the term “SMLG”, followed by file name which briefly describes its content, followed by a version number (or the term “FINAL”), followed by the short name of the organisation which prepared the document (if relevant). Each dataset that will be collected, processed, or generated within the project will be accompanied by a brief description. Moreover, a list of names is provided for the main documents generated by the project (see Table 1).

The main data expected to be generated in the SMARTLAGOON project are shown below. These data have been identified after consultation with the different SMARTLAGOON partners about the data they expected to generate in their tasks. This list may be expanded during the lifetime of the project as the research progresses. The main technical building blocks of SMARTLAGOON include the:

- **Efficient ICT techniques for data collection:** the project will develop efficient tools and algorithms for crawling socio-environmental information in real time. From the hardware perspective, efficient IoT infrastructures with novel AI-based sensing technologies will be developed and eventually deployed in Mar Menor lagoon and its watershed.
- **Process-based modelling and simulation techniques:** the project will develop and strengthen modelling capabilities and overcome limitations of existing modelling tools. At the project onset a modelling scope analysis will be conducted to allow for developing linkages between the different models and broadening the range of questions that could be explored and answered.
- **Machine learning (ML) driven models:** new approaches to creating environmental models will be developed based on the capacity of ML techniques to estimate nonlinear relationships between environmental variables. These techniques will be combined with physically based models to build more efficient hybrid predictive models.
- **Participatory system dynamics modelling for strategic policy and management analysis:** using the project's strengthened modelling capabilities in a multi-model set-up, SMARTLAGOON will analyse, through systems dynamic modelling and scenario studies, multiple dimensions of future pathways, environmental risks, and management practices.
- **Citizen engagement, stakeholder dialogue and interactions:** multiple participatory activities (workshops, expert seminars, citizen science activities) will be organized with stakeholders and citizens to capture perspectives and knowledge and actively involve them in an iterative process aimed at both understanding needs, as well as presenting results and co-developing results and products.

The kind of data that SMARTLAGOON will collect is mainly environmental parameters, but includes also socio-demographic and economic parameters. Furthermore, when using the mobile applications and integrating the data into a web portal, other data is also collected, including personal data.

The research objectives require qualitative and quantitative data that are not available from other sources, than the ones listed here below:

❖ **Personal Data**

- Name and surname
- Location/positioning, GPS
- Occupation and education (pottentially)
- Image / voice

❖ **Public Datasets:**

- ERA5 - a dataset available for public use, which provides hourly estimates of many atmospheric, land and oceanic climate variables.
- HWSD – a dataset available for public use, which provides a 30 arc-second raster database with over 15 000 different soil mapping units worldwide.
- GlobCover - a dataset available for public use from the European Space Agency initiative, which provides global land cover maps.
- SRTM – a dataset available for public use, which provides digital elevation data of the world, produced by NASA originally.
- Open data sources such as INE, AEMET, SIAM from IMIDA among others.
- A comprehensive inventory of sustainability criteria and socio-economic indicators for the case of Mar Menor lagoon/region: Inventory of ~30 policy documents (relevant strategies, regulations), including a brief characterisation and analysis regarding their potential impact on the sustainability of the Mar Menor lagoon/region.
- Inventory of parameters and relationships used in the Mar Menor System Dynamics model: An overview of the parameters and underlying hypotheses used to build and run the Mar Menor System Dynamics model, including a brief characterisation and sources (literature, experts judgement, etc).

❖ **Surveys (anonymized data)**

- Stakeholders needs/requirement/expectations of the modelling tools which will be developed in SMLG.

❖ **Workshops**

- Group modelling workshops - results of stakeholders' views on sustainable development of Mar Menor region: Structured overview of knowledge and views of selected stakeholders on the interlinkages between various activities and the well-being of Mar Menor ecosystem/region, intervention points for policy instruments for the selected sector(s), policy-related challenges concerning Mar Menor region sustainability.

❖ **Observations**

- Socio-economic data on the Mar Menor Region Secondary quantitative data and qualitative.
- Environmental data on the Mar Menor Region Secondary quantitative data, potentially some text e.g., pollution data.
- Environmental web maps services Secondary GIS maps E.g., Natural protection areas, Algal bloom information from Sentinel, etc.
- Data elicited from stakeholder interactions Primary Sound (recordings of interviews), text (transcriptions), interview summaries (text).
- Data elicited from local regulators/ policy makers Primary Sound, summaries (text), questionnaires (e.g., Google Forms) Workshop / focus group.
- Perceptions about the Mar Menor Region Primary Text, images, videos Through the social sensing activity.

