

Project acronym: **SPOT**
Project full title: **Social and innovative platform on cultural tourism and its potential towards deepening Europeanisation**

SPOT – DELIVERABLE

Data Management Plan

Deliverable number: D5.4

Due date: 30 June 2020

Nature¹: Other (ORDP – Open Research Data Pilot)

Dissemination Level²: PU

Work Package: WP5 Project and consortium management

Lead Beneficiary: MENDELU

Contributing Beneficiaries: all partners

Authors: Prof. Milada Šťastná, Dr. Alexander Chvorostov, Jan Zloch

Acknowledgement:

This project is funded by the *European Union's Horizon 2020 research and innovation programme* under **Grant Agreement No. 870644**

Disclaimer:

The content of this deliverable reflects only the authors' view. The *European Commission* and its *Research Executive Agency* are not responsible for any use that may be made of the information it contains.

¹ Nature: R = Report, P = Prototype, D = Demonstrator, O = Other

² Dissemination level:

PU = Public

PP = Restricted to other programme participants (including the Commission Services)

RE = Restricted to a group specified by the consortium (including the Commission Services)

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

Purpose and scope of the deliverable:

The objective of D5.4 (the Data Management Plan) is to describe what data the SPOT project will collect and proceed, whether and how it will be exploited or made accessible for verification and re-use and how it will be curated and preserved after the end of the project.

Document history

<i>Version</i>	<i>Date</i>	<i>Description</i>
0.0	06-Mar-20	Template designed
0.1	14-Apr-20	Draft outline produced
0.2	10-May-20	The first consolidated file produced
0.3	25-May-20	The final version of the consolidated file produced
1.0 (final)	26-June-20	Approved by PMB and the Coordinator
	30-June-21 (scheduled)	Annual update 2021
	30-June-22 (scheduled)	Annual update 2022
	December 2022 (scheduled)	Final SPOT DMP

Contents

0 Preamble: Background requirements to H2020 projects with regard to data management.....	6
0.1 H2020 requirements to data management.....	6
0.1.1 Data Management Plan – General Definition.....	6
0.1.2 Research Data Management Plans during the Project Life Cycle	6
0.2 H2020 Open access & Data management.....	6
0.3 Further H2020 recommendations regarding DMP.....	7
0.4 SPOT project and Open Data Access Pilot.....	8
1 SPOT DMP: Data Summary	9
1.1 What is the purpose of the data collection/generation and its relation to the objectives of the project?	9
1.2 What types and formats of data will the project generate/collect?.....	10
1.3 Will you re-use any existing data and how?.....	12
1.4 What is the origin of the data?.....	13
1.5 What is the expected size of the data?	18
1.6 To whom might it be useful ('data utility')?	18
2 SPOT DMP: FAIR data	19
2.1 Making data findable, including provisions for metadata.....	19
2.1.1 Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism?	19
2.1.2 What naming conventions do you follow?	19
2.1.3 Will search keywords be provided that optimize possibilities for re-use?	20
2.1.4 Do you provide clear version numbers?.....	20
2.1.5 What metadata will be created?.....	20
2.2 Making data openly accessible.....	20
2.2.1 Which data produced and/or used in the project will be made openly available as the default?	20
2.2.2 How will the data be made accessible (e.g. by deposition in a repository)?	20
2.2.3 What methods or software tools are needed to access the data?	21
2.2.4 Is documentation about the software needed to access the data included?.....	21
2.2.5 Is it possible to include the relevant software (e.g. in open source code)?.....	21
2.2.6 Where will the data and associated metadata, documentation and code be deposited?	21
2.2.7 Have you explored appropriate arrangements with the identified repository?	22
2.2.8 If there are restrictions on use, how will access be provided?	22
2.2.9 Is there a need for a data access committee?	22
2.2.10 Are there well described conditions for access?	22
2.2.11 How will the identity of the person accessing the data be ascertained?	22
2.3 Making data interoperable.....	22
2.3.1 Are the data produced in the project interoperable?	22
2.3.2 What data and metadata vocabularies, standards or methodologies will you follow to make project data interoperable?	22
2.3.3 Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?.....	23
2.3.4 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?	23
2.4 Increase data re-use (through clarifying licenses).....	23
2.4.1 How will the data be licensed to permit the widest re-use possible?	23
2.4.2 When will the data be made available for re-use?	23

2.4.3	Are the data produced and/or used in the project useable by third parties, in particular after the end of the project?	23
2.4.4	How long is it intended that the data remains re-usable?	23
2.4.5	Are data quality assurance processes described?	23
3	SPOT DMP: Allocation of resources	24
3.1	What are the costs for making data FAIR in your project?	24
3.2	How will these be covered?	24
3.3	Who will be responsible for data management in your project?.....	24
3.4	Are the resources for long term preservation discussed?	25
4	SPOT DMP: Data security	26
4.1	What provisions are in place for data security?	26
4.2	Is the data safely stored in certified repositories for long term preservation and curation?.....	26
5	SPOT DMP: Ethical aspects	27
5.1	Are there any ethical or legal issues that can have an impact on data sharing?	27
5.2	Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?	27
6	SPOT DMP: Other issues	28
6.1	Do you make use of other national/funder/sectorial/departmental procedures for data management?	28
6.2	Other SPOT project specific issues	28

ABBREVIATIONS

Acronym	Full name
CA	Consortium Agreement
DoA	Description of Actions
DMP	Data Management Plan
FAIR	Findable, Accessible, Interoperable, and Reusable
FPM	Framework Project Management
GA	Grant Agreement
ICT	Information & Communication Technologies
MS	Milestone
OA	Open Access
PC	Project Coordinator
PM	Project Manager
S/T	Scientific/ Technical
SC	Steering Committee
SPOT	Social and Innovative Platform on Cultural Tourism
SSL	Secure Sockets Layer
TL	Task Leader
WP	Work Package
WPL	Work Package Leader

0 Preamble: Background requirements to H2020 projects with regard to data management

0.1 H2020 requirements to data management

Detailed guide and rules to data management in H2020 are published here:

https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

In particular, the rules state the following:

0.1.1 Data Management Plan – General Definition

Data Management Plans (DMPs) are a **key element** of good data management. A DMP describes the data management life cycle for the data to be collected, processed and/or generated by a Horizon 2020 project. As part of making research data findable, accessible, interoperable and reusable (**FAIR**), a **DMP should include information** on:

- the handling of research data during & after the end of the project,
- what data will be collected, processed and/or generated,
- which methodology & standards will be applied,
- whether data will be shared/made open access, and
- how data will be curated & preserved (including after the end of the project).

