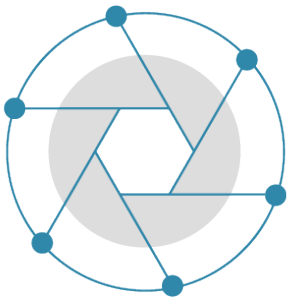
Please cite as: M. Biscotto, G. Gurtner, P. Hernández, U. Kluge, P. López, D. Valput, "Data Management Plan (D1.6)", CAMERA, Madrid, Spain, 2020.



CAMERA

D1.6

D1.6 DATA MANAGEMENT PLAN

Project number: 769606
Project acronym: CAMERA
Deliverable number: D1.6
Deliverable title: Data Management Plan

This project has received funding from the European Union's Horizon 2020 research And innovation programme under grant Agreement n°769606

H2020 - CAMERA

CAMERA – COORDINATION AND SUPPORT ACTION FOR MOBILITY IN EUROPE: RESEARCH AND ASSESSMENT

CAMERA Consortium:

1	innaxis	INNAXIS (INX)	Contact: Paula López-Catalá plc@innaxis.aero
2	UNIVERSITY OF WESTMINSTER	University of Westminster (UoW)	Contact: Andrew Cook cookaj@westminster.ac.uk
3	EUROCONTROL	EUROCONTROL (ECTL)	Contact: Peter Hullah peter.hullah@eurocontrol.int
4	Bauhaus Luftfahrt Neue Wege.	Bauhaus Luftfhart s. V. (BHL)	Contact: Annika Paul annika.paul@bauhaus-luftfhart.net
5	deepblue CONSULTING & RESEARCH	Deep Blue srl (DBL)	Contact: Micol Biscotto micol.biscotto@dblue.it

Document number: D1.6

Document title: : Data Management Plan

Version: 1.0

Status: Final

Work Package: 1

Deliverable Type: Report

Delivery date: October 2020

Responsible Unit: Innaxis

Contributors: Ulrike Kluge (BHL), Gerald Gurtner (UoW), Damir Valput (INX), Paula López (INX), Pablo Hernández (INX), Micol Biscotto (DBLUE)

Dissemination Level: CO

This project has received funding from the European Union's Horizon 2020 research And innovation programme under grant Agreement n°769606

Table of Contents

Abstract

This document provides the plan for managing the data generated and collected during and after the project in compliance with the Open Data Pilot requirements. It includes a catalogue of the datasets that will be part of the projects, the management and protection procedures and the mechanisms to make the data FAIR according to the level of confidentiality of each dataset. The DMP will be continously monitored and updated to include changes in the datasets managed or adapt to new procedures if required.

Executive Summary

The Data Management Plan has been designed as a living document to serve as a guidance on how to align with the Open Data Pilot requirements.

The first section identifies the datasets that are used for the scope of the project, mostly composed by CORDIS data from both FP7 and H2020 projects and, accordingly, public datasets. The project also includes data from semi-structured interviews conducted to ellaborate the performance framework. If additional datasets were included in the analysis to achieve the project objectives, the list will be updated. The objective is to ensure a complete list of datasets is maintained on a regular basis.

The second section outlines how to make the CAMERA data Findable, Accessible, Interoperable and Re-usable (FAIR) considering the nature of the CAMERA datasets, together with the necessary allocation of resources to achieve the FAIR goals (section 3). As the vast majority of the data is obtained from a public repository, the procedures to make these data and its metadata FAIR, are simple but clearly described for the interested audience. These section also covers the open access to CAMERA results.

The final sections define the security and ethical issues that might be associated to the project. The security section also includes a definition of the security of the data needed based on the public (CORDIS) or private (interviews) nature of the data. The Ethics section explains how CAMERA deals with personal data protection, which is only treated for events organization but never as part of the analysis.

0 Introduction

The current deliverable is the data management plan for the CAMERA project participating in the Open Research Data Pilot. CAMERA (Coordination and support Action for Mobility in Europe: Research and Assessment) is a Horizon 2020 Mobility Coordination and Support Action (CSA) to assess and report on the status of air transport mobility research and innovation in Europe and the ability of this research to accomplish the above-mentioned mobility goals.

CAMERA will quantitatively and qualitatively evaluate the research activities undertaken on air mobility over the last decade, identify the current and future gaps and innovation bottlenecks in the context of the mobility goals, and formulate appropriate recommendations. The findings and results of CAMERA will mainly support the European Commission and ACARE Working Group 1 ("Meeting Social & Market Needs"). In order to fulfill its goals, CAMERA will handle and analysze a variety of public and private data that will be managed according to its level of confidentiality. CAMERA aims at sharing with the whole community the research data collected and generated during the project always and unless otherwise required by the confidentiality conditions of some of the datasets.