- Meteorological data.
 - Water temperature data.
 - Water quality data.
 - Stream flow data.
 - Stream water quality data.
 - Video data type.
 - Stream flow data.
 - Lake and streams water surface data: for the video data we will be looking at a small sections of the stream and lake, focusing entirely on the water surface. With this in mind it is very unlikely that any personal images will be captured, but in the unlikely event of a person being visible on an image, the person will be pixelated.
- ❖ **Social media data** (mainly Twitter, YouTube, Telegram and RSS feeds).
- ❖ **Models and software**
- Social Sensing model for microtext processing.
 - Environmental model setups for project locations.
 - Model-output from environmental models for Erken and Mar Menor.
 - Datasets coming out from System Dynamics modelling.

Moreover, SMARTLAGOON will collect posts by people that directly submit messages to their app (no personal data will be required).

The qualitative data generated in SMARTLAGOON project will be useful for scientific use only, at least within the scope of the project. For its later use, it could benefit all the stakeholders of the Mar Menor, different Companies, citizens, Murcia province and other Spanish Provinces.

TYPE	CODE
MONITORING BODIES	
COORDINATION TEAM	SMLG_CTX_ TITLE OF DOC_ DD/MM/YEAR
PROJECT STEERING COMMITTEE	SMLG_PSCX_ TITLE OF DOC_ DD/MM/YEAR
EXECUTIVE BOARD	SMLG_EBX_ TITLE OF DOC_ DD/MM/YEAR
PROJECT INNOVATION COMMITTEE	SMLG_PICX_ TITLE OF DOC_ DD/MM/YEAR
WPL MEETING	SMLG_WPX_ WPMX_ TITLE OF DOC_ DD/MM/YEAR
COMMUNICATION WORKING GROUP	SMLG_CWGX_ TITLE OF DOC_ DD/MM/YEAR
ADVISORY BOARD	SMLG_ABX_ TITLE OF DOC_ DD/MM/YEAR
ETHICS EVALUATION COMMITTEE	SMLG_EEC_ TITLE OF DOC_ DD/MM/YEAR
INTERNAL REPORTING	
INTERIM PROGRESS REPORT (M9)	SMLG_IPRX_ PARTNER SHORT NAME_ FINANCIAL_ M1_ M9
	SMLG_IPRX_ PARTNER SHORT NAME_ TECH_ M1_ M9

FIRST PERIODIC REPORT (M18)	SMLG_IPRX_PARTNER SHORT NAME_FINAN- CIAL_M1_M18
	SMLG_IPRX_PARTNER SHORT NAME_TECH_M1_M18
INTERNAL REPORT (M27)	SMLG_IPRX_PARTNER SHORT NAME_FINAN- CIAL_M1_M27
	SMLG_IPRX_PARTNER SHORT NAME_TECH_M1_M27
SECOND PERIODIC REPORT (M36)	SMLG_IPRX_PARTNER SHORT NAME_FINAN- CIAL_M1_M36
	SMLG_IPRX_PARTNER SHORT NAME_TECH_M1_M36
THIRD REPORT (M48)	SMLG_IPRX_PARTNER SHORT NAME_FINAN- CIAL_M1_M48
	SMLG_IPRX_PARTNER SHORT NAME_TECH_M1_M48
FINAL REPORT (M48)	SMLG_IPRX_PARTNER SHORT NAME_FINAN- CIAL_M1_M48
	SMLG_IPRX_PARTNER SHORT NAME_TECH_M1_M48
TIMESHEET	SMLG_TS_PARTNER SHORT NAME_STAFF NAME_MX
WP DOCUMENTS	SMLG_WPX_TITLE OF DOC
DELIVERABLES	SMLG_WPX_DX.X_TITLE OF DELIVERA- BLE_DD/MM/YEAR

Table 1 List of names for main documents generated in the project

3. Fair Data

3.1 Open Data

Open access (OA) refers to the practice of providing online access to scientific information that is free of charge to the end-user and reusable.

Open access to scientific publications means free online access for any user. Under these definitions, 'access' includes not only basic elements - the right to read, download and print – but also the right to copy, distribute, search, link, crawl and analyze.

The open access mandate comprises 2 steps:

- depositing publications in repositories
- providing open access to them

SMARTLAGOON publications will be in a format that can be used and understood by a computer.