A DMP is required for all projects participating in the extended ORD pilot unless they are opt-out of the ORD pilot. However, projects that opt-out is still encouraged to submit a DMP voluntarily.

0.1.2 Research Data Management Plans during the Project Life Cycle

First version: Once a project has had its funding approved and has started, you **must submit the first version of your DMP** (as a deliverable) within the first 6 months of the project. The Commission provides a DMP template in the annex, the use of which is recommended but voluntary.

Updates: The DMP needs to be updated throughout the project whenever significant changes arise, such as (but not limited to):

- new data,
- changes in consortium policies (e.g. 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 should be updated as a minimum in time with the periodic evaluation/assessment of the project.

- If there are no other periodic reviews foreseen within the grant agreement, then such an update needs to be made in time for the final review at the latest.
- Furthermore, the consortium can define a timetable for review in the DMP itself.

0.2 H2020 Open access & Data management

Requirements regarding the open access and data management in H2020 project are published here:

https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm

What Is Open Access (OA)?

Open access can be defined as the practice of **providing on-line access to scientific information that is free of charge to the reader**. In the context of R&D, open access typically focuses on access to 'scientific information' or 'research results', which refers to two main categories:

- Peer-reviewed scientific research articles (primarily published in academic journals)
- Research data

OPEN ACCESS TO PUBLICATIONS

Under Horizon 2020, each **beneficiary must ensure open access to all peer-reviewed scientific publications relating to its results**.

Beneficiaries can freely choose between the most appropriate route towards open access for them:

Self-archiving (also referred to as 'green' open access) means that a published article or the final peer-reviewed manuscript is archived (deposited) in an online repository before, alongside or after its publication. Repository software usually allows authors to delay access to the article ('embargo period') If this route is chosen beneficiaries must ensure open access to the publication within a maximum of six months (twelve months for publications in the social sciences and humanities).

Open access publishing (also referred to as 'gold' open access) means that an article is immediately provided in open access mode (on the publisher/journal website). Publishers sometimes charge so-called Article Processing Charges (or APCs) to make articles open. Such **costs are eligible for reimbursement during the duration of the project** as part of the overall project budget. In the case of gold open access publishing, open access must be granted at the latest on the date of publication and you also have to deposit a copy in a repository.

OPEN ACCESS TO RESEARCH DATA

Research data is information (particularly facts or numbers) collected to be examined and considered and to serve as a basis for reasoning, discussion or calculation.

Open access to research data - the right to access and reuse digital research data under the terms and conditions set out in the Grant Agreement.

Besides, these requirements provide the following definitions:

Data set: The Open Research Data Pilot applies primarily to the data needed to validate the results presented in scientific publications. Other data can also be provided by the beneficiaries voluntarily.

Data Management Plan: Participating projects will be required to develop a Data Management Plan (DMP), in which they will **specify what data will be open**: detailing what data the project will generate, whether and how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved.

0.3 Further H2020 recommendations regarding DMP

The Research Data Alliance provides a [Metadata Standards Directory](#) that can be searched for discipline-specific standards and associated tools.

The [EUDAT B2SHARE](#) tool includes a built-in license wizard that facilitates the selection of an adequate license for research data.



Useful listings of repositories include:

[Registry of Research Data Repositories](#)

Some repositories like [Zenodo](#), an [OpenAIRE](#) and CERN collaboration, allow researchers to deposit both publications and data while providing tools to link them.

Other useful tools include [DMP online](#) and platforms for making individual scientific observations available such as [ScienceMatters](#).

0.4 SPOT project and Open Data Access Pilot

SPOT project has opted to take part in the H2020 Open Research Data Pilot (ORD Pilot) and hence produces the Data Management Plan (DMP) as required. DMP is first issued by 30 June 2020 and will be reviewed and updated annually.

1 SPOT DMP: Data Summary

1.1 What is the purpose of the data collection/generation and its relation to the objectives of the project?

This document, D5.4 Data Management Plan (DMP) is a deliverable of the SPOT project, which is funded by the European Union's H2020 Programme under Grant Agreement #870644. SPOT's main goal is to develop a new approach to understanding and addressing cultural tourism and to promote the development of disadvantaged areas. Specifically, it will identify different layers of data and capitalise on existing practice. The Consortium strongly believes in the concepts of open science, and in the benefits that the European innovation ecosystem and economy can draw from allowing the reuse of data at a larger scale. Besides, there is a need to gather experience in real conditions, structural performance, and operating data in cultural tourism. Therefore, this project proposes delivering besides the obtained and collected data also the innovation tool (SPOTIT) to society.

The SPOT project participates in the Pilot on Open Research Data launched by the European Commission along with the H2020 programme. The use of a Data Management Plan is required for all participating projects. The DMP is not a fixed document; on the contrary, it will evolve during the lifespan of the project. This first version of the DMP includes an overview of the datasets to be produced by the project and the specific conditions that are attached to them. The next versions of the DMP will get into more detail and describe the practical data management procedures implemented by the SPOT project. The expected types of research data related to cultural tourism that will be collected or generated along the project lie in the following categories: 1) Cultural tourism quantitative and qualitative data 2) Spatial data and 3) Case study data.

All data is necessary to fulfil **the main project objective** which is to provide a new approach to cultural tourism to reflect patterns of travel in the 21st century, but also **specific project objectives** focused on the scientific evidence base, database of good practice examples, place-based identities, social and economic cohesion, the role of local stakeholders, different challenges facing contrasting types of cultural tourism and developing an innovative tool for cultural tourism.

Purpose of the Data Management Plan

The purpose of the DMP is to provide an analysis of the main elements of the data management policy that will be used by the Consortium concerning the project research data. The DMP covers the complete research data life cycle. It describes the types of research data that will be generated or collected during the project, the standards that will be used, how the research data will be preserved and what parts of the datasets will be shared for verification or reuse. It also reflects the current state of the Consortium agreements on data management and must be consistent with exploitation and IPR requirements.

