smashHit

Smart Dispatcher for Secure and Controlled sharing
of Distributed Personal and Industrial Data

Data Management Plan – Final Version

Public Report D8.8 Data Management Plan

This document contains information about the data the project has generated, whether and how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved.

DECEMBER 2022

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. *871477*

Foreword

The Deliverable 8.8, “Data Management Plan” contains information about the data the project has generated, whether and how it will be exploited or made accessible for verification and re-use, and how it will be curated and preserved. The purpose of the Data Management Plan is to provide an analysis of the main elements of the data management policy that will be used by the consortium with regard to all the datasets that are generated by the project. This deliverable also outlines how the research data collected, or generated, will be handled after the smashHit project, describes which methodology for data collection and generation will be followed, and whether and how data will be shared.

Table of Contents

1 Building a data management plan in the context of H2020

1.1 Purpose of the smashHit Data Management Plan

smashHit is a Horizon 2020 project participating in the Open Research Data Pilot. This pilot is part of the Open Access to Scientific Publications and Research Data programme in H2020[1]. The goal of the programme is to foster access to data generated in H2020 projects.

Open Access to Research Data refers to the right to access and reuse digital research data under the terms and conditions set out in the Grant Agreement. In this way data become re-usable, and the benefit resulting from public investment in the H2020 research programme will be increased.

The EC provides a document with guidelines[2] for project that participate in the Open Research Data Pilot. The guidelines address aspects like research data quality, sharing and security. The first Data Management Plan (DMP) was developed within the first 6 months of the project and the current one considers the evolution of the data sets within the project and foresees the sharing of the datasets after the project end.

The DMP describes the types of data that was generated or gathered during the project, the standards used, the ways how the data will be exploited and shared for verification or reuse, and how the data will be preserved.

This document has been produced following these guidelines and aims to provide a consolidated plan for smashHit partners that identifies the datasets and establishes the data management policies the project has followed and will follow after the project end. Therefore, the project partners have:

- Developed and kept an up-to-date Data Management Plan
- Deposited the project data in a research data repository
- Ensured third parties can freely access, mine, exploit, reproduce and disseminate published data (if possible in the dataset nature)
- Identified or provided the tools needed to validate the data

Among other data sets specified in the DMP (see Table 1), the industrial use cases have provided the data and metadata needed to validate the scientific publications.

1.2 Background of the smashHit DMP

The smashHit DMP is written in reference to Article 29.3 in the Model Grant Agreement called “Open access to research data” (research data management). Project participants must deposit their open data in a research data repository and take measures to make the data available to third parties. The third parties should be able to access, mine, exploit, reproduce and disseminate the data. This should also help to validate the results presented in scientific publications. In addition, Article 29.3 suggests that participants will have to provide information, via the repository, about any tools and instruments that might be made available by the project for the validation of the project outcomes.

The DMP was important for tracking all data produced during the smashHit project and after. Article 29 states that project beneficiaries do not have to ensure access to parts of research data

[1] http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf
[2] http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

if such access would undermine the achievement of project goals, for example with respect to exploitation or commercialisation of project outcomes. In such cases, the DMP must contain the reasons for not providing access.

2 Data Management Plan Overview

The DMP covers the complete research data cycle of the smashHit project as described in Figure 1. In Step 1 of the DMP, smashHit produces raw data (generated through measurements and simulations, collected through market researching, etc.). The data is then processed and analysed into more usable forms; i.e. reports, publishable documents, data tables, codes, etc. In Step 2, the data is preserved using appropriate naming rules and metadata schemes. The project's open access policy is applied to determine which datasets shall be made accessible (shared) for re-use in Step 3. The publicly accessible datasets are then re-used by the public for verification.