Data Management Plans (DMPs) are a key element of good data management. The DMP describes the data management life cycle for all datasets considered under the framework of the CAMERA projects. The data management procedures cover data collection, storage, preparation, processing and visualization, when applicable. The data processes and FAIR principles described in this deliverable will drive the data usage at all these stages including beyond project duration. As part of making research data findable, accessible, interoperable and re-usable (FAIR), this DMP includes information on:

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

The current deliverable will be updated is necessary and appropriate to adapt to potentially changing needs during the project.

1 Data summary

The data used for the macro-level assessment of mobility research in CAMERA is acquired from the EU Open Data Portal (EU OpenData) and pertains to the CORDIS repository. Specifically, since the focus of CAMERA is on the status of the mobility research and research initiatives performed in a bit over the last decade under FP7 and H2020 frameworks, we downloaded publicly available free data from the following two repositories:

1. CORDIS - EU research projects under Horizon 2020 (2014 - 2020) (CORDIS H2020);
2. CORDIS - EU research projects under FP7 (2007 - 2013) (CORDIS FP7).

The datasets in those two repositories contain the information on the projects funded by the European Union under the seventh framework programme for research and technological development (FP7, 2007 - 2013), i.e. under the Horizon 2020 framework programme for research and innovation (H2020, 2014 - 2020). In addition to those two repository, we rely on the reference data contained in the auxiliary repository "CORDIS reference data" (CORDIS ref) which contains information on the reference lists to which CORDIS data links (e.g. FP7 or H2020 programmes or subprogrammes, topic keywords, participating countries, etc.).

1.1 FP7 data

The list of tables contained in this dataset that covers the projects funded under the framework programme FP7 is:

- **fp7projects** - Contains the public grant information for each project, including the following information: Record Control Number (RCN), project ID (grant agreement number), project acronym, project status, funding programme, topic, project title, project start date, project end date, project objective, project total cost, EC max contribution (commitment), call ID, funding scheme (type of action), coordinator, coordinator country, participants, participant countries.
- **fp7organisations** - This table lists the participating organisations, covering data fields such as: project Record Control Number (RCN), project ID, project acronym, organisation role, organisation ID, organisation name, organisation short name, organisation type, participation ended (true/false), EC contribution, organisation country.
- **fp7reports** - Covers periodic or final report summaries (or publishable summaries) from the projects; this has been included since October 2018.

The Table 1.1. gives the overview of the main characteristic of this data source.

ID	Dataset	Detail	Confidentiality Level	Role in CAMERA	Volume*
A1	fp7projects	Contains the public grant information for each project.	PU	macro-modelling and analysis	25,778 entries
A2	fp7organisations	Lists the participating organisations and their information.	PU	macro-modelling and analysis	144,113 entries
A3	fp7reports	Covers periodic or final report summaries (or publishable summaries) from the projects.	PU	macro-modelling and analysis	22,201 entries

Table 1.1: Datasets covering FP7 research initiatives and their main characteristics

1.2 H2020 data

The list of tables contained in this dataset that covers the projects funded under the framework programme H2020 is:

- **h2020projects** - Contains the public grant information for each project, including the following information: Record Control Number (RCN), project ID (grant agreement number), project acronym, project status, funding programme, topic, project title, project start date, project end date, project objective, project total cost, EC max contribution (commitment), call ID, funding scheme (type of action), coordinator, coordinator country, participants, and participant countries.
- **h2020organisations** - This table lists the participating organisations, covering data fields such as: project Record Control Number (RCN), project ID, project acronym, organisation role, organisation ID, organisation name, organisation short name, organisation type, participation ended (true/false), EC contribution, organisation country.
- **h2020reports** - Covers periodic or final report summaries (or publishable summaries) from the projects; this has been included since September 2018.
- **h2020projectDeliverables** - List of deliverables from the projects, included since May 2019.
- **h2020projectPublications** - List of publications from the projects, included since May 2019.
- **h2020-erc-pi** - List of personal investigators (PIs) involved in the H2020-funded projects. In CAMERA this dataset was extracted in November 2018.
- **h2020-msca-fellows** - List of MSCA fellows involved in the H2020-funded projects. In CAMERA this dataset was extracted in November 2018.

Table 1.2 gives the overview of the main characteristic of this data source.