The DMP is not a fixed document but will evolve during the lifespan of the project, particularly whenever significant changes arise such as dataset updates or changes in Consortium policies. This document is the first version of the DMP, delivered in Month 6 of the project. It includes an overview of the datasets to be produced by the project and the specific conditions that are attached to them. The next versions of the DMP will get into more detail and describe the practical

data management procedures implemented by the SPOT project concerning the IT tools developed in WP1. At a minimum, the DMP will be updated in Month 18 and Month 30 respectively.

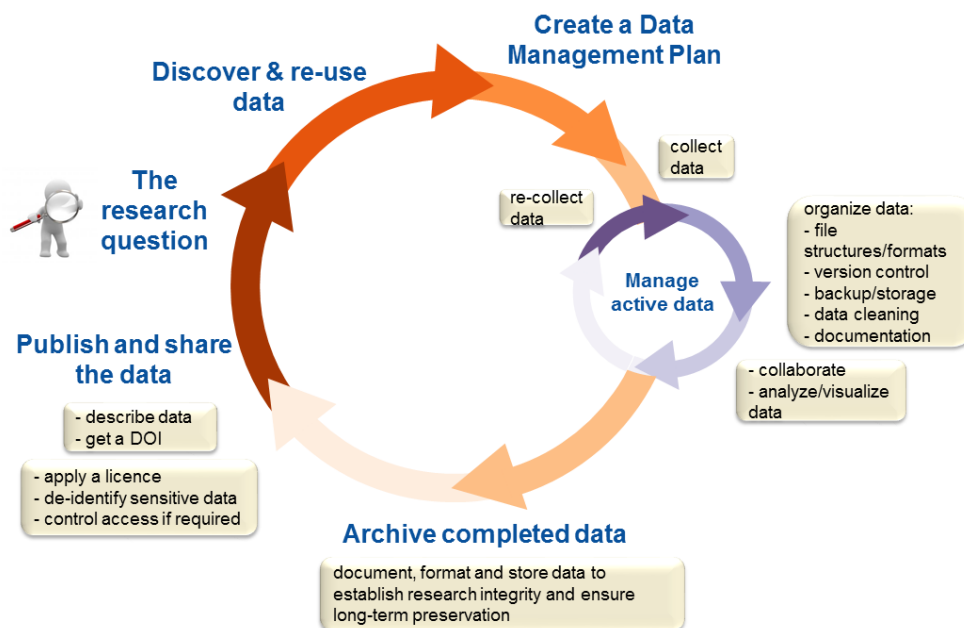


Figure 1. Research Data Management Lifecycle diagram. Adapted from: The University of California, Santa Cruz (2018) Data Management LibGuide, Retrieved from <https://guides.library.ucsc.edu/datamanagement> [1]

1.2 What types and formats of data will the project generate/collect?

For this first release of DMP, the data types that will be produced during the project are focused on the Description of the Action (DoA) and the results obtained in the first months of the project.

According to such consideration, Table 1.1 reports a list of indicative types of research data that SPOT will produce. These research data types have been mainly defined in WP1, including data structures, sampling and processing requirements, as well as relevant standards. This list may be adapted with the addition or removal of datasets in the next versions of the DMP to take into consideration the project developments. A detailed description of each dataset is given in the following sections of this document.

Table 1.1 List of indicative types of research data.

#	Dataset	Lead partner	Related WP(s)
1	Cultural tourism quantitative and qualitative data	WR	WP2, WP3
2	Spatial data	BIU	WP1, WP2
3	Case study data	All partners	WP1, WP2, WP3

Research data linked to exploitable results will not be put into the open domain if they compromise its commercialisation prospects or have inadequate protection, which is an H2020 obligation. The rest of the research data will be deposited in an open access repository. When the research data is linked to a scientific publication, the provisions described in DoA will be followed. Research data needed to validate the results presented in the publication should be deposited at the same time for “Gold” Open Access or before the end of the embargo period for “Green” Open Access. Underlying research data will consist of selected parts of the general datasets generated, and for which the decision of making that part public has been made. Other datasets will be related to any public report or be useful for the research community. They will be selected parts of the general datasets generated or full datasets (i.e. up to 2 years of key operating data), and be published as soon as possible.

Procedures to Collect, Process and Manage Data according to DoA

The research is not intended to collect personal data, including sensitive one, and neither its processing. The participating organisations in the SPOT are aware that research data management deserves ethical attention and are transparent about the way their research is done, and how data are handled. All data management will comply with applicable national and international law. The SPOT participants will commit to following the principles described in the Personal Data Act (523/1999) and General Data Protection Regulation EU Directive 95/46/EC.

SPOT proposal includes general procedures on how the data management plan (DMP) of the project will be produced. The DMP will comply with the existing framework for data management plans extensively applied in EU Universities and research organisations. The DMP will also comply with all national and EU legislation. The DMP will apply to all researchers of the SPOT consortium. The plan for data-handling will include aspects such as:

- Data safety procedures which are compliant with privacy by design and protection of confidentiality (including GDPR).
- Procedures for data collection, storage, protection, retention, transfer if any, destruction or reuse.
- Choice of data, sharing and ownership of data, choice of software.
- Use of methods when observing participants.

Subject information (i.e., personal identifiers) will be stored separately from their data. Researchers who want access to the data will be delivered spreadsheets. No spreadsheets will contain personal identifiers about any of the subjects (only subject numbers will be used as identifiers).

Responsibility

Each SPOT partner has to respect the policies set out in this DMP. Datasets have to be created, managed and stored appropriately and in line with applicable legislation. The Project Coordinator has a particular responsibility to ensure that data shared through the SPOT website are easily available, but also that backups are performed and that proprietary data are secured. WR, as WP1 leader, will ensure dataset integrity and compatibility for its use during the project lifetime by different partners. Validation and registration of datasets and metadata is the responsibility of the partner that generates the data in the WP. Metadata constitutes an underlying definition or description of the datasets, and facilitate finding and working with particular instances of data. Backing up data for sharing through open access repositories is the responsibility of the partner possessing the data. Quality control of these data is the responsibility of the relevant WP leader, supported by the Project Coordinator. If datasets are updated, the partner that possesses the data has the responsibility to manage the different versions and to make sure that the latest version is available in the case of publically available data. WP1 will provide naming and version conventions. Last but not least, all partners must consult the concerned partner(s) before publishing data in the open domain that can be associated with an exploitable result.