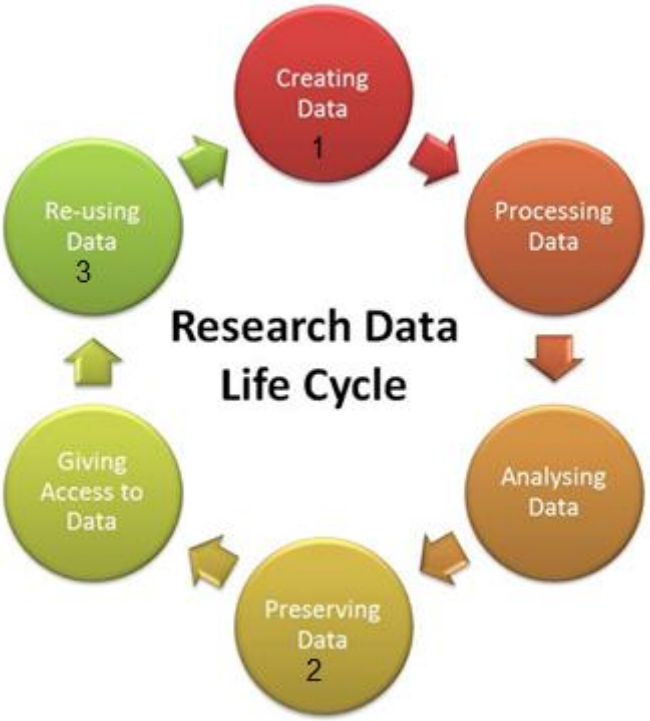

Figure 1 Research data life-cycle

This Data Management Plan presents the datasets that are produced or utilised during the project lifetime. Information on their author, curator, access and relation to the workplan are given when possible.

3 General Provisions for FAIR and IPR

A set of General Provisions have been established that provide the rules to ensure that the project datasets are **F**indable, **A**ccessible, **I**nteroperable, **R**eusable (FAIR) and that the IPR is followed. These are described in the following sections.

3.1 Making Data Findable

3.1.1 Metadata

Metadata are data on the research data themselves. They enable other researchers to find data in an online repository and is, as such, essential for the reusability of the dataset. By adding rich and detailed metadata, other researchers can better determine whether the dataset is relevant and useful for their own research. As described in the project Grant Agreement (Article 29.2), the bibliographic metadata should include all of the following:

- The terms “European Union (EU)” and “Horizon 2020”
- The name of the action, acronym and grant number
- The publication date, and length of embargo period, if applicable
- A persistent identifier

Note: All publications resulting from smashHit project must acknowledge the financial support of the EU by inclusion of the statement: “smashHit project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 871477.”

3.1.2 Persistent and unique identifier

DOI and Creative Common’s license numbers are used as persistent identifiers on open data repositories.

3.1.3 Data discovery

All datasets are referenced and indexed in the most common internet search engines. Datasets will be therefore immediately discoverable using simple web search capabilities. In addition, Open and Embargoed datasets will be discoverable on the designated open repositories for the project (see Section 3.2).

3.1.4 Data identification

All Open and Embargoed datasets have a DOI. This DOI is the same as the publication for which the data have been used for (e.g. technical publication), if the publisher requires data upload contextually to the publication.

Restricted and Closed datasets do not have a DOI, however, their metadata may have. The metadata references the dataset to which they refer.

3.1.5 Keywords

All Open and Embargoed datasets are tagged with the keywords “H2020”, “smashHit”, “871477”, and at least one additional keyword indicative of the content of the dataset.

3.2 Making Data Accessible

3.2.1 Projects datasets

The smashHit project datasets are first stored and organized in a database by the data owners (personal computer or on the organisation's secure server) and, where appropriate, on the project database. Some datasets, for which the Consortium declares no confidentiality or IPR issues, are also stored in Zenodo, the open access repository of the Open Access Infrastructure for Research in Europe (OpenAIRE). In such cases, data access policy are unrestricted. An embargo period is incurred if collected datasets are linked to green open access publication. Access levels to datasets are specified as:

- O = Open
- E = Embargo, followed by expiry date
- R = Restricted
- C = Closed

3.2.2 Accessibility of files

Most data files are accessible with common and free software:

- .png, .tif, .jpg, .raw: Any image viewer such as XnView, IrfanView, GIMP etc.
- .pdf: Acrobat reader
- .xls, .xlsx: Google sheets, LibreOffice Calc
- .doc, .docx, .odt:, txt Google docs, LibreOffice Writer
- .xml, .csv, .sql: any text editor, or database import tools/viewers.

Additional filetypes that may be agreed during the progress of the project will be described accordingly within the smashHit services which produce or use them.

3.2.3 Repositories

All datasets that are produced within the smashHit project, and for which the Consortium declares no confidentiality or IPR issues, are uploaded to Zenodo, with their respective access rights.

