D9.4 Data Management Plan (DMP)

ZLC

Deliverable	D9.4-Data Management Plan
Work package	WP9- Project Management
Work package title	Data Management Plan
Date	27.01.2021
Authors	ZLC, CERTH
Status	Final
Dissemination level	ORD

Table of Contents

1 Executive Summary

This deliverable describes the SPROUT project Data Management Plan. This deliverable outlines how data will be handled during the project life.

The data collection, sharing and storing process is described in this document, following a methodology aligned with the H2020 guidelines on data management. Specifically, the document describes how data will be managed in the project pilots, and the urban mobility shared data space. It also describes the interaction with the European Transport Mobility Forum.

In addition, data security and personal data protection questions are contained in this deliverable.

2 Introduction

The SPROUT Project Data Management Plan (DMP) describes the data management life cycle for the data to be collected, processed and/or generated during and after the end of the project execution.

The DMP includes information about the type of data to be collected, handled or generated during and after the end of the Project. This deliverable includes the methodology and standards to be used, regulation about open data and how the data will be preserved even after the end of the Project.

The SPROUT project participates in the Open Research Data Pilot proposed by the Commission to ensure that the project data and results are FAIR (findable, accessible, interoperable and reusable). The DMP is aligned and serves the SPROUT aim to contribute data to the open research pilot. Data sets which are candidates for sharing will be checked to ensure that:

- They are not confidential, that they do not include personal or commercially sensitive information.
- permission from the relevant stakeholders and/or data subjects has been obtained.
- sharing the data does not damage exploitation or IP protection prospects.

Accordingly, datasets will be reviewed by the *Data Protection Officer* and must be approved before becoming candidates for contribution to the open research pilot. These parties and the originating consortium partner(s) will then agree on licencing procedures, for example creative commons or public domain. Following in principle approval, SPROUT will then make the dataset available through the existing relevant and suitable open access repositories. Where data must be embargoed to support IP protection or exploitation, a time line for its release will be provided.

From the project Grant Agreement:

29.3 Open access to research data

Regarding the digital research data generated in the action ('**data**'), the beneficiaries must:

(a) deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user — the following:

    (i) the data, including associated metadata, needed to validate the results presented in scientific publications, as soon as possible;

    (ii) not applicable;

    (iii) other data, including associated metadata, as specified and within the deadlines laid down in the 'data management plan' (see Annex 1);

(b) provide information — via the repository — about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (and — where possible — provide the tools and instruments themselves).

This does not change the obligation to protect results in Article 27, the confidentiality obligations in Article 36, the security obligations in Article 37 or the obligations to protect personal data in Article 39, all of which still apply.

As an exception, the beneficiaries do not have to ensure open access to specific parts of their research data under Point (a)(i) and (iii), if the achievement of the action's main objective (as described in Annex 1) would be jeopardised by making those specific parts of the research data openly accessible. In this case, the data management plan must contain the reasons for not giving access.

Figure 1 Grant Agreement Number 814910- SPROUT. Article 29.3.

The SPROUT project will comply with the General Data Protection Regulation, the European Code of Conduct for research Integrity, and national regulations in SPROUT partner countries. Moreover, SPROUT will also comply with the following set of ethical requirements:

- NEC-Requirement No.1 (D1.1)
- H-Requirement No.2 (D1.2)
- NEC- Requirement No. 3 (D1.3)

The structure of the SPROUT DMP is based on the Data Management template provided by the EC, which is available at https://ec.europa.eu/research/participants/data/ref/h2020/gm/reporting/h2020-tpl-oa-data-mgt-plan_en.docx.

3 SPROUT Data Management Plan

Data collection / generation in this project are enablers for executing it according to the Description of Action (DOA). The content of this DMP has been adapted to the required development of the project across the different phases and needs.

The DMP will be updated over the course of the project whenever significant changes arise, such as (but not limited to):

- new data
- changes in consortium policies (e.g. new innovation potential, decision to file for a patent)
- changes in consortium composition and external factors (e.g. new consortium members joining or old members leaving).

The DMP will be updated as a minimum in time with the periodic evaluation/assessment of the project.

The SPROUT coordinator, ZLC, will be the data controller for the project. According to its privacy policy[1], the data controller can be reached at <u>privacy@zlc.edu.es</u>.