1.3 Will you re-use any existing data and how?

Data from available databases

The themes have been selected to draw out some of the most significant potential challenges posed for transmission and benefit from cultural tourism in Europe covered by the cultural tourism platform, whilst also reflecting the expertise of the consortium partners. Within each theme, research will focus on several illustrative examples located in the case study areas. The methods employed in each of the individual thematic WPs are also outlined in more detail below. They include the collection and analysis of both quantitative and qualitative data.

Quantitative data is employed to measure and enumerate key aspects of Europeanization impacting on involved countries, for example, tourism and the promotion of the cultural tourism destinations through global networks. Data will primarily be sourced from published sources, networks, agencies, galleries, libraries, archives, websites, social media platforms, museums, and other public institutions, including Eurostat, ESPON, national statistical offices and different authorities. Preliminary checks have confirmed availability, and where appropriate data is not currently available in existing datasets, actions for the collection of new primary data have been built into WP programmes. The analysis of quantitative data will be assisted by visualization methods, building on existing models.

Qualitative data collection methods will include semi-structured interviews with a range of subjects, including key informants, community representatives, European/national/regional/local authorities, inhabitants etc.; biographical interviews with tourists; group interviews; and document analysis. The qualitative methods will enable the motivations and experiences of actors involved in engaging with globalisation processes and their effects on local, municipal level in particular countries to be explored. Additionally, soliciting critical assessments and evaluations of key issues, and assisting in the

identification and documentation of examples of good practice within the cultural tourism platform will also be explored.

Data for the innovative tool

In WP3 a new user-friendly Innovation Tool will be developed, tested and disseminated. A GIS-based website which integrates interactive maps and other interfaces will help create interactive communities that enable sharing and exchange of ideas. Different map layers will contain relevant cultural tourism information, with some layers being created specifically for any given region. The layers will be designed as dynamic ones meaning that the system updates existing data layers itself while linking to various information resources and systems including publicly available remote sensing data of the EU.

See, for instance, e.g. the Copernicus Program <https://www.copernicus.eu/en>, Europeana thematic Collections <https://www.europeana.eu/portal/en>, Cultural routes <https://www.coe.int/en/web/cultural-routes>.

Personal data

In case personal data will be collected through interviews or surveys, all procedures will be in line with the ethical principles as defined in the ethics deliverables D.6.1 - 6.8. Data will be collected according to minimisation principles, anonymised and securely stored on a relevant server with the periodical backup.

1.4 What is the origin of the data?

Data will be collected on the municipal, national, but also regional and European level. Majority of the data are available for free from the databases provided by the national bodies involved in the tourism or statistical offices. Data, which is not available in such databases will be collected personally by the researchers on the local and national level, where appropriate. The survey, or interview will be the main tool to collect this kind of data where Informed Consent Forms and Information Sheets for Participants in the national language will be presented before the action starts. Data will be collected in accordance with the Data template prepared and presented as in D1.1.

Following data will be gathered:

1. There will be an analysis based on existing data and published analysis. The deliverable will be a report in which reference will be done to the relevant publications and data. In many cases, the original data sources will be best fitted to meet FAIR data principles. It is, for example, not necessary to republish EUROSTAT data. EUROSTAT data is better findable at EUROSTAT and EUROSTAT provides metadata. However, in cases in which different data sources are combined in a novel data set this will be made accessible to a data repository. Currently, it is not known what data this will be. We expect that the data will have less stand-alone value than it will have in the context of the reporting. This means that potential data users aiming to dig deeper in the subjects studied will access most of this data through the reports. This target group will be reached by dissemination actions foreseen in WP4.

2. Fifteen case studies will be executed. In these case studies semi-structured in-depth interviews will be held. Interviews will be held in local languages. In most regions, these statements must be translated in English to incorporate them in the reports.

The informed consent forms have been presented as part of deliverable D6.1. In the case studies, other data will also be gathered from a variety of sources. The case study reports will use adequate reference methods to these data sources. The case study reports will be reviewed by the WP leader, assisted by task leaders of this WP and the coordinator of SPOT, on personal data issues.

Other data (usually textual analysis using references to external sources) will be published as part of a deliverable or will be kept confidential (interview transcripts in various original languages) as it involves personal data.

Datasets and Publications for DMP

In this section, a list of all existing or foreseeable results for dissemination is presented, separated into public deliverables, publications and open research data. For each result and in accordance to the FAIR data management guideline [4] FAIR, 2013 we provide a description, name the standards used for storage and metadata (to make data findable & interoperable), and define which open access platform is chosen. In summary, the SPOT partners will comply with the ethical principles as set out in Article 34 of the Grant Agreement, which asserts that all project activities must be carried out in compliance with E.U. legislation towards data handling and preservation, aligned with the recent EC GDPR requirements.

Deliverable Number	Deliverable Title	WP number	Lead beneficiary	Type	Dissemination level	Due Date (in months)
D1.2	Common Report of cultural tourism as developed in the project	WP1	6 - WR	Report	Public	24
D1.3	Report of the information and statistical data	WP1	6 - WR	Report	Public	24
D1.4	Report of the consolidated interviews in case study areas	WP1	6 - WR	Report	Public	24
D1.5	Report of good practices across case study regions	WP1	6 - WR	Report	Public	24
D2.1	Report of Policies, practices and strategies.	WP2	9 - UNIABDN	Report	Public	18
D2.2	Summary Report on stakeholder involvement	WP2	9 - UNIABDN	Report	Public	24

D2.3	Summary Report of impact evaluations of cultural tourism on target areas for types of cultural tourism	WP2	9 - UNIABDN	Report	Public	30
D2.4	Summary Report on the role of cultural tourism for the development of place identities and appreciation of “otherness” and impacts on minorities	WP2	9 - UNIABDN	Report	Public	30
D2.5	Policy Guidelines and Briefings	WP2	9 - UNIABDN	Report	Public	30

SPOT Project Public deliverables

We are considering the SPOT Project public deliverables as part of the data management plan. The following table presents the list of public deliverables of the SPOT project.