ID	Dataset	Detail	Confidentiality Level	Role in CAMERA	Volume*
B1	h2020projects	Contains the public grant information for each project.	PU	macro-modelling and analysis	25,531 entries
B2	h2020organisations	Lists the participating organisations and their information.	PU	macro-modelling and analysis	131,151 entries
B3	h2020reports	Covers periodic or final report summaries (or publishable summaries) from the projects.	PU	macro-modelling and analysis	13,151 entries
B4	h2020projectDeliverables	List of deliverables from the projects.	PU	macro-modelling and analysis	70,269 entries
B5	h2020projectPublications	List of publications from the projects.	PU	macro-modelling and analysis	113,341 entries
B6	h2020-erc-pi	List of personal investigators (PIs).	PU	macro-modelling and analysis	5,128 entries
B7	h2020-msca-fellows	List of MSCA fellows.	PU	macro-modelling and analysis	4,567 entries

Table 1.2. Datasets covering H2020 research initiatives and their main characteristics

1.3 Reference data

CORDIS reference data include auxiliary information which can be very useful to further describe the datasets covered in 1.2 and 1.1. Some of the fields included in this dataset include programmes topics, funding schemes (types of action), organisation types and countries.

The list of tables contained in the reference data source is:

- cordisref-countries
- cordisref-FP7programmes
- cordisref-H2020programmes
- cordisref-H2020topics

- cordisref-organizationActivityType
- cordisref-projectFundingSchemeCategory
- cordisref-sicCode

Table 1.3 gives the overview of the main characteristic of this data source.

ID	Dataset	Detail	Confidentiality Level	Role in CAMERA	Volume*
C1	cordisref-countries	Country information.	PU	support in analytics	1,503 entries
C2	cordisref-FP7programmes	Information on FP7 programmes: titles and codes.	PU	support in analytics	6,233 entries
C3	cordisref-H2020programmes	Information on H2020 programmes: titles and codes.	PU	support in analytics	909 entries
C4	cordisref-H2020topics	List of topic codes that appear in the H2020 projects database.	PU	support in analytics	3,786 entries
C5	cordisref-organizationActivityType	List of organisation types by their code.	PU	support in analytics	5 entries
C6	cordisref-projectFundingSchemeCategory	Information on the field "funding scheme", contained in some tables in FP7 and H2020 database.	PU	support in analytics	187 entries
C7	cordisref-sicCode	Additional information on the field "subject", contained in some tables in FP7 and H2020 database.	PU	support in analytics	426 entries

Table 1.3. Reference datasets from the CORDIS database used in CAMERA

Notes:

**These datasets are being updated on continuous bases. Their characteristics as described in the tables above date correspond to the datasets updated on May 2020 and used in CAMERA as such, and they may vary as the database at the EU Open Data Portal evolves.*

1.4 Interview data

In 2018, 17 semi-structured expert interviews were conducted for supporting the development of the CAMERA Performance Framework (published in D2.1). We interviewed experts mostly from the industry side, covering all companies along the transport value chain, such as airlines, airports, public transport providers, but also mobility experts. The interview findings aimed to provide a practical, real-world view on current and future mobility challenges, which were integrated into the development of the Performance Framework and respective KPAs and KPIs. All interviews followed a pre-defined interview guideline. After consent by the interviewee, the interviews were recorded and transcribed by Bauhaus

Luftfahrt. Within conducting the interviews, we generated a) audio data of the interviews, b) textual data of the interview transcripts.

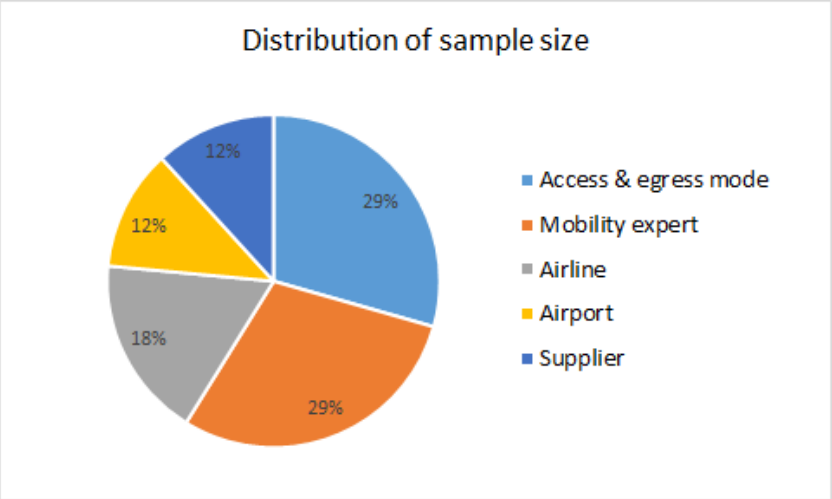


Figure.1.1: Overview of interviewed experts and distribution of sample size for the CAMERA project (N=17).