As it is stated in the Grant Agreement, the Project Coordinator (ZLC) in cooperation with the local Ethics and Data Protection Managers will perform ethics and data protection internal audits to ensure compliance with the Ethics requirement and Data Management Plan. Local Ethics and Data Protection Managers are the scientific partners in each 1st layer pilot city: Fundación Valenciaport (VPF), Ningbo Supply Chain International Institute China (NSCIIC), Centre for Budapest Transport (BKK), Venice International University (VIU), Israel Institute of Technology (TECHNION), and Institute of Logistics and Warehousing (ILiM).

3.1 Data Summary

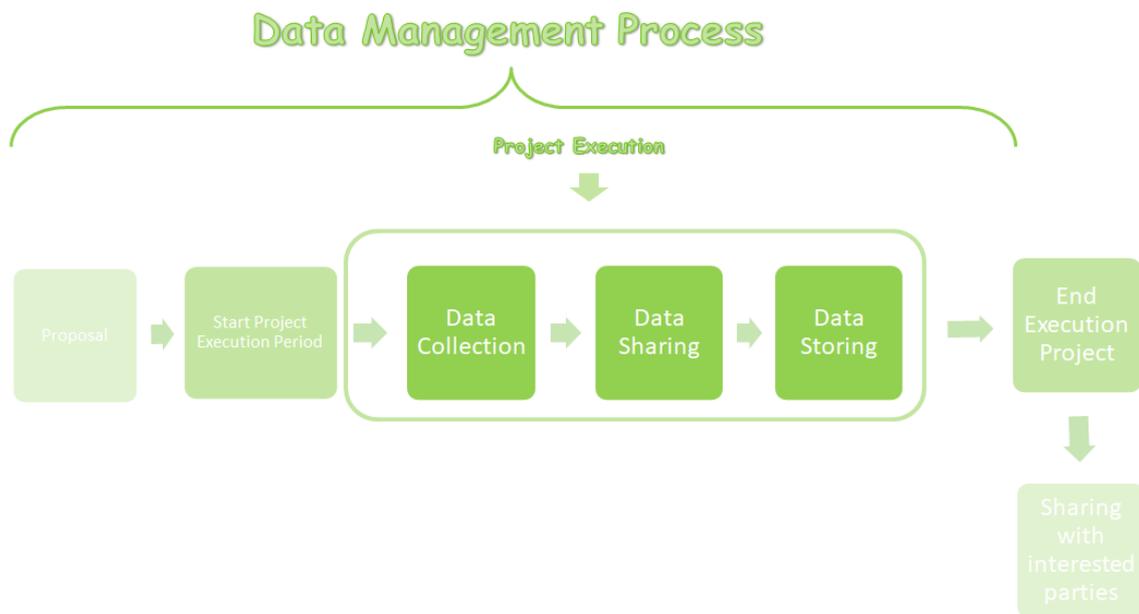Figure 1 illustrates the data management process in the SPROUT project.

---

[1] https://www.zlc.edu.es/privacy-policy/

Figure 2 Data Management Process

3.1.1 Data collection, sharing, storing, protection, retention and destruction

SPROUT is a Research and Innovation Action whose ultimate goal is to develop a city-led innovative policy response that is widely applicable to European cities. Project results will be disseminated during and after the project execution period to all potential stakeholders: cities, transport operators, logistics service providers, academia, policy makers, users and others.

Data collection is the process of gathering information that will be used during project execution. There are two types of data to be gathered during the SPROUT project:

- Quantitative data will be collected mainly to capture the situation before any SPROUT intervention and the situation after the pilots' implementation. Such quantitative data are mainly part of a catalogue of KPIs (key performance indicators). Quantitative data from the 1st layer cities will also be collected in the framework of the pilots (WP4), in order to develop an evaluation framework to guide its activities as well as process and outcome evaluation, assessing their impacts.
  Quantitative data collected from secondary sources (e.g. municipal statistics), will be based on primary data generation activities (e.g. traffic counts) and will have to be gathered through surveys and other similar methods.
  Data sources to be used include local/regional transport models, planning documents, travel surveys, statistics of planning agencies, infrastructure managers and transport operators, traffic management data, open data, crowdsourcing etc.