The majority of the data will be shared in open-source formats. Any software tool (not a part of the confidential deliverables mentioned in section 3.2) capable of decoding data structures and developed

under the SMARTLAGOON project will be released under an open source license and made accessible at the end of the project by a [GitHub](#) repository linked to the [Zenodo](#) dataset. The released code will remain hosted on GitHub and linked into the same dataset(s) with a specific DOI, so it can be open to reuse to decode the datasets.

Most of the generated data will be open-access articles that will have their own DOI. All the datasets produced within the project will be deposited in the Zenodo. In Zenodo the datasets and software generated on the project can be attached with an associated DOI.

This naming convention is the same used to identify the deliverables in the project.

When datasets are deposited in the Zenodo, it will generate keywords used for the description of the dataset.

3.2 Accessible Data

The majority of the data will be shared in open-source formats, so the data in the database will be made openly available as there is no private information stored in it. The data collected during the project will be available for 4 years after the finalization of the project. It will be linked to the project website and its maintenance will be the responsibility of UPV.

There are some deliverables that are confidential, which are those generated by companies. They have an associated exploitation plan:

D2.2	Sensing tech. video	2	PHO	DEM	CO	M18
D2.5	Crowd platform	2	UPV	DEM	CO (RES)	M48
D5.2	Webportal photo-based products	5	PHO	OTHER	CO	M36
D5.3	Short-term forecasting catchment	5	WIT	OTHER	CO	M24

The data repositories will be accessible through different GitHub repository, which some of them will be linked to the Zenodo dataset. A SMARTLAGOON organisation has been created on GitHub. This organisation will be structured in several GIT repositories that will store all the project documentation. The internal project management information will be stored in private repositories. Only authorized personnel by the repository manager, shall have access to documents of a confidential nature. The permissions will be provided through “teams” previously configured in GitHub. These teams are made up of the different project members interested in each type of information and will be required to have a GitHub account. A small GitHub tutorial is provided in Annex 2.

The SMARTLAGOON organisation’s public information will be directly accessible through this link <https://github.com/SMARTLAGOON> and available to the general public. This includes all the data generated by the SMARTLAGOON research listed above.

There is no documentation included about the software needed to access the data. Some software developments will be open uploaded into the GitHub community. But bear in mind that there are some software's that are confidential. Data and associated metadata, documentation and code of SMARTLAGOON project will be deposited at GitHub and Zenodo.

3.3 Interoperable Data

This version of the SMARTLAGOON DMP does not include a compilation of all the metadata about the data being produced in SMARTLAGOON project, but there are already several domains considered in the project which follows different rules and recommendations. This is a very early-stage identification of standards:

- Microsoft Office 2010 for text-based documents (or any other compatible version) .doc, .docx, .xls, .xlsx, .ppt, .pptx. Also, especially where larger datasets need to be dealt with, .csv and .txt file formats will be used. All finished and approved documents will also be made available as .pdf documents.
- Latex editor for scientific documents will be Overleaf; a collaborative cloud-based LaTeX editor used for writing, editing and publishing scientific documents.
- Illustrations and graphic design will make use of Microsoft Visio (Format: .vsd), Photoshop (Format: different types possible, mostly .png), and will be made available as .jpg, .psd, .tiff and .ai files.
- PFDs, PIDs and layouts will preferentially use inkscape.org, an open-source software for vector graphics. (Format: .svg), and will be made available as .png, .jpg and .pdf files. • MP3 or WAV for audio files.
- Quicktime Movie or Windows Media Video for video files.

These file formats have been chosen because they are accepted standards and in widespread use. Files will be converted to open file formats where possible for long-term storage.

Regarding Metadata will be comprised of two formats – contextual information about the data in a text based document and ISO 19115 standard metadata in an xml file. These two formats for metadata are chosen to provide a full explanation of the data (text format) and to ensure compatibility with international standards (xml format).

3.4 Reusable Data

For all datasets deposited on a public data repository, like ZENODO, the access is unlimited.

Some or all the following restrictions may be applied with (CREATIVE COMMONS LICENSING) on the dataset:

- Attribution: requires users of the dataset to give appropriate credit, provide a link to the license, and indicate if any changes we made.
- Non-Commercial: Prohibits the use of the dataset for commercial purposes by third parties out of the SMARTLAGOON project.
- Share A like: requires the others to use the same license as the original on all derivative works based on the original data.
- Access to the software information generated by the project will be Creative Commons licensed, with restricted access to confidential information.