Table 1.4.gives the overview of the main characteristic of this data source.

ID	Dataset	Detail	Confidentiality Level	Role in CAMERA	Volume*
D1	CAMERA Overview Experts	Information on all contacted experts, status of feedback, interviewee, job position, contact details, data of conducted interview	R	support in interview process	60 entries
D2	Overviewsample size	Anonymized and aggregated distribution of interviewed sample size per mode (access & egress mode, mobility expert, airline, airport, supplier)	PU	support deliverables	1 pie chart
D3	Interview audio data	Interview audio data in mp-3	R	support in analytics	21 audio data
D4	Interview transcript	Interview transcripts in word	R	support in analytics	17 documents

Table 1.4

2 FAIR data

The FAIR Data Principles are a set of guiding principles in order to make data findable, accessible, interoperable and reusable (Wilkinson et al.). These principles provide guidance for scientific data management in research projects. They directly address data owners/producers and data publishers to promote maximum use of research data. The FAIR principles strengthen the development of infrastructures and services that enable a systemic change of science practices to Open Science. They also emphasize the capacity of machines (with none or minimal human intervention) to find, access, interoperate, and reuse data.

Following the European Commission's Open Research Data Pilot, CAMERA commits to adhere to FAIR principles in order to ensure that all the data used in the project and produced as a results of research performed in CAMERA is soundly managed and shared. For research data to be FAIR, the following principles need to be adhered to:

- FINDABLE = Data and their supplementary materials have sufficiently rich metadata and unique and persistent identifier.
- ACCESSIBLE = Metadata and data are understandable to humans and machines. Data is deposited in a trusted repository.
- INTEROPERABLE = Metadata use a formal, accessible, shared and broadly applicable language for knowledge representation.
- REUSABLE = Data and collections have a clear usage licenses and provide accurate information on provenance.

In the following four sections of this DMP, the data used in CAMERA are analysed against the four principles of the FAIR Data Framework.

2.1 Findable

In order for data used in and produce by a research project to be re-usable, the first step is to make them easy to find. This refers both to metadata and data, and the data should be relatively easily findable for both humans and machines. For example, machine-readable metadata are essential for automatic discovery of datasets and services.

CAMERA is relying on the data contained on the European Union Open Data Portal (EU ODP) for the macro-analysis of the research initiatives under FP7 and H2020. Concretely, the datasets are obtained from the CORDIS repository of the EU ODP and can be found at the following URIs:

- EU Research Projects under FP7 (2007 2013): https://data.europa.eu/euodp/data/dataset/cordisfp7projects
- EU Research Projects under H2020 (2014 - 2020): https://data.europa.eu/euodp/data/dataset/cordisH2020projects
- CORDIS Reference data: https://data.europa.eu/euodp/data/dataset/cordisref-data

EU ODP provides access to open data published by EU institutions and bodies. All the data that can be found via this catalogue are free to use and reuse for commercial or non-commercial purposes.

Principles of findability in the context of CAMERA

F1. (Meta)data are assigned a globally unique and persistent identifier

The data used in CAMERA's macro-analysis can be easily found in the FP7 and H2020 data catalogues by following the Uniform Resource Identifiers (URIs) provided above. A URI is a string of characters that unambiguously identifies a particular resource by following a predefined set of syntax rules, and due to that an interested user can unequivocally locate the datasets we rely on in CAMERA.

Moreover, each research project in the database of FP7 and H2020 research projects is assigned a unique and persistent (across the whole database) code (RCN) which enables unambiguous localisation and identification of each of the research projects analysed in CAMERA. Those unique project codes are kept throughout CAMERA's deliverables, and specifically, annual Mobility Reports in which we present the results of our analysis; and thus the results can easily be traced back to each data entry (research initiative analysed) in the original database in the EU Open Data Portal.

F2. Data are described with rich metadata (defined by R1)

The datasets are succinctly described on the CORDIS database at the EU Open Data Portal where they are contained, and additional metadata is available in the CORDIS Reference dataset. The reference data describe parameters from the first two datasets such as H2020 and FP7 programmes and topic codes, subprogrammes, organisations participating in analysed project activities by their type and activity the perform, etc.

Any results of CAMERA that are going to be made freely and openly available in a digital form are going to be accompanied by rich metadata that is going to instruct the user how to find, access and use the data presented.

F3. Metadata clearly and explicitly include the identifier of the data they describe

The metadata and the dataset they describe are contained in separate and easily findable files. The association between a metadata file and the dataset is made explicit through the usage of common identifiers in the metadata (reference data) and data of interest for the analysis.

CAMERA team will apply the same methodology for describing data produced as a project result.