- Qualitative: Related data gathering methods include semi-structured interviews, open questions in surveys, notes, moderation cards and flip-chart scribblings from workshops with stakeholders, scenario writing etc.

This data will be collected in various formats, depending on the type of data and the specific research phase:

- Numerical values, captured either on paper during face-to-face surveys or electronically in case of online surveys.
- Quantitative data will be stored in suitable electronic databases, most likely XLS files.
- Visual data will be captured and stored mostly as digital photographs and scans.
- Qualitative (textual) data will be recorded on paper or in MS Word files
- Data analysis through System Dynamics Analysis software.

SPROUT will conform to the relevant EU regulations and directives (e.g. the Directive on data collection and electronic communications and the Data Protection Directive) relating to the collection and storage of big data information, applying all possible means for the protection of privacy and ethics in personal data collection, processing and dissemination, such as the Privacy Enhancing Technologies (PETs)[2]. SPROUT will pay special attention to the General Data Protection Regulation (EU) 2016/679 ("GDPR") that is the regulation in EU law on data protection and privacy for all individuals within the European Union (EU) and the European Economic Area (EEA). It also addresses the export of personal data outside the EU and EEA areas.

Furthermore, article 19 "Ethical Principles" of Regulation 1291/201360 is primary to the SPROUT research. All the research and innovation activities carried under Horizon 2020 shall comply with ethical principles and relevant national, EU and international legislation, including the Charter of Fundamental Rights of the European Union and the European Convention on Human Rights and its Supplementary Protocols.

The following principles define the SPROUT data management approach:

- Compliance with Legislation: Any real data collected for research and demonstration purposes will be handled in accordance with the Data Protection legislation in the concerned countries and each company handling the data will be registered to handle this type of information with their data protection authority.
- Use Limitation: All information leading to person identities will be encrypted and protected according to EU best practices e.g. using reference numbers instead of actual names.
- Security Safeguards: Personal data will be kept secure from potential abuse, during all required processing before storing, until personal identities are eliminated.
- Openness: All collection processes will be transparent on how data is collected, used, and shared.
- Accountability: SPROUT will be accountable to comply with all above principles.

---

[2] https://ec.europa.eu/commission/presscorner/detail/en/MEMO_07_159

Apart from the data generated by the SPROUT project itself, it will be necessary for the project partners to collect, process, and protect data:

- Personal data such as name and contact details and
- Data coming from cities. These would be related to urban population and economy, urban land use and accessibility, urban traffic, urban passenger and active transport characteristics, urban freight characteristics, etc.

The data collection will be part of the early project phase ("data collection"):
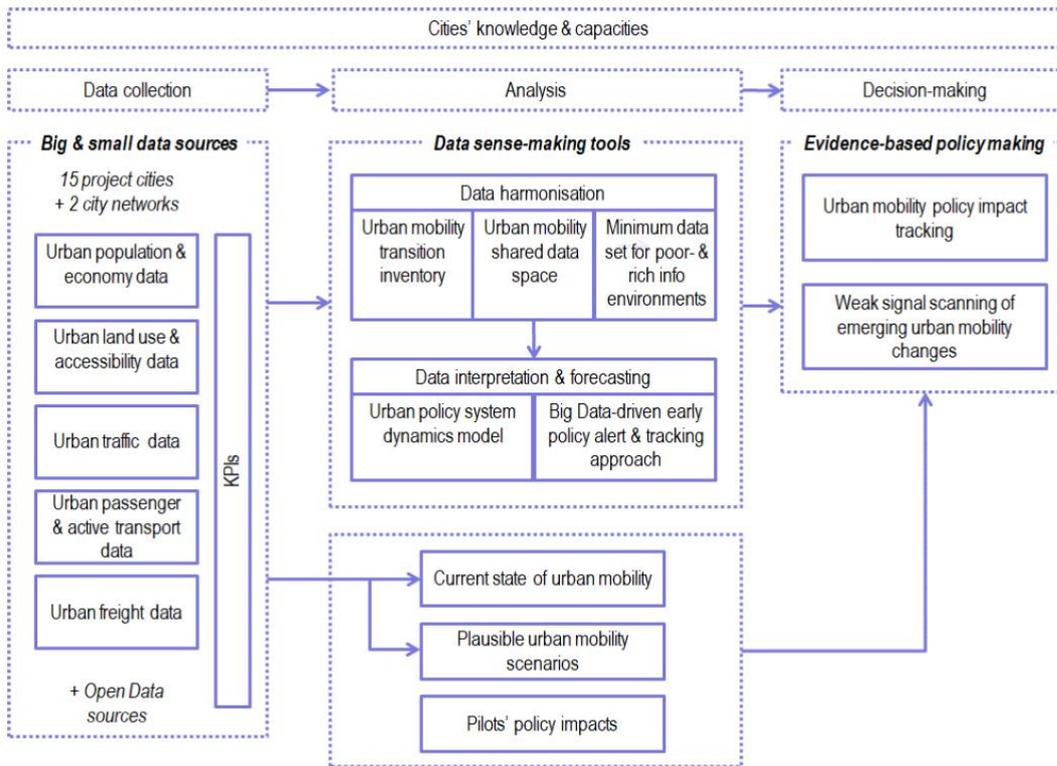