3.3 Making Data Interoperable

3.3.1 Data and metadata vocabularies

The Consortium looked to use standard vocabularies within the datasets that are widely accepted by the respective communities. Controlled vocabularies were used in descriptive metadata fields to support consistent, accurate, and quick indexing and retrieval of relevant data. Keywords (see section 3.1.5) and their synonyms are used for indexing and subject headings of the data and metadata.

3.3.2 Mapping of non-common ontologies

Non-common, non-obvious ontologies are explained in the metadata of the individual datasets.

3.4 Making Data Reusable

3.4.1 Data licensing

Creative Common Licensing is used to protect the ownership of the datasets. Both Share-Alike and NonCommercial-ShareAlike licenses are considered for the parts of datasets that can be made publicly available.

3.4.2 Third party access

Third parties must request access rights to Restricted and Closed data according to the procedure indicated in the respective metadata. Access rights to third parties are evaluated by the data owner on a case-by-case basis. In general, third parties have to state the purpose concerning why access rights are needed, and sign a Non-Disclosure Agreement.

3.4.3 Reasons for restriction

Restricted and closed data are raw experimental results which still need to undergo rigorous evaluation, data which might undermine the beneficiaries' IP protection and commercialisation strategies if published, or data containing sensitive information.

3.4.4 Long-term reusability

Unless otherwise indicated, all Open and Embargoed data is available for indefinite time and curated until two years after project conclusion. Restricted and Closed data is available and curated until two years after project conclusion.

3.5 Ensuring Proper Commercialisation and IPR Protection of Generated Results

For all data types, the Consortium has examined the aspects of potential conflicts against commercialisation and the IPR protection issues of the knowledge generated before deciding which information needs to be made public and when. The decision process, summarized in Figure 2, is overseen by the Exploitation Manager.

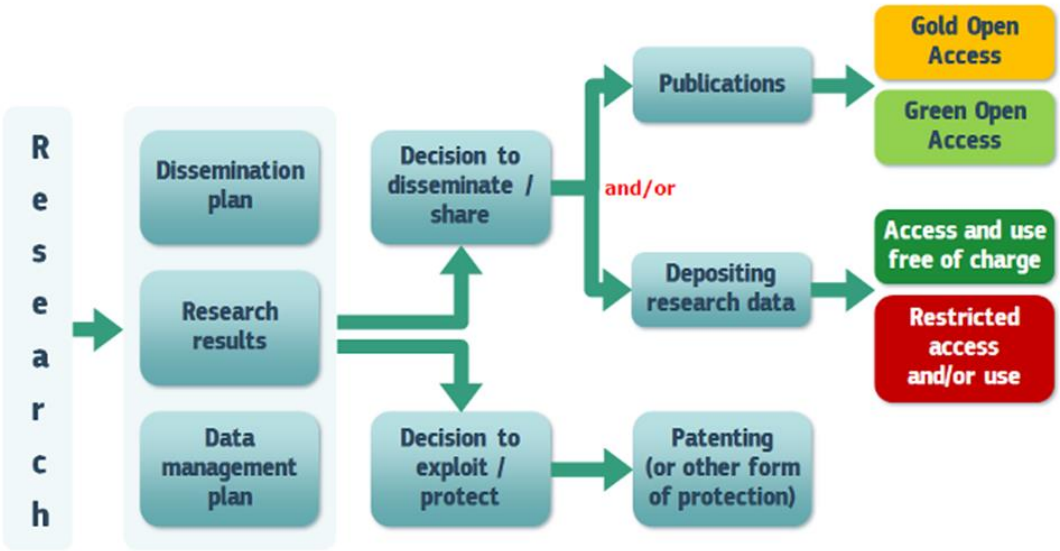

Figure 2: Open access to research data and publication decision diagram (from Guidelines to the Rules on Open Access to Scientific publications and Open Access to Research Data in Horizon 2020)

The Grant Agreement (Article 29.3) states “As an exception, the beneficiaries do not have to ensure open access to specific parts of their research data if the achievement of the action’s main objective would be jeopardized by making those specific parts of the research data openly accessible.” Such an exception applies to the smashHit project when the project findings present high innovation levels (possibility of commercialisation). In this case, the consortium has considered two forms of protection:

- to withhold the data for internal use, or
- to apply for a patent or other IP protection measure in order to commercially exploit the invention and have in return financial gain.