Deliverable Number	Deliverable Title	WP number	Lead beneficiary	Type	Dissemination level	Due Date (in months)
D3.3	SPOTIT tool (Final version of the tool)	WP3	2 - BIU	Other	Public	30
D4.1	Setting up project website	WP4	5 - KRTK	Websites, patents filling, etc.	Public	2
D4.2	Electronic newsletter I	WP4	5 - KRTK	Other	Public	12
D4.3	Electronic newsletter II	WP4	5 - KRTK	Other	Public	24
D4.4	Electronic newsletter III	WP4	5 - KRTK	Other	Public	36
D4.5	Project leaflet	WP4	5 - KRTK	Other	Public	6
D4.6	Establishing Web-based Resource Centre	WP4	5 - KRTK	Websites, patents filling, etc.	Public	24
D4.7	Report from Podcasts/ Social Media posts/Blogs/Public Engagement Events/ Posters/Infograms (I)	WP4	5 - KRTK	Report	Public	8
D4.8	Report from Podcasts/ Social Media posts/Blogs/Public Engagement Events/ Posters/Infograms (II)	WP4	5 - KRTK	Report	Public	16
D4.9	Report from Podcasts/ Social Media posts/Blogs/Public Engagement Events/ Posters/Infograms (III)	WP4	5 - KRTK	Report	Public	24
D4.10	Report from Podcasts/ Social Media posts/Blogs/Public Engagement Events/ Posters/Infograms (IV)	WP4	5 - KRTK	Report	Public	32

Deliverable Number	Deliverable Title	WP number	Lead beneficiary	Type	Dissemination level	Due Date (in months)
D5.4	Data Management Plan	WP5	1 - MENDELU	ORDP: Open Research Data Pilot	Public	6
D5.5	Policy Report 1	WP5	1 - MENDELU	Report	Public	26
D5.6	Policy Report 2	WP5	1 - MENDELU	Report	Public	36
D5.7	Progress Report	WP5	1 - MENDELU	Report	Public	24

The template for the management of public deliverables is provided:

1 INITIAL EXAMPLE DMP
Project Name: Example plan (Horizon 2020 DMP) - Initial Example DMP
Project Identifier: Example-DMP-ID-1
Principal Investigator / Researcher: researcher name
Project Data Contact: +420 XXX XXX XXX, name@institution.xx
Description: The example project is based on ABC and aims to investigate research questions XYZ. Data will be collected for the purpose of quantitative analysis and evidence-based conclusions drawing.
Funder: European Commission (Horizon 2020)
2 FOR EACH DATA SET SPECIFY THE FOLLOWING:
Data set reference and name
Example test dataset: examplename1.ods; ID: example1
Data set description
Example test dataset will be collected by the project research team. The collected dataset will encompass different methodological approaches. Apart from the research team, the dataset will be useful for

Standards and metadata
We intend to share our dataset in a publicly accessible disciplinary repository using descriptive metadata as required/provided by that repository. Files and folders will be versioned and structured by using a name convention consisting of the project name, dataset name and ID.
Data sharing
Example: We aim to publish our software code along with publishing our dataset in a disciplinary repository. The analysis will be performed using freely available open source software tools.
Archiving and preservation (including storage and backup)
Example: As specified by the "rules of good scientific practice" we aim to preserve data for at least ten years. Approximated end volume of the example test dataset is X GB. Associated costs for dataset preparation for archiving will be covered by the project itself. Long-term preservation will be provided and associated costs covered by a selected disciplinary repository. During the project, data will be saved daily with backup on a separate server. Backup will be checked at intervals of two weeks.

1.5 What is the expected size of the data?

The only minimal set of data will be collected to assure the significance of the expected results. All the data will be collected according to the minimisation principles, see D6.6.

A data minimization policy will be adopted that means no data which is not strictly necessary to the running of SPOT will be collected or processed. Whenever participants are requested to submit personal data, they will be informed that this data is stored and processed for the project only. Participants' data will not be mined or used for any purpose other than those explicitly and needed for running the interview.

The data to be processed in SPOT activities may include personal data (self-reported or collected), data collected from the participants (e.g. responses during the interview or collected data through questionnaires), data on behaviours related to cultural tourism, etc. All acquired data will be stored on computer media (SPOT databases, data storages, and hard disks drive), protected with state-of-the-art security measures, accessed only by selected and restricted personnel of partners.

1.6 To whom might it be useful ('data utility')?

Data generated by SPOT project, which is publicly available, will be available to all groups of people involved or interested in cultural tourism, including different types of stakeholders, businesses and research organizations, but also any other kind of visitors.

The database (as a report) is a public deliverable of the project. This report will be available by the systems the European Commission uses to provide access to these public deliverables. It is text data. Next to the report, a spreadsheet (data file in csv format) with text boxes will be made available. This data will not only be useful for people currently interested in cultural tourism, but also for future researchers that study cultural tourism.

2 SPOT DMP: FAIR data

2.1 Making data findable, including provisions for metadata

2.1.1 Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism?

The data files to be included in an open data repository will be made openly available through the ZENODO which is the open access repository of the Open Access Infrastructure for Research in Europe (<https://zenodo.org/>).

Data, which can be made openly accessible, will be submitted to the repository within 6 months after the approval of the deliverable they have contributed to.

All raw data (including transcripts of interviews, which will not be provided openly) will be retained for at least five years on the servers of the project partner responsible for gathering this raw data (this data is often in the local language of that partner, and it is better not to transfer unnecessary personal data between partners) for validation. Each partner should ensure that there will be people curating the raw data during this period. The servers for storing the raw data will have necessary security mechanisms, such as password protection, restricted access only to the project partner.

Access to this raw data will only be provided for validation. Access will only be provided based on a positive review of the partner that is keeping the data, and if such a does not exist at this partner, based on a positive review by the coordinator.

The data that can be made openly available (database; inventory of dream futures; other data) will be shared through the ZENODO. Part of the data may also be shared through other repositories (in that case, links will be provided). The best practice to make data discoverable can be through standard naming practices and use of metadata such as keywords, location and date of.

Access to data (including transcripts of interviews) for the validation of the results will be possible under strict access conditions. The responsible project partners will ensure access upon request for this purpose based on a review by the coordinator.

2.1.2 What naming conventions do you follow?

SPOT project will follow standard university naming convention [5]. It is useful to establish a best practice for file naming as part of managing both paper and electronic records. The benefits of naming conventions include finding files more easily, creating uniformity, making sorting more predictable, giving clues to the contents of files and folders without a close examination, and controlling versions. Below are guidelines we will follow:

- Create names that will allow useful sorting.
- Keep names as short as possible and make them easy to read (Windows, OS X, and Linux all limit names to 255 characters).
- Include only alphanumeric characters.
- Use camel case to distinguish words (e.g. SenateVoteRound01.docx).
- Avoid spaces, abbreviations, and most symbols except underscore “_” and hyphen “-” (hyphens should only be used in the root filename, preferably for dates).