Figure 3 SPROUT data collection process (from SPROUT's Grant Agreement).

Indeed, SPROUT follows a participatory innovation approach, involving urban mobility policy makers, economic operators, and researchers in the form of an Open Innovation Community on Urban Mobility Policy.

This is therefore an interaction approach that uses, among others, methods based on participation and shared views of experts, focus groups and stakeholders, for instance in the form of scenario building workshops or the application of the Multi-Actor Multi-Criteria Analysis (MAMCA).

Secondary data will also be collected during the project execution period (through desk research, technology and innovation scanning, etc.).

The data sharing process will be arranged mainly using the project repository OneDrive, but also through mail exchanges among the people involved. Public documents will be shared through the SPROUT website. Also the open access procedure will apply for scientific publications that will derive from the project.

Participation data (e.g. number of followers) will also be gathered through SPROUT Social Networks and Website, as they provide data from visitors, contact forms and users of Twitter, LinkedIn also Research Gate in any case.

SPROUT will encrypt and anonymise the data, in order to protect personal identities, and it will decisively not implement processing algorithms leading to "personally identifying" results. Each information sheet will be given a reference number and will not be tied to the subject (person giving the information). Information gathered may include confidential commercial information. Stored data will be complying with data protection principles but will include all the required information for the scope of this research.

Additionally, SPROUT participants will:

- deny unauthorised persons to access the data-processing equipment used for processing the research datasets (equipment access control);
- prevent the unauthorised reading, copying, modification or removal of data media (data media control);
- ensure that persons authorised to use the data-processing system only have access to the data covered by their access authorisation (data access control).

Data will be stored at the project repository, Website, and Social Networks. Additionally, contact information data will be stored at project partners' servers and CRM.

For these purposes SPROUT will manage a dedicated contacts database and will ensure that personal contact details are not used elsewhere without specific agreement, thereby complying with data protection regulations.

The centre-piece legislation applicable to personal data processing activities carried out within the SPROUT project is the General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679). The regulation includes definitions, principles, legitimate basis for processing personal data, rights of data subjects and obligations of data controllers.

In order to comply with this regulation, the following disclaimer will appear in SPROUT's website when registering through the contact form to the newsletter and contact section:
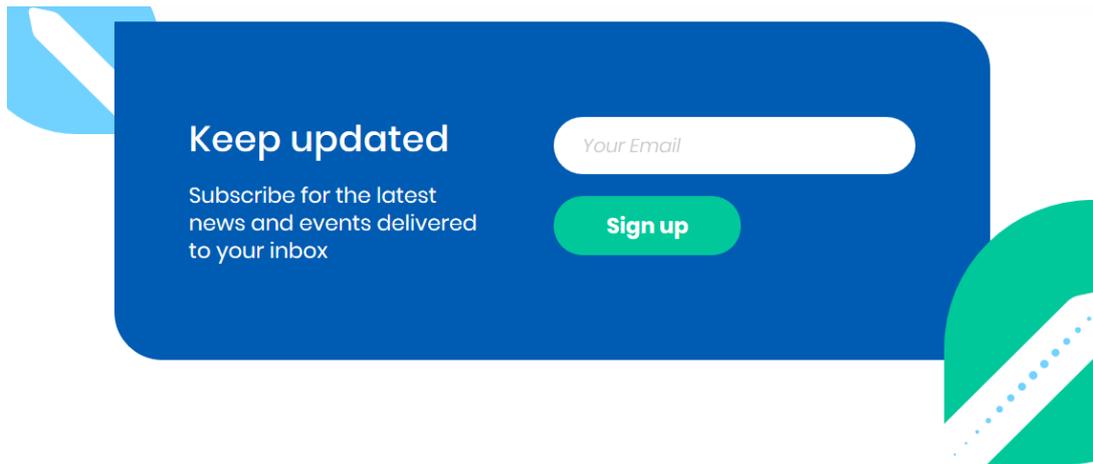
By subscribing to the SPROUT newsletter, you consent POLIS to store and treat your data with the purpose of sending you information related to the SPROUT project. Your data will not be given to third parties unless there is a legal obligation or it's essential for the execution of any organized events that you choose to attend. If you want to make use of your rights or want more information, please read POLIS's privacy policy.