In the former case, appropriate IPR protection measures (e.g. Non-Disclosure Agreement) must be taken for data sharing outside the consortium. In the latter case, publications are delayed until the patent filing or other IP protection measures are completed. Otherwise, the results are made “Open Access” by depositing the research data into the designated Zenodo repository service or by publishing in journals adhering to suitable “Open Access” (‘green’ or ‘gold’). In parallel, public deliverables are stored on the following locations:

- the smashHit project website;
- the smashHit page on CORDIS website where all public deliverables submitted to the European Commission are hosted.

4 Datasets generated in the smashHit project

smashHit has produced data in a wide range of R&D activities and the expected datasets are summarized in Table 1. This list considers all datasets created or used during the project lifetime. Once generated (or collected), these datasets are stored in several formats, which are: Documents, Images, Data and Numerical Codes.

Table 1: Data generated.

No.	Dataset reference and name	Accessibility[3]	Main partners involved
1	DS1 – smashHit Core ontology	O	ATB, UIBK, LUH, UBO, ATOS, TOG
2	DS2 - Knowledge graphs with instance data representing the contracts	C[4]	UIBK
3	DS3 - Vehicle data from ECUs	C	VW
4	DS4 - Insurance customer facing consent flow	C	LexisNexis
5	DS5 - Insurance customer facing vehicle data flow	C	LexisNexis
6	DS6 - Vehicle information – static vehicle build	C	LexisNexis
7	DS7 - Vehicle information – configurable vehicle build	C	LexisNexis
8	DS8 - Vehicle information – Telematics [Mileage; ADAS sensor trigger/state; UBI; PAYD; PHYD]	C	LexisNexis
9	DS9 - Traffic information for the travellers	O	INFT
10	DS10 - Ferry arrival information	R	INFT
11	DS11 - End user feedback and observation data	R	INFT
12	DS12 - Traffic light data	C	INFT
13	DS13 - City TMC - traffic information	O	INFT
14	DS14 - Data set for traceability	O C	UBO, LUH
15	DS15 - Vehicle position data set for corner view scenario	R	VW

All datasets described in Table 1 above are described in the following sub-sections.

[3] Options for accessibility, see also section 3.2: (i) O = Open; (ii) E = Embargo, followed by expiry date; (iii) R = Restricted; (iv) C = Closed

[4] Closed data, may be made available also for third-parties upon permissions

In addition to the datasets in Table 1, the following public outputs as software programs will be made “Open Access” (to be provided free of charge for public sharing) and, as such, the software code datasets are included in the Open Research Data Pilot, and will thus be managed according to the present Data Management Plan:

- D5.4 – Public optimised smashHit Framework (INFT)
- D6.5 – Demonstrator of Services Using Integrated CPP and Insurance Data (VW)
- D7.5 – Demonstrator of Services Using Integrated Traffic, Smart City and CPP Data (FVH)

4.1 DS1 - smashHit Core Ontology

Dataset reference and name	DS1 – smashHit Core Ontology
Purpose and relation to the objectives of the project	The dataset comprise the smashHit Core schemas / ontology, that models and represent the contractual and consent data management aspects, particularly, for sensor data.
Data types	.owl
File formats	files, query end points (e.g. for SPARQL)
Reuse of existing data	Other ontologies and vocabularies are reused.
Data production methods	Manual, semi-automatic
Data utility	Servers, file systems, graph database management
Potential for reuse	Very high, in the project and beyond the project
Usage rights	Open data
Data retention	No data retention
Curator	ATB, UIBK

4.2 DS2 - Knowledge graphs with instance data representing the contracts

Dataset reference and name	DS2 – Knowledge graphs with instance data representing the consent and contracts
Purpose and relation to the objectives of the project	Instance data produced as a result of automated contracting environment, including the details on consent, contract, etc.
Data types	.rdf, .rdfs, .owl, .json
File formats	files, query end points (e.g. for SPARQL).
Reuse of existing data	Other ontologies and vocabularies will be reused.
Data production methods	Automatic, semi-automatic, manual
Data utility	Servers, file systems, graph database management
Potential for reuse	Very high, useful and usable in many applications
Data retention	Data is retained in accordance to the period of consent given by the user for their personal data storage
Usage rights	Closed data, may be made available also for third-parties upon permissions
Curator	UIBK

