DISCLAIMER

This deliverable has been submitted but has not been approved by the EC yet

"Enhanced data management techniques for real time logistics planning and scheduling"

Deliverable D2.1: DNET – Report on identified and specified data sets in the form of a data management plan

Dissemination level:
☒ Public ☐ Confidential, only for members of the consortium (including the Commission Services)

Version number: 1.0

Submission deadline: 31/12/2018

www.logistar-project.eu

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

DOCUMENT INFORMATION

Authors

Name	Organisation
Nenad Gligoric	DNET
Martin Kaltenböck	SWC

Reviewers

Name	Organisation
Enrique Onieva	Deusto
Naia Merino	Deusto

Document control

Version	Date	Comment
V0.1	24/08/2018	First version of deliverable contents prepared by DunavNET
V0.2	1/11/2018	Revision of the introduction from DEUSTO (Naia Merino)
V0.3	14/11/2018	Restructure of the sections, adding dataset and reference names
V0.4	15/11/2018	Dataset description section 3.2 added
V0.5	16/11/2018	Data sharing section 3.4 added
V0.6	05/12/2018	Adapted to new template; refined & reviewed all sections; added data collection procedure (in section 1); expanded section 2 (overall dataset structure); expanded 3.2; added section 3.5. information security guidelines; added 5 ethical aspects and 6 references sections.
V0.7	11/12/2018	Ethics mentor activities included. Added list of abbreviations and acronyms. (Naia Merino)
V0.8	12/12/2018	Final editorial changes before submission

Document approval

Version	Date	Partners
1.0	12/12/2018	

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

BASIC PROJECT INFORMATION

Horizon 2020 programme

H2020 - Mobility for Growth- 5-2-2017. Innovative ICT solutions for future logistics operations

Grant Agreement No. 769142

TABLE OF CONTENTS

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

Executive Summary

This document describes the Data Management Plan (DMP) for the LOGISTAR project. The DMP provides an analysis of the main elements of the data management policy that will be used throughout the project by the project partners, with regard to all the datasets that will be generated, harvested and/or used by the project. Documentation of this plan is a precursor to the trials and pilot activities. The format of the plan follows the Horizon 2020 template [1].

In more detail this document explains and describes:

1. the LOGISTAR data identification and collection approach,
2. the LOGISTAR overall dataset structure, including an overview of identified data sources and datasets
3. the LOGISTAR overall data management plan and policy including
    a. the policies for dataset reference and naming,
    b. the dataset description (metadata scheme),
    c. relevant standards and metadata,
    d. guidelines for (secure) data sharing and
    e. information security guidelines,
4. the approach for data archiving and preservation, and finally
5. Ethical aspects in regard to data management in the LOGISTAR project.

As data management is an ongoing process along the duration of the LOGISTAR project and data management in the project is taking place in a dynamic environment, this document on hand is seen as a living document, this means that the document will be developed and maintained continuously over time.

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

1. Data Collection Procedure

This Data Management Plan (DMP) has been prepared by taking into account the template of the "Guidelines on Data Management in Horizon 2020"[1].

Elaboration of the DMP will allow LOGISTAR partners to address all issues related to management of data collected during the project as well as ethics. DMP is planned as a deliverable for M6. However, it is a living document which will be updated throughout the project based on the project progress.

The consortium will comply with the requirements of Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation) on the protection of individuals with regard to the processing of personal data and on the free movement of such data.

Type of data, storage, recruitment process, confidentiality, ownership, management of intellectual property and access: The Grant Agreement and the Consortium Agreement are to be referred to for these aspects. The procedures that will be implemented for data collection, storage, and access, sharing policies, protection, retention and destruction will be according to the requirements of the national legislation of each partner and in line with the EU standards.

The Steering Committee of the project will also ensure that EU standards are followed. Informed consent will be provided to all participants in the project trials and pilots.

All collection of sensitive data will be done with full consideration of data protection principles and industry standards, and will satisfy data protection requirements in accordance with EU and non-EU directives and national implementations thereof. Due to nature of services it is NOT likely that personal data will be captured and processed. In case that there will be sensitive/ personal data, collection and processing will be done according to the applicable data protection provisions, such as Regulation (EU) 2016/679 on the protection of individuals with regard to the processing of personal data and on the free movement of such data including article 29 working group 8/2010 opinion and Directive 2002/58 on Privacy and Electronic Communications.