- Format dates to enhance sorting; for proper sorting, date order should be YYYY-MM-DD (e.g. Minutes_2013-02-15.docx).
- Use the filename for version control (e.g. CollectionPolicy_rev2013-02-20.docx, Minutes_draft_2015-08-22.pdf, Minutes_final_2016-05-20.pdf).
- Consider putting the initials of the author in the filename (e.g. HowToFile2013-02-20_dh.docx).

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

A pre-agreed set of words arranged in a hierarchical order (with the option to agree to add other words as the project progresses) will be used so as to ensure a level of standardisation across data from different partners.

2.1.4 Do you provide clear version numbers?

It is not decided yet.

2.1.5 What metadata will be created?

It is not known yet.

2.2 Making data openly accessible

2.2.1 Which data produced and/or used in the project will be made openly available as the default?

During the lifecycle of the SPOT project, datasets will be stored and systematically organised in a database tailored to comply with the requirements of WP1 (for more details on the database architecture, please see D1.1 Template for the data collection). The database schema and the queryable fields will be also publicly available to the database users as a way to better understand the database itself. In addition to the project database, relevant datasets will be also stored in ZENODO, which is the open access repository of the Open Access Infrastructure for Research in Europe. Data access policy will be unrestricted since no confidentiality or IPR issues are expected regarding the datasets. All collected datasets will be disseminated without an embargo period unless linked to a green open access publication. Data objects will be deposited in ZENODO under:

- Open access to data files and metadata and data files provided over standard protocols such as HTTP and OAI-PMH.
- Use and reuse of data permitted.
- Privacy of its users protected.

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

All raw data (including transcripts of interviews, which will not be provided openly) will be retained for at least five years on the servers of the project partner responsible for gathering this raw data (this data is often in the local language of that partner, and it is better not to transfer unnecessary personal data between partners) for validation. Each partner should ensure that there will be people curating the raw data during this period. The servers for storing the raw data will have necessary security mechanisms, such as password protection, restricted access only to the project partner.

Access to this raw data will only be provided for validation. Access will only be provided based on a positive review the partner that is keeping the data, and/or based on a positive review by the coordinator.

The data that can be made openly available (database; inventory; other data) will be shared through ZENODO. Part of the data may also be shared through other repositories (in that case, links will be provided).

2.2.3 What methods or software tools are needed to access the data?

Anonymised data collected during the interviews or surveys will be accessible by available programs such as MS Office (Word, Excel). Data stored at ZENODO will be accessible through the instructions provided on its website.

Certain datasets may be accessible only using specific software, but this should be avoided as much as possible. Each case must be explained and justified in the data management plan for the particular project.

The SPOTIT software could be included to allow for results to be reproduced, however its utilization will be limited as this may interfere with commercialisation. We suppose that information regarding the functioning of the tool will not be open access.

2.2.4 Is documentation about the software needed to access the data included?

As for now, there is no particular documentation about the software needed to access the data.

It may be in certain cases.

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

This is not clear yet.

2.2.6 Where will the data and associated metadata, documentation and code be deposited?

The SPOT project database will be designed to remain operational for 5 years after project end. By the end of the project, the final dataset will be transferred to the ZENODO repository, which ensures sustainable archiving of the final research data. Items deposited in ZENODO will be retained for the lifetime of the repository, which is currently the lifetime of the host laboratory CERN and has an experimental programme defined for the at least next 20 years. Data files and metadata are backed up on a nightly basis, as well as replicated in multiple copies in the online system. All data files are stored along with an MD5 checksum of the file content. Regular checks of files against their checksums will be made.

SPOT Project Portal

The partners in the SPOT consortium decided early to set up its own project-related webpage. This webpage describes the mission and the general approach of the project and its development status. A dedicated section for downloads is used to publish reports and white papers. All documents are published using the portable document format (PDF). All downloads are enriched by using simple metadata information like the title and the type of the document. The webpage is designed and developed by the partner of the consortium KRTK. All webpage-related data will be backed

on a regular basis (once per month). All information on the SPOT website can be accessed without creating an account, though a private section will also be available linked with the private account of the project. The SPOT Project Portal will be available during the project runtime and will still be available for at least two years after the official project end.

SPOT Web link: <http://www.spotprojecth2020.eu/>

2.2.7 Have you explored appropriate arrangements with the identified repository?

Under the process.

2.2.8 If there are restrictions on use, how will access be provided?

No restrictions are expected as ZENODO is open access and publicly available platform used by several H2020 projects.

2.2.9 Is there a need for a data access committee?

There is no need for a data access committee.

2.2.10 Are there well described conditions for access?

ZENODO provides very clear instructions and manual on how to access particular data stored in there.

2.2.11 How will the identity of the person accessing the data be ascertained?

ZENODO provides very clear instructions and manual on how to access particular data stored in there, including the protection of the identity.

2.3 Making data interoperable

2.3.1 Are the data produced in the project interoperable?

Datasets, which can be made openly accessible, will be made publicly available within 6 months of European Commission's approval of the deliverable that is based on the dataset, allowing the necessary time to publish.

ZENODO (where the datasets will be deposited) ensures data quality and curation (manual curation at the time of deposition, and automated curation and checks for data integrity after the deposit). Metadata will be provided. Research data will be available for at least 5 years from the time of data deposition.

Access to data (including transcripts of interviews) for the validation of the results will be possible under strict access conditions. The responsible project partners will ensure access upon request for this purpose based on a review by the project coordinator.

2.3.2 What data and metadata vocabularies, standards or methodologies will you follow to make project data interoperable?

Regular standards and methodologies are going to be followed to make project data interoperable.

Projects will use established European and international standards and methodologies for the storage, exchange and dissemination of project data. INSPIRE (the European Directive on Infrastructure for Spatial Information) [6] compliance will be used wherever possible. The SPOT call will explicitly solicit projects that address the development of such standards.

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

Standard vocabularies will be used to the extent that they exist or will be developed in the projects.

2.3.4 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?

The projects will be required to document if they find it unavoidable to use uncommon or generate specific ontologies or vocabularies. If mapping to more commonly used ontologies will be possible the projects will be required to establish such a mapping.

2.4 Increase data re-use (through clarifying licenses)

2.4.1 How will the data be licensed to permit the widest re-use possible?

It will be specified.