The data collected during the project will be available for 4 years after the finalization of the project. It will be linked to the project website and its maintenance will be the responsibility of the UNIBO partner.

4. Allocation of resources

The costs for making data FAIR in SMARTLAGOON are the following

- **Data archiving at GitHub and ZENODO:** free of charge
- **Copyright licencing with Creative Commons:** free of charge.
- **Project App and website operation:** Digital Twin hosted on UPV servers. Website of the project will be hosted by a company to be hired by UNIBO.
- **Open Access publications:** All partners have allocated budget resources for open access publications. Moreover, Main contribution will be published through Open Research Europe initiative (<https://open-research-europe.ec.europa.eu/>). Open Research Europe offers free of charge and fast publication and open peer review for research stemming from Horizon 2020 funding across all subject areas.

The main responsible for data management in SMARTLAGOON will be the UPV, which will oversee guaranteeing the resources for the long-term preservation of the data. The Executive Board will oversee making decisions on what data to keep and for how long depending on the commitments established with the commission (at least 4 years after the end of the project) and the potential benefit for society. During the project data will be updated regularly as new results are submitted by the SMARTLAGOON partners.

5. Data Security

For the duration of the project, datasets will be stored on the GitHub/Zenodo storage system restricted to the project consortium. Every partner is responsible to ensure that the local copies of the data are stored safely and securely and in full compliance with European Union data protection laws. After the completion of the project, all the responsibilities concerning data recovery and secure storage will go to the repository storing the dataset.

As an initial step, only the SMARTLAGOON Partners will have access to the cloud storage where dataset and metadata are filed.

GitHub has the following security measures aimed at guaranteeing security in the logical access to the platform where the deliverables are stored:

- Advanced Universal Second Factor (U2F) double factor authentication.
- Secure Shell (SSH) and Hypertext Transfer Protocol Secure (HTTPS) protocols.
- Encryption and digital signature using the GNU Privacy Guard (GPG) tool.
- Encryption of access credentials using the bcrypt hash function.

Likewise, there are also some good security practices identified at GitHub in relation to:

- Identification, detection and prevention against technical vulnerabilities.

- Security in the management of databases.
- Management of security incidents.
- Carrying out external security audits.
- Security in the Cloud.

Following, publications and articles, the dataset deliverables and the final results will be shared through ZENODO and other database to promote the data making FAIR.

All data files will be transferred via secure connections and in encrypted and password-protected form (for example with the open source 7-zip tool providing full AES-256 encryption: <http://www.7-zip.org/> or the encryption options implemented in MS Windows or MS Excel). Passwords will not be exchanged via e-mail but in personal communication between the partners.

Finally, SMARTLAGOON, will ensure that any application to be used must be analysed and tested for compliance with data protection regulations and information security measures, considering security regulations and best practices. [CCN Security Guide](#)

6. Ethical Aspects

Ethics is an integral part of research from the beginning to the end. It is only by getting the ethics right that research excellence can be achieved. Ethical research conduct implies applying fundamental ethical principles and legislation to scientific research in all possible domains of research. Key ethical issues concerning research activities are identified and defined here according to EU and national directives. These issues are examined from the Project point of view and include recruitment of participants, information to participants, informed consent and data handling during the planned research activities.

The project activities will be carried out respecting the ethical implications, the regulations expressed in international texts and codes of practices, in particular the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons about the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation).

All the experimental data shall be collected and used in compliance with the European Parliament's directive and legislations.

Expected tools for data acquisition will be surveys, questionnaires, interviews and workshops.

Data acquisition will be performed in the context of WP2, “Efficient data collection and sensing”, WP3, “Innovative modelling of environmental processes”, WP4 “Socio-environmental dynamics modelling”, WP5 “Digital twin development”, and WP7 “Citizen’s engagement, collaboration, communication, dissemination and exploitation”. Involved persons will be adequately informed of the

project aims, expected results and limits of the research in the information supplied before getting the informed consent.

In these work-packages, the following personal data will be processed:

- ❖ Names and surnames
- ❖ Occupations and educational background (potentially)
- ❖ Email addresses
- ❖ Images/voices

An internal independent Ethics Evaluation Committee (EEC) is appointed to oversee the ethical concerns involved in this research. A report by the Ethics Advisor shall be submitted to the project consortium with the project reports in the monthly EBM (Executive Board meetings).