F4. (Meta)data are registered or indexed in a searchable resource

The datasets contained in the CORDIS database can be found using the search engine provided by the EU Open Data Portal (EU OpenData).

2.2 Accessible

Once a user finds the data they want, they need to know how the data can be accessed. All the data used in CAMERA to perform the macro analysis of the status of the mobility research in Europe is openly available by the default and can be accessed at the EU Open Data Portal. Moreover, the database at the EU ODP is updated on regular (monthly) bases. Since the data from the EU ODP can be freely used in all commercial and non-commercial purposes, all the datasets used for the macro-analysis in CAMERA can be shared without any legal or contractual restrictions.

Principles of accessibility

A1. (Meta)data are retrievable by their identifier using a standardised communications protocol.

The data and metadata from the CORDIS database is accessible through the EU Open Data Portal using one of the two programming interfaces (APIs) to search for datasets: a REST API and a SPARQL endpoint. All of the portal core functionalities are available through the application programming interface (API), which encompasses most of what you can do with the web interface. The information retrieved can then be used by an external code to transform, update or reference and provide new input for further calls to the API. The protocol is open, free, and universally implementable.

Another way to access the EU ODP is the machine-readable SPARQL endpoint, which allows queries on the RDF (Resource Description Framework) descriptions of datasets.

For more information, go to https://data.europa.eu/euodp/en/developerscorner.

A2. Metadata are accessible, even when the data are no longer available

N/A

2.3 Interoperable

Interoperability concerns the needs of and approach to integrating data with other data. Moreover, it concerns the needs of data to be fed into or used in applications, storages or in some type of processing. As well as humans should be able to exchange and interpret each other's data, computers should be able to do the same: read the data without the need for specialised or ad hoc algorithms, translators, or mappings. Interoperability understands that different computer system (at least) have knowledge of the other system's data exchange formats. For this to be realisable, the use of commonly used controlled vocabularies and ontologies, as well as good data models is needed.

Principles of interoperability

I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

The data used in CAMERA is available in either a common textual (.txt) format or tabular format (.csv or .xlsc file format). All those formats are standard file formats that can be read by almost any machine and/or operating system today. Textual (.txt) format is mostly used for metadata, and thus can be easily open and read by humans as .txt files are readable and contain plain text.

CSV files contain "comma-separated values"; i.e. it is essentially a delimited text file that uses a comma to separate values. XLSX files are standard Microsoft Excel Open XML Spreadsheet files and can be easily manipulated by programmes such as Microsoft Excel or OpenOffice Calc (freely available, open source). As both file formats are textual, they can be easily readable by a number of software tools and read and expected by humans. Moreover, both CSV and XLSX files can be read and manipulated by a number of Python libraries, such as for example Pandas library for the programming language Python which specialises in reading tabular data from such files.

The results obtained in CAMERA are going to be saved and distributed in the same, widely accepted and used file formats.

I2. (Meta)data use vocabularies that follow FAIR principles

The controlled vocabulary used to describe datasets is documented and resolvable using unique and persistent identifiers. For more information, see the section 2.1 on findability.

I3. (Meta)data include qualified references to other (meta)data

A qualified reference is a cross-reference that explains its intent. For more information on cross-referencing the datasets obtained from the EU ODP, see the sections 2.1 and 2.2.

CAMERA commits to creating as many meaningful links as possible between (meta)data resources to enrich the contextual knowledge about the data, and to follow good data model principles in the creation of datasets presenting project results. Whenever that is possible, CAMERA will specify if one dataset builds on another data set, if additional datasets are needed to complete the data, or if complementary information is stored in a different dataset.

2.4 Re-usable

This constitutes the ultimate goal of FAIR framework: that the research data can be reused in subsequent projects and research initiatives, hence contributing to the goal of reproducibility in research. In order for this to be achieved, (meta)data have to be well-described so that they can be replicated in different settings. To increase the re-usability of the data, CAMERA will intend (whenever that is possible) not just provide just metadata on the data used that allows discovery of that data, but also metadata that richly describes the context under which the data (results of the project, for example) was generated. This may include the experimental protocols, algorithms, modelling approaches, and similar, used in CAMERA. In fact, a number of deliverables in CAMERA is envisioned in order to provide accurate and detailed information on the modelling and analytical methodologies used in CAMERA:

- Deliverable *2.1 Establishment of performance framework* outlines the set of Key Performance Indicators (KPIs) used in CAMERA for assessing the status of the mobility research in Europe;
- Deliverable *3.2. Mobility modelling assessment: implementation*, will outline the modelling methodology used in the analysis of the research initiatives on both the macro and micro level;
- Deliverable *4.2. Qualitative and quantitative assessment implementation*, will present the analysis methodology used in CAMERA to extract insights on the mobility research projects analysed, and to detect gaps and bottleneck in European mobility research;
- Annual *Mobility Reports* (4 planned) which, in addition to the gathered insights, present (up to a certain level of details), the analytical methods used (sometimes with more detailed explanation in Technical annexes or separate reports).