Figure 4 Disclaimer on SPROUT website https://sprout-civitas.eu/subscribe-to-newsletter

3.1.2 Data re-use

The only data to be re-used by SPROUT is secondary statistical data, predominantly from municipal statistics.

3.1.3 Origin of the data

As sources of secondary data we foresee municipal statistics and also national databases.

3.1.4 "Data utility"

Public technical deliverables are defined in the Description of Action (DOA) and the data included there can be reused according to current and future needs.

Potential stakeholders to reuse data from SPROUT deliverables are experts involved along the project activities, members of SPROUT Open Innovation Community on Urban Mobility Policy, members of the European Transport and Mobility Forum, policy makers, industrial companies and other research institutions as input for their future research.

3.1.5 SPROUT Pilots

SPROUT 1st layer cities are the ones where the project pilots will run, the sustainability & policy impacts of innovative/emerging transport solutions will be assessed, and city-specific policy responses to harness these impacts will be tested and assessed. The 1st layer cities include: Valencia (Spain), Padova (Italy), Kalisz (Poland), Budapest (Hungary), Tel Aviv

(Israel), and Ningbo (China). The project will involve key stakeholders from pilot cities in preparing and implementing the pilots.

Data to be generated could include:

- mobility (freight & passenger) data
- data on the operational, financial and economic impacts of the new mobility services to be tested in the pilots
- data on the policy impacts of the new mobility services to be tested in the pilots derived from local workshops and interviews

SPROUT will get informed consent from anyone participating in the pilot scenarios so that the purpose(s) for the data collection (Purpose Specification) are agreed. Moreover, data privacy preserving solutions will be developed to ensure that participating partners share information and data with the required level of security and privacy. More specifically, a set of an informed consent forms and an information sheet will be shared (distributed) to the pilot participants.

This set of documents will be comprehensive and will contain at least the following:

1. The scope of research – purpose of personal data collection.

2. A description of the procedures to be followed and tests or measurements to be carried out.

3. Information about who is organising and funding the research.

4. A description of any reasonably foreseeable risk, discomfort or disadvantages.

5. A description of any benefits to the subject or to others which may reasonably be expected from the research avoiding inappropriate expectations.

6. A statement describing the procedures adopted for ensuring data protection/ confidentiality/privacy including duration of storage of personal data.

7. A reference to whom to contact for answers to questions about the research and research subjects' rights.

8. A statement offering to participants the opportunity to ask questions and to withdraw at any time from the research without consequences.

9. An explanation of what will happen with the data or samples at the end of the research period and if the data/ samples are retained or sent/sold to a third party for further research.

10. Information about what will happen to the results of the research.

Informed consent forms and an information sheets will be translated to local languages. More information can be found in D1.2 H-requirement No. 2.

3.1.6 SPROUT Urban Mobility Shared Data Space

In WP6, SPROUT will design an urban mobility shared data space for both passenger and freight transport, as a blueprint for setting up an ecosystem-based mechanism to provide urban mobility data in a harmonised way, while defining a minimum set of data to drive evidence-

based urban mobility policy making, customised to the data capabilities of both rich- and poor data environments.

The type of data and indicators collected would cover aspects related to SPROUT cities such as: transportation metrics (mode shares, average daily trips, average emissions/vehicle, %ZEVs, PT infrastructure, active mobility composition, infrastructure availability for sustainable mobility, etc.). No personal information will be gathered.