2.4.2 When will the data be made available for re-use?

Datasets, which can be made openly accessible, will be made publicly available within 6 months of European Commission's approval of the deliverable that is based on the dataset, allowing the necessary time to publish.

2.4.3 Are the data produced and/or used in the project useable by third parties, in particular after the end of the project?

Yes, no limits or restrictions have been identified yet. The general rule will be that data produced in the projects shall be useable by third parties.

2.4.4 How long is it intended that the data remains re-usable?

The intention that the data remains re-usable is minimum 5 years after the project ends with possible prolongation.

2.4.5 Are data quality assurance processes described?

Data Quality Assurance Plan describes routine measures to assure data quality. It illustrates that data possess 5 key high quality attributes: – Validity – Reliability – Precision – Integrity – Timeliness.

Validity: *Face validity*-data must be true representations of the indicator of interest, and the indicator must be a valid measure of the result. *Attribution.* Changes in the indicator can be plausibly associated with USAID interventions. *Measurement Error.* Sampling and non-sampling errors.

Reliability: Data are considered reliable if the methods by which they are collected and analyzed remain stable over time. Are data collection and analysis methods documented in writing and being used to ensure the same procedures are being followed each time?

Precision: Precise data have a sufficient detail to present a fair picture of what is actually happening. – Is the margin of error reported along with the data? – Limited biases – Is the data collection method/tool being used to collect the data fine-tuned or exact enough to register the expected change?

Integrity: Data that have integrity are protected by a system that reduces the possibility of bias (either by transcription error or deliberate manipulation). Are procedures or safeguards in place to minimize data transcription errors? – Is there independence in key data collection, management, and assessment procedures? – Are mechanisms in place to prevent unauthorized changes to the data?

Timeliness: Are data available frequently enough to inform program management decisions? Are the data reported the most current practically available? Are the data reported as soon as possible after being collected?

To get successful data quality assurance, following steps should be followed:

- Adequate staff capacity, supervision and accountability
- Complete documentation of processes/ protocols readily available to collectors and processors
- Routine cross checking-mechanisms
- Clear strategy to respond to problems
- Adequate financial and logistical resources to ensure timely performance

3 SPOT DMP: Allocation of resources

3.1 What are the costs for making data FAIR in your project?

The costs of data management are part of the project costs of the Horizon 2020 project and data management is part of the regular management structure of the project. It is an integrated part of the activities of this project. The technical coordination of this project is also responsible for the technical coordination of data management.

3.2 How will these be covered?

The costs of archiving the data will stay on the project partner responsible for keeping the data. Costs related to making the data open access are eligible to the Horizon 2020 grant by which the project is financed. It is expected that these costs will be low, as the project does not need large data storage capacity.

3.3 Who will be responsible for data management in your project?

The Data Management Plan presented aims to identify the project outputs to be disseminated as well as to decide on way and means of their Open Access (if applicable). To ensure it, a dedicated time slot will be reserved at each of the project plenary meetings and, if needed, at selected consortium audio conferences. Individual responsibilities on data management in the project consortium are:

- Data Management Plan Leader (MENDELU) – to prepare and lead related discussions at the relevant project meetings and to maintain the channels for dissemination of project outcomes.
- Scientific and Technical Project Manager (MENDELU) – to identify data collected by the project and technical project outcomes eventually suitable for publication
- Dissemination leader (KRTK) – to identify publications suitable for publication in the considered repositories and maintain SPOT inputs for the Open Access

- Each individual partner – to identify own project results suitable for publication Moreover, each SPOT partner has to respect the policies set out in this DMP. Datasets have to be created, managed and stored appropriately and in line with applicable legislation. Validation and registration of datasets and metadata is the responsibility of the partner that generates the data in the WP. Metadata constitutes an underlying definition or description of the datasets, and facilitate finding and working with particular instances of data. Backing up data for sharing through Open Access repositories is the responsibility of the partner possessing the data. Quality control of these data is the responsibility of the relevant WP leader, supported by the Data Management Plan Leader. If datasets are updated, the partner that possesses the data has the responsibility to manage the different versions and to make sure that the latest version is available in the case of publicly available data. Last but not least, all partners must consult the concerned partner(s) before publishing data in the open domain that can be associated to an exploitable result.

3.4 Are the resources for long term preservation discussed?

This issue is under the process.

4 SPOT DMP: Data security

4.1 What provisions are in place for data security?

During the research project, all data will be stored on local servers maintained and automatically backed up by the partners responsible for the activities. Research organisations, working with the research data, in the project have facilities to comply with requirements to storage and access to research data. If necessary, these facilities may be used in a regional context (as defined in the DoA) by other partners who do not have these facilities.

Interview data will be gathered in a local context. Interviewers will upload audio files (of interviews) and text files of transcriptions at least once per working day to a secure data server. If in a remote rural context of a specific case study no workable safe connection to the server is available, this upload may be postponed a few days until such a connection is available or a safety copy needs to be created and stored on other location than where the original data is stored.

Only personnel working in the project has access to data files. Responsible for data access at each partner are the team leaders.

The coordinator will provide additional advice, as needed, on data management during the research project.

Detailed information regarding the security measures can be found in D6.8.

4.2 Is the data safely stored in certified repositories for long term preservation and curation?

Data will be stored in certified repository ZENODO for minimum of 5 years after the project ends.

5 SPOT DMP: Ethical aspects

5.1 Are there any ethical or legal issues that can have an impact on data sharing?

During this project, it is expected to interview human participants and therefore working with personal research data. Details of all the steps undertaken to ensure appropriate levels of data protection are outlined in the dedicated ethics section of the DoA and ethics deliverables.

During the process of transcribing the interviews, each partner will be responsible for de-identifying the human participants as much as possible. If full anonymization is not possible in the context provided, pseudonymization will be used. All the de-identification conditions and protection mechanism will be clarified in the consent form provided to all participants.

Besides, there are no other specific ethical or legal issues that can have an impact on data sharing. For details related to ethical aspects see the ethics deliverables D6.1-D6.8. or ethics chapter in the Description of the Action (DoA) on page 90 (Part B2).