Due to the nature of the investigation activities to be performed, ethically relevant incidental findings are not expected. In case of incidents, details shall be collected and provided to the Ethics Advisor.

In case of other ethics issues, copies of ethical approvals by the competent authorities shall be submitted to the EC.

Data breach prevention will be accurately discussed and analysed in Deliverable 1.2 (POPD – Requirement no. 2).

The Ethics Evaluation Committee (Fundación Universitaria San Antonio, UCAM) and the Data controllers of all the members of the project.

7. Other Issues

Boot that captures information from social networks and media

SMARTLAGOON will use the Twitter API to crawl information from this social network. It is a set of programmatic endpoints that can be used to learn from and engage with the conversation on Twitter.

This API allows you to find and retrieve, engage with, or create a variety of different resources including the following:

- Tweets
- Users
- Direct Messages
- Lists
- Trends
- Media
- Places

Most activity on Twitter and other social network is public, and this information is the only one that is going to be used in this project. We consider this due to the fulfilling of these social network's Privacy Policies and because we acknowledge that certain types of private information published by

users on social media may carry higher risks than others if they are shared without permission. Our aim is to protect individuals from coming to physical harm because of their information being shared. By default, applications can only access public information on Twitter.

Capturing images and videos in the project

The consent procedure is a critical aspect to participate in an investigative activity. Before requesting consent, the investigator shall ensure that the potential participant, or her/his legal representative, has received written, and if desirable oral, information. This information should be provided in such a way that it is probable that the potential participant, or her/his legal representative, understands the contents. Furthermore, s/he should be given sufficient time to make a proper decision on the requested consent.

Participants shall be informed that they are free to withdraw from participation at any point, that their personal data will remain confidential, and that collected data will be analysed for the entire group of participants, rather than individually, thus securing their privacy and anonymity (as far as possible).

Data integrated in the Apps of the project.

The data of the Apps used in the project will be stored in AWS. AWS is designed to help you build a secure, resilient, efficient and high-performing infrastructure for your applications. The world-class security experts who oversee AWS's infrastructure also create and maintain a broad selection of innovative security services, which can help to meet the security and regulatory requirements.

AWS data protection services provide encryption and key management and threat detection that continuously monitor and protect the accounts and workloads of the project. In addition, AWS security configuration will be carried out in accordance with information security best practice guidelines.

Annex I: Tutorial on Zenodo – Open digital repository

(All) Research. Shared.— your one stop research shop!

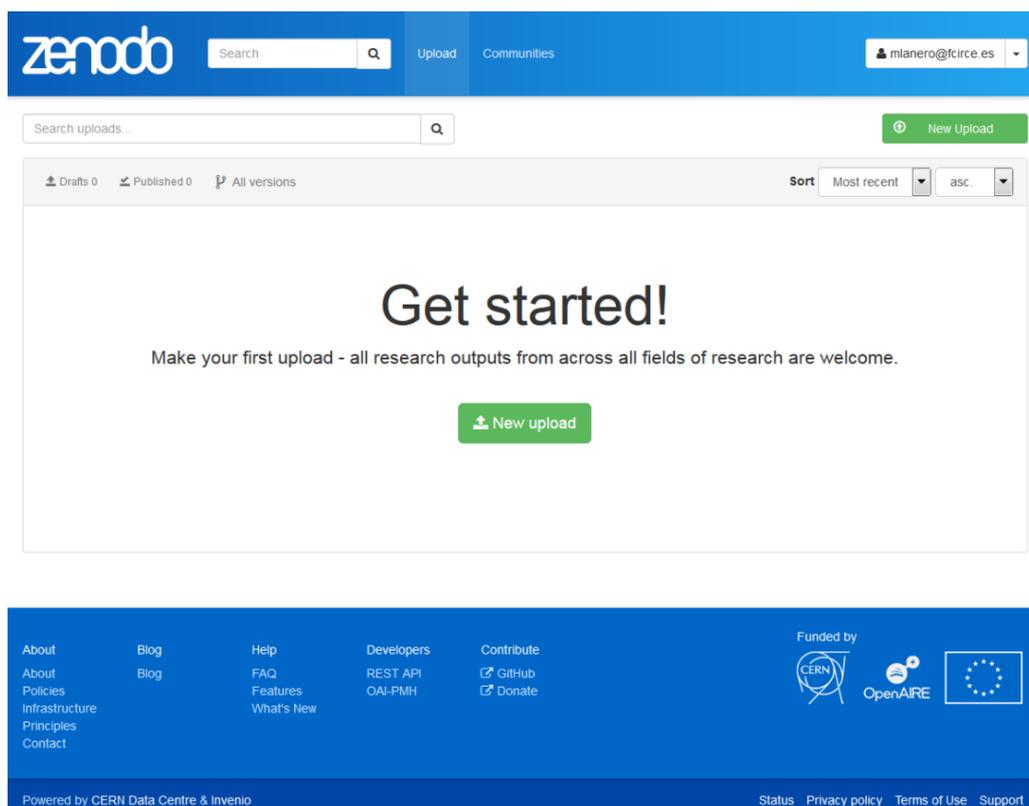
All research outputs from across all fields of research are welcome! Zenodo accepts any file format as well as both positive and negative results. We choose to promote peer-reviewed openly accessible research, and we curate the uploads posted on the front-page.