The data of the 'urban mobility data share' will be used to develop a multigranular (big & small) data-driven approach to provide descriptive & predictive analytics for: (i) scanning weak signals (early indicators) of emerging urban mobility changes; (ii) tracking cities' progress towards the achievement of specific policy goals. The data sets included in the 'urban mobility data share' will not be made available to the public, unless the cities providing them are considering them as open data.

3.1.7 Interaction with the European Transport Mobility Forum

The Open Innovation Community on Urban Mobility Policy (OIC) has been developed and hosted on the European Transport and Mobility (ETM) Forum platform. The ETM Forum is an informal association and online collaboration platform connecting stakeholders to discuss and advocate for user-centric approaches and cross-modal cooperation in transport research, planning and innovation. The Forum was set up as an outcome of the Mobility4EU Horizon2020 project in March 2019. The forum is based on working groups discussing specific topics related to user-centric transport. The online platform (www.etmforum.eu) provides a central portal for information, communication and collaboration between stakeholders that joined the ETM Forum, as well as an entry point for newcomers.

A specific working group titled 'Policies to harness the impacts of new urban mobility solutions (SPROUT Open Innovation Community)' has been created in the ETM Forum platform, with the aim of hosting the SPROUT OIC and their discussions. Stakeholders involved in the OIC will register and become active members of this working group. By registering to this online collaboration tool, OIC members will become a member of the ETM Forum and agree to the Statute[3], the privacy policy[4] and terms of service[5].

---

[3] https://web.etmforum.eu/wp-content/uploads/ETMF-Statutes-V1.5_clean.pdf
[4] https://community.etmforum.eu/privacy
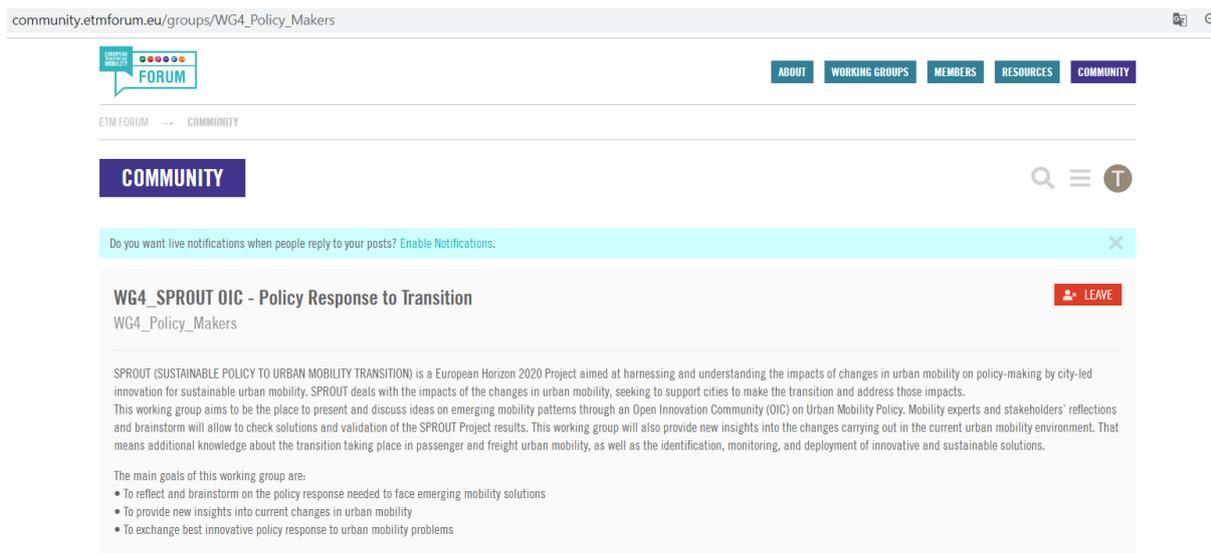[5] https://community.etmforum.eu/tos

Figure 5 Working Group dedicated to the SPROUT OIC in the ETM Forum.

3.2 FAIR data

3.2.1 Making data findable, including provisions for metadata

Deliverables will be named with easily findable and self-explanatory tittles to facilitate their location & identification. The information will be uploaded to SPROUT's shared OneDrive folder (available to consortium members). Public information and public deliverables will be accessible at the Resources section on the website.