4.3 DS3 - Vehicle data from ECUs

Dataset reference and name	DS3 - Vehicle data from ECUs
Purpose and relation to the objectives of the project	Gathered vehicle and driving data from several electronic controls units to compute scores for vehicle insurance quotes (BC1)
Data types	.json
File formats	files
Reuse of existing data	Existing data might serve as sandbox data for test rack emulation
Data production methods	Manual, semi-automatic
Data utility	Servers, file systems
Potential for reuse	May not be reused
Data retention	Data is not stored, only transactional
Usage rights	Closed
Curator	VW

4.4 DS4 - Insurance customer facing consent flow

Dataset reference and name	DS4 - Insurance customer facing consent flow
Purpose and relation to the objectives of the project	Consent authentication by Nat. ID or Passport or if not possible by fName, lName, DoB, Current Address. Validation of consent.
Data types	Text string, captured on Insurers website and passed to LexisNexis.
File formats	Webservice between insurers and LexisNexis.
Reuse of existing data	This could be a new enquiry or an insurance renewal.
Data production methods	Data is provided by the end customer, on the insurers website* and sent through the insurer to LexisNexis (*Terms and conditions plus related consent material available within the site)
Data utility	Requires consent, to be collected by the insurer for onward validation.
Potential for reuse	No reuse – this is a one-time transactional event that requires validation.
Data retention	Data is not stored, only transactional
Usage rights	Closed data
Curator	LexisNexis

4.5 DS5 - Insurance customer facing vehicle data flow

Dataset reference and name	DS5 - Insurance customer facing vehicle data flow
Purpose and relation to the objectives of the project	Vehicle VIN or VRN relating to the insurance policy. Plus details of other 'named drivers' for the vehicle. Linking of VIN/VRN to existing vehicle records, and if no records exist vehicle data retrieved from third party data providers. Return of data would be subject to valid consent.
Data types	Alphanumeric string
File formats	Web service between insurers and LexisNexis
Reuse of existing data	This could be a new enquiry or an insurance renewal.
Data production methods	Data is obtained from a third party data provider and provided to the insurer by LexisNexis. The origin of the data may be the OEM itself (vehicle build data) or data generated by the driver (mileage & telematics).
Data utility	Legitimate interest of the insurer may apply depending on the dataset requested. Where consent is required, this needs to be collected from the end-customer.
Potential for reuse	Data may be stored and reused for insurance use cases
Data retention	Data may be retained based on the legal regulation of the insurance sector of each market (audit purposes)
Usage rights	Closed data
Curator	LexisNexis

4.6 DS6 - Vehicle information – static vehicle build

Dataset reference and name	DS6 - Vehicle information – static vehicle build
Purpose and relation to the objectives of the project	The exact vehicle specification (especially ADAS features) allows the insurer to provide a more precise and accurate insurance quotation that takes account of the reduced risk.
Data types	Vehicle build information linked to the unique VIN. Raw data retrieved from the OEM systems and converted into a normalized set of ADAS Features and a numeric ADAS rating by LexisNexis.
File formats	Various – mainly .json; .csv
Reuse of existing data	Subject to data contract, raw data or derived product data retained by LexisNexis.
Data production methods	The 'raw data' is ingested by LexisNexis and run through a process of normalisation to create a consistent output file listing the Vehicle Build features and a Rating Indicator.
Data utility	Creation of an actionable insight data product from the raw build data.
Potential for reuse	Where the data is 'static' it is possible to reuse it subject to appropriate consent.
Data retention	Default to 12-years but can be set per data source / local market.
Usage rights	Closed data
Curator	LexisNexis – proprietary product containing I.P. / Trade Secret information.

4.7 DS7 - Vehicle information – configurable vehicle build

Dataset reference and name	DS7 - Vehicle information – configurable vehicle build
Purpose and relation to the objectives of the project	The exact vehicle specification (especially ADAS features) allows the insurer to provide a more precise and accurate insurance quotation that takes account of the reduces risk. Dynamic data can confirm any over-the-air updates and remote deployment of features.
Data types	Vehicle build information linked to the unique VIN. Raw data retrieved from the OEM systems and converted into a normalized set of ADAS Features and a numeric ADAS rating by LexisNexis. Data requested from OEM source in real-time.
File formats	Various – mainly .json
Reuse of existing data	Due to the dynamic nature of this information, it must be refreshed.
Data production methods	The 'raw data' is ingested by LexisNexis and run through a process of normalisation to create a consistent output file listing the Vehicle Build features and a Rating Indicator – accurate in real-time.
Data utility	Creation of an actionable insight data product from the raw build data.
Potential for reuse	None.
Data retention	Retained in log files for audit, support and analysis only.
Usage rights	Closed data
Curator	LexisNexis – proprietary product containing I.P. / Trade Secret information.