Citeable. Discoverable.— be found!

Zenodo assigns all publicly available uploads a Digital Object Identifier (DOI) to make the upload easily and uniquely citeable. Zenodo further supports harvesting of all content via the OAI-PMH protocol.

Communities — create your own repository

Zenodo allows you to create your own collection and accept or reject uploads submitted to it. Creating a space for your next workshop or project has never been easier. Plus, everything is citeable and discoverable!



The data must be classified according to given categories such as: dataset (i.e., tables of numerical data), image and others

Finally, the portal prompts for additional meta-data such as authorship of data and sharing policies. The structure of the data-set must be specified here as well

The screenshot shows a portion of the Zenodo metadata form. At the top is a required text input field. Below it is the 'Language' section, which includes a dropdown menu currently showing 'e.g.: eng, fr or Polish'. A note below the dropdown states: 'Optional. Primary language of the record. Start by typing the language's common name in English, or its ISO 639 code (two or three-letter code). See [ISO 639 language codes list](#) for more information.' The 'Keywords' section features a text input field with a search icon and a close icon, and a link below it that says '+ Add another keyword'. The 'Additional notes' section has a text area with a pencil icon and a note that says 'Optional.'

Safe— more than just a drop box!

Your research output is stored safely for the future in same cloud infrastructure as research data from CERN's Large Hadron Collider and using CERN's battle-tested repository software Invenio, which is used by some of the world's largest repositories such as INSPIRE HEP and CERN Document Server.

Reporting — tell your funding agency!

Zenodo is integrated into reporting lines for research funded by the European Commission via OpenAIRE. Just upload your research to Zenodo, and we will take care of the reporting for you. We plan to expand this feature with further funding agencies in the future, so stay tuned!

The screenshot shows the 'Funding' section of the Zenodo form. It has a title 'Funding' and a dropdown menu set to 'recommended'. Below the title is a paragraph: 'Zenodo is integrated into reporting lines for research funded by the European Commission via [OpenAIRE](#). Specify grants which have funded your research, and we will let your funding agency know!'. The 'Grants' section includes a dropdown menu with 'European Commission (EU)' selected and a text input field with the placeholder 'Start typing a grant number, name or abbreviation...'. A note below states: 'Optional. OpenAIRE-supported projects only. For other funding acknowledgements, please use the Additional Notes field. Note: a human Zenodo curator will need to validate your upload - you may experience a delay before it is available in OpenAIRE.' There is a '+ Add another grant' link at the bottom.

Flexible Licensing — not everything is under Creative Commons

Zenodo encourages you to share your research as openly as possible to maximize use and re-use of your research results. However, we also acknowledge that one size does not fit all. Therefore, we allow for uploading under a variety of different licenses and access levels.

The screenshot shows the 'License' and 'Communities' sections of the Zenodo form. The 'License' section is titled 'License' and has a dropdown menu set to 'required'. It features four radio button options for 'Access right': 'Open Access' (selected), 'Embargoed Access', 'Restricted Access', and 'Closed Access'. A note below states: 'Required. Open access uploads have considerably higher visibility on Zenodo.' The 'License' dropdown is set to 'Creative Commons Attribution 4.0'. A note below states: 'Required. Selected license applies to all of your files displayed on the top of the form. If you want to upload some of your files under different licenses, please do so in separate uploads. If you cannot find the license you're looking for, include a relevant LICENSE file in your record and choose one of the 'Other' licenses available ('Other (Open)', 'Other (Attribution)', etc.). The supported open licenses in the list are harvested from [opendefinition.org](#). If you think that an open license is missing from the list, please [contact us](#).' The 'Communities' section is titled 'Communities' and has a dropdown menu set to 'recommended'. It includes a paragraph: 'Any user can create a community collection on Zenodo ([browse communities](#)). Specify communities which you wish your upload to appear in. The owner of the community will be notified, and can either accept or reject your request.' The 'Communities' section has a text input field with the placeholder 'Start typing a community name...' and a '+ Add another community' link below it.