Data collected from respondents of surveys, interviews and focus groups will remain anonymous. For this purpose, we will offer an Informed Consent Form to all respondents, which will include a written promise to this effect.

Other data, in particular primary quantitative data on traffic counts etc. collected for the purposes of the project, will be clearly labelled with suitable metadata and will be made available on the SPROUT website.

Academic papers that will result from the project will, by default, be equipped with a DOI – this will be taken care of by publishers.

File names should consist of:

Deliverable number D8.6

Project Acronym SPROUT

Short Deliverable name _Dissemination strategy_

Two digit version number _v01

Name initials _TC

A folder in the shared OneDrive has been created in order to keep a copy of older versions of the documents.

Deliverables will be named with easily findable and self-explanatory tittles to facilitate their location & identification. Public information and public deliverables are available in SPROUT's website. The consortium shared folder contains information restricted to consortium members.
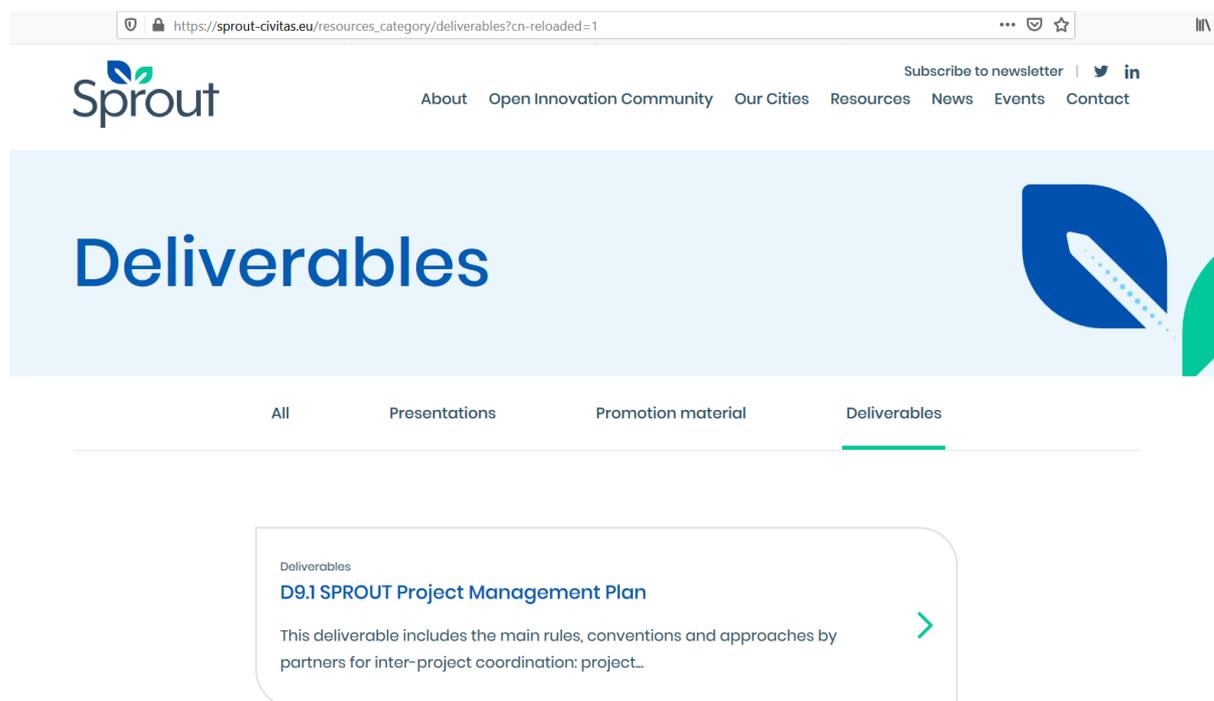

Figure 6 Screenshot of the area devoted to publishing public documents in SPROUT's website

3.2.2 Making data openly accessible and interoperable

During the project development, all the data included in public deliverables will be public and all this data will be posted at SPROUT's website.

SPROUT project is executed by a consortium of partners located at different countries so certain sensitive data may be stored internally in the partners' servers following their own internal criteria, always complying with the regulations in place.

SPROUT website contains a repository for public documents as a result of the project execution.

Project information will be accessible from all devices (computer, laptop, smart phone, tablets, etc.). Documents will be handed preferably in pdf. format or any other one that allows access to the majority of users.