4.8 DS8 - Vehicle information – Telematics

Dataset reference and name	DS8 - Vehicle information – Telematics [Mileage; ADAS sensor trigger/state; UBI; PAYD; PHYD]
Purpose and relation to the objectives of the project	Data points are collected from the vehicle at the start, during and at completion of a journey, and transmitted to LNRS. LNRS transform and enrich the data points to derive products relevant to their B2B markets. Customers may use the products in the pricing and management of insurance policies as well as the evaluation of extreme events such as accidents and harsh driving behaviour.
Data types	Vehicle safety and ADAS sensor data, GPS data, CAN Bus data and unique identifiers applied by the OEM to the file headers to ensure the integrity of the data. Generally transmitted to LNRS as ASCII based .json, .xml or .csv
File formats	Dependant on the OEM, primarily .json or .xml
Reuse of existing data	The data may be stored by LNRS and used to process scores that are evaluated over agreed periods of time
Data production methods	Vehicle telemetry data is transmitted from the data provider (OEM or OEM's platform provider) and ingested by LexisNexis. The data is then passed through a process of normalisation before being enriched and derived into digitised data products. The products are either requested by the customer at the point of need, or supplied to the customer as a scheduled batch. These processes are recorded for billing. The format of this data may vary and may include json, xml or csv files
Data utility	Creation of individual attributes and scores that indicate a real-time risk profile to insurers
Potential for reuse	The data owner may request to have their data re-used as part of the insurance renewal process or as part of an insurance claim
Data retention	Saleable (PII linked) data as part of a policy based product - retained for up to 7 years as part of legislative and audit procedures Anonymized data – retained indefinitely for internal R&D / evaluation purposes
Usage rights	Closed data
Curator	LexisNexis – derived proprietary products containing I.P. / Trade secret information.

4.9 DS9 - Traffic information for the travellers

Dataset reference and name	DS9 - Traffic information for the travellers
Purpose and relation to the objectives of the project	The recommended routes out from harbour for personal cars as well as to bigger commercial vehicles and buses. Also congestion, incidents, warnings or roadworks on one's route options.
Data types	.json, .csv
File formats	files
Reuse of existing data	Historical data can be used for statistics
Data production methods	Integrating existing data sources, from e.g. public transport, Helsinki City open data, Finnish Transport Infrastructure Agency open data.
Data utility	Can be used to provide real time information and statistics
Potential for reuse	City may use for statistical purposes
Data retention	Data will be stored in INFT server at least to the end of the project
Usage rights	Open data (CC BY 4.0)
Curator	INFT

4.10 DS10 - Ferry arrival information

Dataset reference and name	DS10 - Ferry arrival information
Purpose and relation to the objectives of the project	The city TMC needs to be aware when the ferry will arrive on quayside and the vehicles start to enter the street network. As exact prediction as possible is needed to prepare e.g. the traffic management processes to minimize the negative impact for the overall traffic.
Data types	.json, .csv
File formats	files
Reuse of existing data	No reuse.
Data production methods	Sensors / Software
Data utility	Useful to adjust traffic management by traffic light system SW.
Potential for reuse	No reuse.
Data retention	Data will be stored in INFT server at least to the end of the project
Usage rights	Restricted
Curator	INFT

4.11 DS11 - End user feedback and observation data

Dataset reference and name	DS11 - End user feedback and observation data
Purpose and relation to the objectives of the project	All stakeholders need more and better traffic information on the Jätkäsaari area. The information quality can be improved by taking new data sources into use: data from the cars moving in the area as well as end users can provide valuable observations and mobility data when equipped with appropriate mobile crowdsourcing applications. Residents who are familiar with the area and know where and when something exceptional probably happens would be a very valuable source of information for other stakeholders.
Data types	.json, .csv
File formats	files
Reuse of existing data	Historical data can be used for statistics
Data production methods	Mobile application, vehicle systems
Data utility	Can be used to plan and execute corrective actions in City organisations.
Potential for reuse	City and project partners may use for statistical purposes
Data retention	Data will be stored in INFT server at least to the end of the project
Usage rights	Restricted
Curator	INFT