Annex II: Tutorial GitHub

What is GitHub?

GitHub is a code hosting platform for version control and collaboration. It lets you and others work together on projects from anywhere.

This tutorial teaches you GitHub essentials like repositories, branches, commits, and Pull Requests.

Step 1. Create a Repository

A repository is usually used to organize a single project. Repositories can contain folders and files, images, videos, spreadsheets, and data sets – anything your project needs. We recommend including a README, or a file with information about your project. GitHub makes it easy to add one at the same time you create your new repository. It also offers other common options such as a license file.

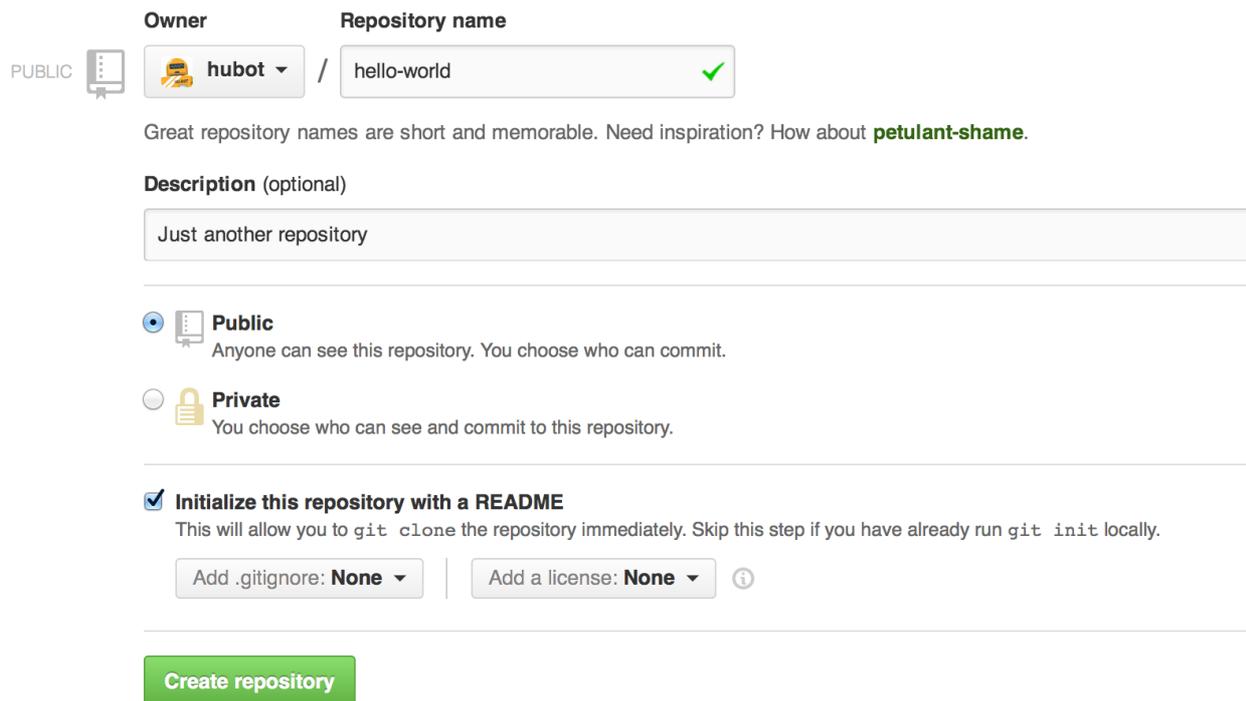
How to create a new repository

In the upper right corner, next to your avatar or identicon, click + and then select New repository.

Name your repository

Write a short description.

Select Initialize this repository with a README.



Owner **Repository name**

PUBLIC   hubot / hello-world ✓

Great repository names are short and memorable. Need inspiration? How about **petulant-shame**.