Regarding interoperability, SPROUT intends to use standard vocabulary for all data types present in our data sets in order to allow interdisciplinary interoperability.

3.2.3 Increase data re-use (through clarifying licences)

The SPROUT project partners are keen to share as much information as possible, as long as their genuine commercial interests are not violated. Care will also be taken to release such information in widely readable, non-proprietary file formats. This is to facilitate the maximum

spread of the knowledge SPROUT will generate. For this reason, SPROUT will always genuinely consider the possibility to release data and deliverables under a creative commons licence.

The project will also establish and promote measures for open access publications in academic contexts by giving SPROUT partners the freedom to choose any of the two main open access publishing modalities: Gold OA (full or hybrid open access journals) and Green OA (self-archiving journal articles in OA repositories).

3.3 Allocation of resources

According to the open access to publications obligations in Horizon 2020, the SPROUT project partners must ensure open access to all peer reviewed scientific publications relating to the project results. The SPROUT partners can choose from one of the following routes:

- Self-archiving, e.g. archiving a published article or the final peer-reviewed version in an online repository before, alongside or after its publication. (Green OA)
- Open access publishing, e.g. placing the publication in open access mode (on the publisher's or journal's website). In this case, article processing charges will be allocated to the other direct costs category in the partners' budget. (Gold OA)

Additionally, the SPROUT OneDrive shared folder has been located in ZLC owned platform free of charge for all project partners. Therefore ZLC will be responsible for data management in the project shared platform. Likewise, POLIS manages SPROUT social networks and public website, including the contact form area. People registering to the contact form will be part of the SPROUT mailing list. This list will be used only for the purpose of the project. Therefore, POLIS will be responsible for data management in those sites. Hosting and maintenance of all these resources will be allocated to the SPROUT budget.

3.4 Data security

Project data will be allocated at each project partner server, and generally at the project shared OneDrive. As the final result of the project will be in the form of public deliverables no sensitive data will be generated, shared or stored.

SPROUT will not involve activities or results raising security issues nor 'EU-classified information' as background or results.

Dual-use items, listed in Annex I and in Annex IV of Council Regulation No 428/2009, or dangerous materials and substances are not going to be used in the SPROUT project.

In any case, a holistic security approach will be undertaken to protect the 3 mains pillars of information security: confidentiality, integrity, and availability. The security approach will consist of a methodical assessment of security risks followed by an impact analysis. This analysis will be performed on the personal information and data processed.

3.5 Ethical Aspects

The Ethics framework will ensure compliance with ethical guidelines set out in the Horizon 2020 Regulation, Grant Agreement, EU legislation, European Research Code of Conduct for Research Integrity and National legislation. The partners will ensure that research will be carried out taking into account reliability, honesty, respect and accountability.

Further ethical aspects are addressed in the project deliverables that comprise ethical requirements D1.1, D1.2, and D1.3.

Ethical issues will be managed by ZLC as Project Coordinator, acting as ethics supervisor, ensuring that all activities and deliverables are in compliance with the EU, national and institutional ethics requirements.

4 Conclusions

This deliverable describes SPROUT's data management plan. It shows how data will be collected, shared and stored during the life of the project. It also describes the process to ensure FAIR data (findable, accessible, interoperable and reusable). The centre-piece legislation applicable to personal data processing activities carried out within the SPROUT project is the General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679).

As a summary, SPROUT project approach has been to make public data available in the project website and storing all additional data in a common repository located at the project OneDrive.

More details can be found in the set of ethical requirements deliverables D1.1, D1.2 and D1.3.

5 References

European Union, Charter of Fundamental Rights of the European Union, 26 October 2012, 2012/C 326/02
https://www.europarl.europa.eu/charter/pdf/text_en.pdf

European Convention on Human Rights
https://www.echr.coe.int/Pages/home.aspx?p=basictexts&c=#n1359128122487_pointer

Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications). https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:en:HTML

Regulation (Eu) 2016/679 Of The European Parliament And Of The Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02016R0679-20160504&from=EN

Regulation (Eu) No 1291/2013 Of The European Parliament And Of The Council of 11 December 2013 establishing Horizon 2020 - the Framework Programme for Research and Innovation (2014-2020). https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32013R1291&from=EN