Principles of re-usability

R1. (Meta)data are richly described with a plurality of accurate and relevant attributes

The datasets collected for the macro-analysis in CAMERA are freely available in the CORDIS database at the EU Open Data Portal. Moreover, at the end of the project, CAMERA team will release the dataset version used to produce the final results of the project, and make it freely available for any commercial

or non-commercial re-use. The data were collected for the modelling and analytical purposes of CAMERA, results of which are presented in the four annual Mobility Reports.

The collected data are in raw format, and have been further processed as part of the modelling process in CAMERA.

Additionally, the CAMERA results are going to be presented and made available for its re-use (in the standard formats as described in the section 2.3). Moreover, CAMERA team will share through an open access research repository the final version of the processed dataset that is necessary and useful to reliably and easily reuse the data and reproduce the results, together with the necessary metadata and data descriptions. The access to the processed dataset can only be granted for those parts of the analysis exclusively using public datasets (i.e. CORDIS data). Finally, algorithmic, analytical and modelling approaches developed in CAMERA are going to be presented in the public deliverables mentioned above so that the re-usability of the data is further facilitated.

For those analysis including non-public datasets, open access to final results will be granted through publications and project deliverables.

Interview Data

The collected interview data did not only provide value for the CAMERA project but also for the scientific community. Next to the Performance Framework, the interview data supported two additional paper publications in the context of the CAMERA project. Firstly, the interview findings were used to support a Delphi study (Delphy study, 2020) in which mobility experts assessed possible future scenarios for door-to-door air travel in 2035. In a second technical paper, the methodological details and the conduction of the interviews were described in depth (Kluge et al., 2020). Both papers are (or are planned to be) published open access.

3 Allocation of resources

CAMERA will reserve a dedicated part of the budget to cover the costs associated to granting open access to the processed datasets and its metadata in order to be shared with the community and comply with the FAIR princiciples.

CAMERA commits to grant open access through a public repository to all its scientific publications immediately when possible, or at the latest 6 months after the original publication. CAMERA will cover its cost in case it was necessary for any particular paper.

4 Data security

As all the data used in **macro-analysis** of the mobility research initiatives in Europe is freely available at the EU Open Data Portal accordingly, no security issues are linked to this dataset.

In case that, during the execution of the project, appears a need for storing any sensitive or personal data in a secure and protected way, the consortium led by Innaxis has at its disposal a data platform **BeSt**, developed in the past several years at Innaxis and maintained by DataBeacon, that enables secure data storage in a protected data lake. DataBeacon holds a strong culture of data protection and confidentiality, and has demonstrated experience in managing and protecting private datasets, including through proprietary encryption techniques for de-identification. Data are stored in virtual private cloud (VPC) environments, each generating a data manifest that publishes the kind of available data. Relying on cryptographic methods, BeSt deploys the Smart Data Fusion (SDF) technology to hide private, identifiable data when that is required, but still allows the data to be identifiable across different datasets.

In order to achieve a long-term data preservation and curation, and thus enable interested parties to find, access and reuse the data produced in CAMERA (results, as well as certain intermediate data), CAMERA will publish (some of) the datasets it produces following the best practices for publishing research data and guidelines outlined in the scope of the FAIR framework (Gewin, V.; White, Ethan P. et al.). CAMERA will ensure that the data are deposited into a reputable repository, i.e. a qualified repository created for sharing scientific data that ensures long-term maintenance of the data, generally enforce community standards, and handle FAIR requirements such as obtaining a unique identifier, maintaining a landing page with appropriate descriptive metadata, and providing programmatic access. The repository will be formed by the end of the project and described in the final report of CAMERA, as well as in the final version of this Data Management Plan.

All the data used in the micro-analysis (with Mercury) are stored securely at the University of Westminster. The access to the data is secured via standard encryption protocols. The credentials for the access is sent via secured email protocols. Output data is stored in the same location with the same protocols.

Conversely, the gathered interview data contain personal information and sensitive data. All data on transcripts and audio recordings are stored locally in a password-protected folder at BHL in Taufkirchen in Germany. This folder is only accessible for the CAMERA-project team members and the IT employees. A data backup is conducted once per day, a shadow copy of the data twice per day. A paper-copy is also stored in a secure place in the office of the CAMERA-project team. This office is locked in case of absence and located on the Airbus campus, which is additionally protected against unauthorized access. The internal data protection officer at BHL approved all described data security measurements.