Description (optional)

Just another repository

Public
Anyone can see this repository. You choose who can commit.

Private
You choose who can see and commit to this repository.

Initialize this repository with a README
This will allow you to `git clone` the repository immediately. Skip this step if you have already run `git init` locally.

Add .gitignore: **None** | Add a license: **None** ⓘ

Create repository

Step 2. Create a Branch

Branching is the way to work on different versions of a repository at one time.

By default your repository has one branch named main which is considered to be the definitive branch. We use branches to experiment and make edits before committing them to main.

When you create a branch off the main branch, you're making a copy, or snapshot, of main as it was at that point in time. If someone else made changes to the main branch while you were working on your branch, you could pull in those updates.

Step 3. Make and commit changes

Now, you're on the code view for your readme-edits branch, which is a copy of main. Let's make some edits.

On GitHub, saved changes are called commits. Each commit has an associated commit message, which is a description explaining why a particular change was made. Commit messages capture the history of your changes, so other contributors can understand what you've done and why.

Step 4. Open a Pull Request

Pull Requests are the heart of collaboration on GitHub. When you open a pull request, you're proposing your changes and requesting that someone review and pull in your contribution and merge them into their branch. Pull requests show diffs, or differences, of the content from both branches. The changes, additions, and subtractions are shown in green and red.

As soon as you make a commit, you can open a pull request and start a discussion, even before the code is finished.

By using GitHub's @mentionsystem in your pull request message, you can ask for feedback from specific people or teams, whether they're down the hall or 10 time zones away.

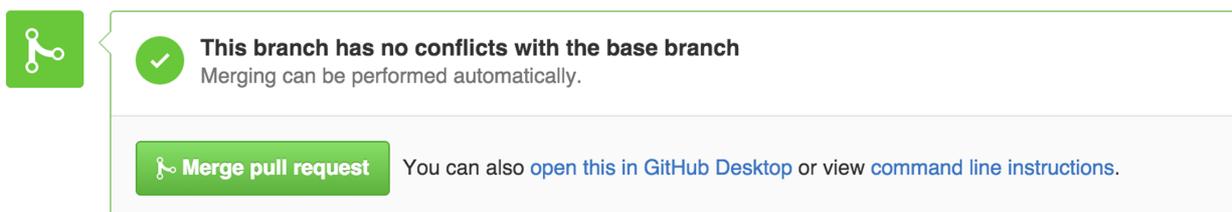
Step 5. Merge your Pull Request

In this final step, it's time to bring your changes together – merging your readme-edits branch into the main branch.

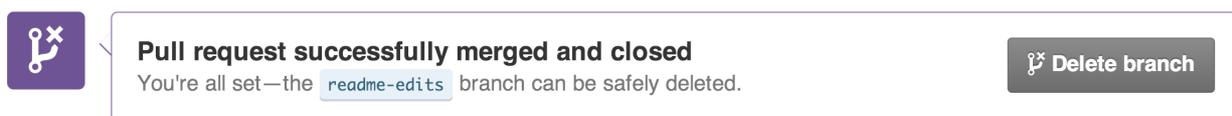
Click the green Merge pull request button to merge the changes into main.

Click Confirm merge.

Go ahead and delete the branch, since its changes have been incorporated, with the Delete branch button in the purple box.

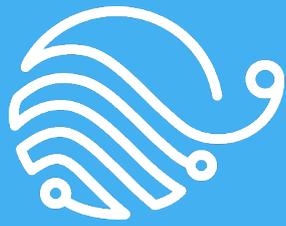


This screenshot shows a green notification box from GitHub. On the left is a green square icon with a white branching diagram. To its right is a green circle with a white checkmark. The text reads: "This branch has no conflicts with the base branch" followed by "Merging can be performed automatically." Below this is a green button with a white branching icon and the text "Merge pull request". To the right of the button is the text: "You can also [open this in GitHub Desktop](#) or view [command line instructions](#)."



This screenshot shows a purple notification box from GitHub. On the left is a purple square icon with a white branching diagram and a white 'x'. To its right is the text: "Pull request successfully merged and closed" followed by "You're all set—the `readme-edits` branch can be safely deleted." On the right side of the box is a dark grey button with a white branching icon and the text "Delete branch".

You can also check out our other Guides, YouTube Channel and On-Demand Training for more on how to get started with GitHub.



SMARTLAGOON

End of Deliverable 6.9



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017861.