Transparency

SPOT framework will ensure the collection, storage, protection, of data collected by Research organizations or individuals, who will take part in the process of the project. Any breach or leak of data to irrelevant parties (e.g. supervisors, managers) may lead to transparency issues. To that end, SPOT Project would provide the necessary feedback in order to minimize the impact of that risk or any other similar. Furthermore, the Project in collaboration with platform coordinators will inform participants and relevant authorities on the details, the scope and the purposes of the data collection process in order to get an Ethical approval consent signed by them. In addition, the SPOT Project would ensure that data will be used only within the goals and objectives of the project. The goal of the project is not to use data for other reasons and this point will be also marked on the information content to be delivered on platform users.

Behaviour

The SPOT consortium has taken into consideration the fact that some people may change their behaviour and/or their professional performance when they know that they are being interviewed. For that reason, the project's purpose and intentions have been made perfectly clear to all participants. In addition, the selection of the participants is based mainly on their high interest and willingness to participate in SPOT project, pointing out that the participation in the project does not result in more work for the involved end users. The ethical scope of the SPOT project has received significant consideration from the very beginning, and as the projects unfolds and evolves it will be one of the aspects that will guide all the procedures. The ethics of the SPOT framework will be carefully treated throughout the lifecycle of the project so that ethical risks will be appropriately addressed.

5.2 Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?

Details regarding data sharing and long term preservation are included in the questionnaire template dealing with personal data and in the informed consent. For details see deliverable D6.1.

6 SPOT DMP: Other issues

6.1 Do you make use of other national/funder/sectorial/departmental procedures for data management?

Regular procedures for data management, as set for each partner's institution will be followed.

6.2 Other SPOT project specific issues

Each partner from the consortium of 15 countries/institutions will follow also their policies for data management. The obligations set in the grant agreement and consortium agreement precede these local policies. However, many of these policies are based on the same principles of providing access to research data to validate results, data provision using the FAIR principles (Findable, Accessible, Interoperable and Reusable), and protection of personal data based on the General Data Protection Regulation of the EU.

Summary

The objective of the document is to report the steps for data management which are to be followed during the execution of SPOT project. The scope of the document is twofold: to define the detailed data management plan towards the dissemination of project outcomes and to report the detailed ethics handling and preservation activities. The current document gives a preliminary information about the data types used and generated by the project consortium partners including focus on the means of sharing data captured by SPOT framework and further specifies the methods of data storage thus providing general view over the complete data management life cycle. As this report is generated at the early stage of the project execution is considered as a living document which will be further updated during the project life time if needed.

SOURCES:

- [1] <https://libguides.jcu.edu.au/research-data-management-toolkit>
- [2] ZENODO. Available at: <http://www.zenodo.org/>
- [3] OpenAIRE H2020 Project. Available at: <https://www.openaire.eu/>
- [4] FAIR. (2013). FAIR Data Management in Horizon 2020.
- [5] File naming guidelines. Available at: <https://library.stanford.edu/spc/university-archives/managing-university-records/file-naming-guidelines>
- [6] INSPIRE (the European Directive on Infrastructure for Spatial Information) <https://inspire.ec.europa.eu/inspire-directive/2>

PERSONAL DATA PROTECTION

DESIGNATIONS

Partners / partnership: In the following text, “partner” / “partnership” designate partner institutions / organisations participating in SPOT and likely to collect data within the framework of the project, namely MENDELU (CZ), BIU (IL), IGAR (RO), IOER (DE), MTA KRTK (HU), WR (NL), TLU (EE), UNI GRAZ (AT), UNIABDN (UK), UAegean (EL), UB (ES), UNIVR (IT), UWR (PL), UL (SI), UKF (SK). The SPOT project partners carefully treat all data collected as confidential and strictly use it under the framework of SPOT project activities in compliance with the EU legal regulations and the H2020 Programme rules. All data collected and managed under the activities of the SPOT project, namely accounts, newsletter subscribers, event registrations, surveys of any kind, is in strict compliance with the new Regulation (EU) 2016/679, General Data Protection Regulation.

WHO IS CONCERNED BY THIS NOTICE?

This notice is addressed to the following public:

- All partners participating in the SPOT project;
- respondents to surveys, interviews;
- registered participants in events;
- account owners in the exchange areas of the project website;
- newsletter subscribers of any outsourced database application.

FOR WHAT PURPOSES DOES THE SPOT PROJECT STORE YOUR DATA?

Project partners collect data to perform their legal obligations as an EU co-funded project. The scope of the data collected is the minimum necessary for each purpose, avoiding as much as possible personal information. However, no personal information is collected without the knowledge and consent of the target audience. No data will be shared with third parties outside the project and the European Commission, other than the external providers or used for unintended purposes without the express consent and prior notification to the interested individuals. When personal data is collected, the purpose will be clearly expressed. The data collected within the framework of the SPOT project will be retained by the project partners and the H2020 Programme until 31/12/2027. Once the retention period has passed, the SPOT project partners, potential sub-contractors / external providers will take adequate measures to delete all personal data collected within the project or extend their use, after receiving respective consent by the data providers, as part of the post-project exploitation activities of the project.

WHAT ABOUT COOKIES AND TRACKING TECHNOLOGIES?

The SPOT web platform collects and stores information using cookies and similar tracking technologies to track web behaviours. We use this information to provide analytics of only statistical nature.

WHAT ARE YOUR LEGAL RIGHTS AS DATA SUBJECT?

Every person may directly require from an organisation holding information about them the data to be corrected (if they are wrong), completed or clarified (if they are incomplete or equivocal), or erased (if this information could not legally be collected). Anyone may oppose that information about them is used for advertising purposes or for commercial purposes; They may also oppose to information concerning them being disclosed to a third party for such purposes. The persons concerned should have the possibility of exercising their right to oppose the disclosure of their data to a third party at the moment the data is collected. The use of automatic calling machines or faxes for advertising purposes is prohibited unless the person has given their prior consent. If you believe that the processing of your personal data constitutes a violation of the legislation in force, you have the possibility to lodge a complaint with the supervisory authority concerned. (check here: https://ec.europa.eu/justice/article-29/structure/data-protection-authorities/index_en.htm)

List of the data subject rights as stated in the Chapter 3 of the GDPR:

- Transparency and modalities
- Information and access to personal data
- Rectification and erasure
- Right to object and automated individual decision-making

WHAT TYPE OF DATA IS COLLECTED AND HOW?

The project partners store data of different types and in several ways: E-mail, postal and telephone contacts; Newsletters subscription; Event registrations; Surveys; Contact forms; Data collected through the SPOT web site.