5 Ethical aspects

CAMERA data does not imply any ethical or legal issues of any type.

The EC Directive 95/46, aims at the protection of the fundamental rights and freedoms of natural persons, and in particular their right to privacy with respect to the processing of personal data, whose analysis is out of the scope of the project. According to the definition of personal data provided in Article 2(a) of EU Directive 95/46/EC, the consortium declares that CAMERA will not collect, store or manage directly or indirectly personal data of any kind apart from those collected for events organizational purposes.

As detailed in *Section 1 - Data summary*, the data used for the CAMERA project is obtained from the CORDIS public repository, including public deliverables, publications and reports from EU funded projects. This dataset complies with national and European regulations in terms of ethical and legal issues and are publicly accessible. Accordingly, no human/personal data is used for the CAMERA analysis.

While personal data remain outside of the CAMERA analysis, personal data will be collected as part of the communication activities of the project, mainly events. CAMERA will comply with applicable data protection legislation, with relevant EU Directives on data protection (Directives 95/46/EC and 2002/58/EC) and the General Data Protection Regulation (GDPR, Regulation No 2016/679, which applies from 25 May 2018). In particular, contact data will be requested from events attendees for organizational purposes only (full name, email, company and affiliation). Personal data will be used only for the purpose of the project, as explained in the consent form that the participants will have to sign when registering to the events. Personal data will be collected, processed and protected according to the General Data Protection Regulation (GDPR) (EU) 2016/679. Participants in the event will have the right to request access to and rectification or erasure of personal data or restriction of processing concerning the data or to object to processing as well as the right to data portability just sending an email to the responsible of data treatment listed hereafter. They will also have the right to lodge a complaint with a supervisory authority. Responsible for data treatment in accordance to GDPR will be Innaxis, as project coordinator.

EVENTS PARTICIPANTS

Beyond mere compliance with the legal requirements, and ensuring that the privacy and data protection rights are fully protected, a privacy policy will be prepared and distributed along with the CAMERA Consent Form and included in the online registration form. Participation in the workshop will be voluntary.

5.1 Data collection and processing:

1. **Collection of personal data:** the personal data collected and further processed are data necessary for the identification of the participants (name/surname, e-mail address, and organisation). CAMERA will ensure that the amount of personal data collected from any form of interaction with participants is kept at a minimum necessary to achieve the workshop's objectives and will incorporate anonymised data as much as possible.
2. **Processing of personal data:** CAMERA will ensure that personal data, processed on the basis of the participants' consent and freely given, can be withdrawn at any time. CAMERA will put measures in place that allow personal data to be deleted from data records. In the event that a user wants to check stored personal data, or modify, correct, or delete it, said user only needs to contact CAMERA (camera.h2020@gmail.com) and explain his or her request.

3. **Disclosure of personal data:** personal data collected and/or processed will be passed on to the local organizer (host organization, if different from the consortium partners) and coorganizers for organizational purposes. The access to these personal data is only granted on the need-to-know principle, without a prejudice to a possible transmission to the bodies in charge of a monitoring or inspection task in accordance with the EU legislation.
4. **Confidentiality:** CAMERA will keep all personal data confidential as well as data given in confidence. CAMERA organizers will disclose the confidential data only to those individuals directly working on the event organization on a need-to-know basis or when required by law. CAMERA organizers will not disclose data to any person or party unrelated to the CAMERA events. Participants' identities will not be disclosed in publications or at meetings.
5. **Data protection and data security:**
    - CAMERA will keep all personal data secure by using all necessary technical and organisational measures. It will not keep personal data for longer than necessary outside the purposes for which it was collected and will destroy or delete any personal data that is no longer required for the purposes of the DSIAW workshop.
    - All hard copy and electronic files containing personal data will be held, transferred, and processed securely. When they are no longer required for the purposes of the research, they will be destroyed in a manner that safeguards confidentiality.
6. **Re-contacting participants:** follow-up contact with a participant will be carried out only if the participant's permission has been obtained at a previous point of data collection. The only exception to this is re-contact for quality control purposes. Re-contact will match the assurances given to participants at the time that permission was gained - e.g. when re-contact was planned to occur, with what purpose and by whom.

The above detailed conditions will be clearly shared with events participants.