4.12 DS12 - Traffic light data

Dataset reference and name	DS12 - Traffic light data
Purpose and relation to the objectives of the project	The data from the traffic signal system will be investigated to improve the understanding of traffic situation in the Jätkäsaari area.
Data types	.json, .csv
File formats	files
Reuse of existing data	Historical data can be used for statistics
Data production methods	Traffic light control system
Data utility	Restricted, industrial data
Potential for reuse	City may use for statistical purposes
Data retention	Raw data is not stored, only transactional. Refined data is stored in INFT server
Usage rights	Closed data
Curator	INFT

4.13 DS13 - City TMC - traffic information

Dataset reference and name	DS13 - City TMC - traffic information
Purpose and relation to the objectives of the project	The city TMC needs to inform various stakeholders if some incidents are detected or expected in the traffic. The incidents may be planned or unplanned. City TMC will use all the data sources and information that they have access to, when they are making the decision whether to publish the information or not.
Data types	.json
File formats	DATEXII
Reuse of existing data	Historical data can be used for statistics
Data production methods	Professionally operated TIC (traffic information collector) tool
Data utility	Can be used to provide real time information and statistics
Potential for reuse	City may use for statistical purposes
Data retention	Data will be stored in INFT server
Usage rights	Open data (CC BY 4.0)
Curator	INFT

4.14 DS14 - Data set for traceability

Dataset reference and name	DS14 - Data set for traceability
Purpose and relation to the objectives of the project	Users need to have more control and transparency on the data they share to incentivise them to share their data.
Data types	.json, .csv
File formats	files
Reuse of existing data	Use of open trajectory dataset from the city of Porto in Portugal (under CC BY 4.0 license[5]), and a private German trajectory dataset. These datasets are used for training and testing of data watermarking components.
Data production methods	Random generation of data for software testing.
Data utility	Can be used to provide software statistics and perform functionalities testing.
Potential for reuse	Data reusable for the same purpose described above, namely statistics and testing.
Data retention	Raw data are not stored, only metadata are processed and stored in an L3S server.
Usage rights	Open data in case of Porto taxi trajectory dataset[6] Restricted data for other, only accessible for smashHit parties involved in the development of data traceability in the context of the project.
Curator	UBO, LUH

[5] https://creativecommons.org/licenses/by/4.0/

[6] https://figshare.com/articles/dataset/Porto_taxi_trajectories/12302165/1

4.15 DS15 - Vehicle position data set for corner view scenario

Dataset reference and name	DS15 – Vehicle position data set for corner view scenario
Purpose and relation to the objectives of the project	vehicle GPS position data during the testing of the corner view (BC2)
Data types	.json
File formats	files
Reuse of existing data	Existing data might serve as sandbox data for test rack emulation
Data production methods	Manual, semi-automatic
Data utility	Servers, file systems
Potential for reuse	May not be reused
Data retention	Data is not stored, only transactional
Usage rights	Restricted
Curator	VW

About smashHit

The vision of smashHit is to overcome obstacles in the rapidly growing Data Economy which is characterized by heterogeneous technical designs and proprietary implementations, which block business opportunities due to inconsistent consent and legal rules among different data sharing platforms actors and operators. The Framework provides methods and tools, such as the smashHit platform, to assure common consent over data shared by using semantic models of consent and legal rules. The new tools include consent management and automatic contracting among the data subjects, data providers, service providers and users. It also includes traceability of use of data.

These tools are critical for enormous volumes on data streaming from the usage of mass products with cyber physical features (e.g. vehicles). These data streams offer new opportunities to build innovative services, but their combination with other personal and industrial data is subject to complex ownership and consent aspects, data streaming from these products might be considered as personal data and there are many different and partially conflicting interests involved regarding the use of data generated by the use of such products. The project has been based on the solutions developed in previous projects (AutoMat, Cross-CPP, CAMPANEO, DALICC etc.). smashHit is driven by two industrial Business Cases involving several existing industrial and personal data platforms owned by the leading data providers in three diverse sectors (automotive industry, insurance, and smart city). Demonstrators of various applications of the developed solutions are being developed and will be shown latest from end of the project on.

The smashHit project vision as well as a range of selected key problems and needs in the state-of-the-art solution for consent and contract handling in the data economy led to the starting point of the smashHit project. Driven by these needs, innovative approaches were compiled to build the smashHit architecture as key element of the project.

More information is available at www.smashHit.eu

Funded by the Horizon 2020 Framework Programme of the European Union