5.2 Workshop material and communication:

Approval to videos, pictures, recordings and dissemination of event results will be formally asked in the registration form. CAMERA will only include voluntary participants who have provided, and fully understand, the informed consent. In particular, the informed consent used in the registration form, includes information on the following subjects:

- **Usage of workshop results:** Workshop results may be published in project reports, journal articles, conference presentations, and via any other mode of scientific exchange and dissemination considered appropriate, while protecting the participants' anonymity. Data collected will be published in anonymous form.
- **Usage of photos and videos:** Photos and videos taken during the workshop activities may be used for dissemination purposes, while protecting the participants' anonymity and in the respect of the consent provided. Photos and videos may be published on CAMERA website, as well as on INNAXIS website, datascience.aero website and Vimeo Channel (https://vimeo.com/user39555379). Photos will also be published on Innaxis and CAMERA twitter and LinkedIn accounts.
- **Usage of recordings**: CAMERA may record the audio of the events in order to facilitate latter analysis.

Participants signing in the CAMERA event will have to give their explicit consent to the conditions shown above prior to take part in the event.

NEWSLETTER SUBSCRIBERS

1. **Collection of personal data:** the personal data collected and further processed are data necessary for the identification of the subscribers and sending of the alert (name/surname, e-mail address, and organisation). Individuals will need to actively opt in to receive CAMERA news/newsletter and, at that time, only a name and email address will be collected for this purpose. Sensitive personal data will not be collected.
2. **Processing of personal data:** CAMERA will ensure that personal data, processed on the basis of the subscribers' consent and freely given, can be withdrawn at any time. CAMERA will put measures in place that allow personal data to be deleted from data records. In the event that a user wants to check stored personal data, or modify, correct, or delete it, said user only need contact CAMERA (camera.h2020@gmail.com) and explain his or her request.
3. Confidentiality: CAMERA will keep all personal data confidential as well as data given in confidence.

INTERVIEW DATA

1. **Data gathering:** the participation in the interview is voluntary and experts can withdraw their consents to it at any time without justification and without disadvantages. The interviews are only recorded after explicit consent (otherwise, a memorial protocol is developed).
2. **Storage of data:** all personal data and gathered interview data is stored securely and locally at Bauhaus Luftfahrt's server. Please see the section on *4. Data Security* for further details.
3. **Disclosure of personal data:** results of the interviews are only used and displayed anonymously. The name of the company (represented by the interviewee) is disclosed in the deliverable (after consent by experts). CAMERA does no publish any personal data such as names, employment details, age, etc.
4. **Confidentiality:** CAMERA will keep all personal data and gathered interview data safe and confidential. Interview participants also received a consent form before the interviews (attached here).

6 References

- (CORDIS H2020) CORDIS - EU research projects under Horizon 2020 (2014 - 2020), Accessed at: https://data.europa.eu/euodp/data/dataset/cordisH2020projects
- (CORDIS FP7) CORDIS - EU research projects under FP7 (2007 - 2013), Accessed at: https://data.europa.eu/euodp/data/dataset/cordisfp7projects
- (CORDIS ref) CORDIS reference data, Accessed at: https://data.europa.eu/euodp/en/data/dataset/cordisref-data
- (EU OpenData) EU Open Data Portal, Access to European Union open data. Available at: https://data.europa.eu/euodp/en/home
- (OpenAire) How to find a trustworthy repository for your data, Guides for researchers. Accessed at: https://www.openaire.eu/find-trustworthy-data-repository
- (Core Certified Repositories) Core Certified Repositories, Accessed at: https://www.coretrustseal.org/why-certification/certified-repositories/
- (Zenodo) Zenodo, Accessed at: https://zenodo.org/
- (Nature Scientific Data) Recommended data repositories. Nature, Scientific data. Accessed at: https://www.nature.com/sdata/policies/repositories
- (Wilkinson et al.) Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18
- (Delphy study, 2020) U. Kluge, J. Ringbeck, S. Spinler (2020) "Door-to-door travel in 2035 - a delphi study". Technological Forecasting and Social Change 157(August), 120096. URL: http://www.sciencedirect.com/science/article/pii/S0040162520309227, doi:10.1016/j.techfore.2020.120096.
- (Kluge et al., 2020) U. Kluge, S. Spinler, J. Ringbeck (2020)“Improving the Delphi Technique: Lessons Learned from a Modified, Two-Round Delphi Survey”, (submitted for publication to *MethodsX*).
- (Gewin, 2016) Gewin, V. Data sharing: An open mind on open data. *Nature* 529, 117–119 (2016). https://doi.org/10.1038/nj7584-117a, Available at: https://www.nature.com/articles/nj7584-117a
- (White et al., 2013) White, Ethan P. et al. Nine simple ways to make it easier to (re)use your data. Vol 6 No 2 (2013): Special Issue - Data Sharing in Ecology and Evolution. Available at: https://ojs.library.queensu.ca/index.php/IEE/article/view/4608.