For this reason, in case of personal data collection and processing, only anonymous user data will be collected and securely stored. Anonymous identification of user-provided information will be leveraged only to confirm the authenticity of users interacting with the system and to prevent malicious behaviour. No need to personally identify users through their information is envisioned nor to include sensitive data. The collected data will be treated anonymously and additionally a various set of measures will be put in place in order to protect user privacy and its data security, by embedding privacy by design principles from the early stage of the project technical start. Where needed, a prompt Privacy Impact Assessment (PIA) exercise will be performed.

[1] H2020 Programme, Guidelines on FAIR Data Management in Horizon 2020 http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

The type, quality and quantity of accessed data will be regulated, by designing and implementing adequate PIR (Privacy Information Retrieval) and PPQ (Privacy Preserving Query) mechanisms.

By referring to the proposed work plan, it is worth noticing that all such measures will be considered at all levels of the technical project development, starting from WP1 (Market research: interviews, user needs functional requirements analysis and network data collection), from WP2 where data gathering and harmonization will be done (overall data storage and data processing) up to WP3, WP4 and WP5 where data will be used to build different algorithms and services. While new required and relevant technologies will be developed as part of the project.

LOGISTAR is already aware of the following existing technological measures necessary to minimize associated privacy risks such as:

- use of secure data storage, encrypted transfer of data over the capturing channels, controlled and auditable access for different classes of data;
- obscuring/removing user identities at the source of field trial data generation to prevent direct user tracing;
- obscuring personal location data through indirect or delayed routing to prevent individual localization as much as possible and limit user tracking through correlation of depersonalized data based on its location.

The procedure for data identification and collection in LOGISTAR has been specified as follows, taking into account the specifics of the project:

1. Evaluation of the overall requirements elicitated in WP1 (Market research: interviews, user needs functional requirements analysis and network data collection)
2. Evaluation of the available requirements specification of WP7 (Use Cases and Living Labs)
3. Development of a metadata schema for LOGISTAR (to manage data monitoring for the project along a unique schema), based on DCAT (Data Catalogue Vocabulary) [2][2]
4. Data monitoring and data identification for the LOGISTAR project (along the ODI Data Spectrum[3]), means open – shared – closed data. Thereby identification of data sources and datasets and collection of respective metadata of these datasets to provide overview and search & browse mechanisms over the LOGISTAR data.
5. Setting up a data catalogue containing metadata (no data!) of the above identified and collected datasets
6. Development of data- and information security guidelines to ensure trusted and secure data sharing between partners and third parties
7. Data acquisition and harvesting by making use of the WP2 data storage layer and harvesting mechanisms. Plus continuous ingestion of new data, as well as updates and maintenance of existing data.
8. The metadata and data stores will be used for data analysis and visualisation (in WPs 3,4,5,7).

[2] https://www.w3.org/TR/vocab-dcat/
[3] https://theodi.org/about-the-odi/the-data-spectrum/

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

As pointed out above the data collection of LOGISTAR follows the ODI Data Spectrum that includes data and information as follows:

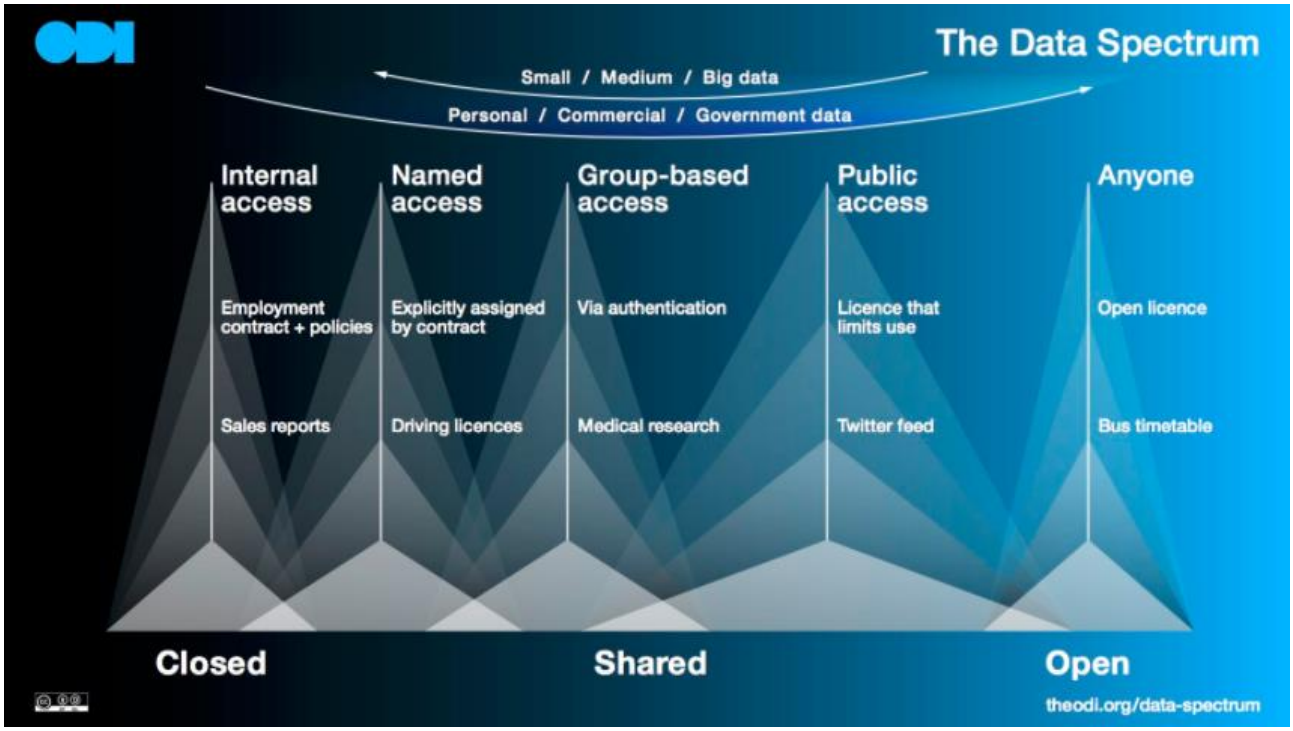

Fig.001: ODI Data Spectrum, https://theodi.org/about-the-odi/the-data-spectrum/

The overall LOGISTAR Data Management approach in LOGISTAR follows the (Linked) Data Lifecycle as follows:

Fig. 002 (Linked) Data Life Cycle

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

2. Overall dataset structure

The Data Management Plan will present in details the procedures for creating 'primary data' as well as its management. Separate datasets will be created for each stakeholder, providing the same structure, in accordance with the guide of Horizon 2020 for the Data Management Plan. Data gathered during validation of LOGISTAR and functionality as preparation for the pilots or for the purpose of scientific publications will be included in this dataset as well.

The consortium will decide and describe the specific procedures that will be used in order to ensure long-term preservation of the data sets. This field will provide information regarding the duration of the data preservation, the approximate end volume, the associated costs and the plans of the consortium to cover the costs.

2.1. Purpose of data management

The main objective of the LOGISTAR project is to allow **effective planning and optimizing of transport operations** in the supply chain by taking advantage of **horizontal collaboration**, relaying on the increasingly **real time available data** gathered from the interconnected environment.

For this, a real-time decision making support tool and a real-time visualization tool of freight transport will be developed, with the purpose of delivering information and services to the various agents involved in the supply chain, i.e. freight transport operators, their clients, industries and other stakeholders such as warehouse or infrastructure managers.

The data management activities and guidelines in LOGISTAR are built on top of this main project objectives – aligned with WP2 (Data Gathering and Harmonisation) where the major objectives are as follows:

- the identification of broad, open and IoT data for the project as well as relevant stakeholder and stakeholders partner data (closed data on i.e. goods),
- the provision of the data acquisition layer of LOGISTAR (broad & open & IoT data) and
- the provision of the metadata / semantic layer of LOGISTAR, and finally
- the provision of the overall data storage layer of LOGISTAR including 3 stores: (i) event store (ii) big data store and (iii) metadata store

Thereby the ultimate goal is to prepare a managed collection of actionable data to be used in other WPs. What includes strategies and mechanisms of secure data storage and sharing so data can be used easily and secure for analytics and visualisation et al.

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

2.2. Sources, Types and Formats of Data

The following sources of relevant data have been identified as relevant data sets for the LOGISTAR project:

- **Closed Data / shared data**
  - Data from Use Case partners (types of data see below) coming from their transport management systems (TMS)
  - 3rd party data (external use case partners) providing data from TMS and/or specific datasets about routes, prices, et al.
  - Simulated transport data from project partners
- **Open Data**
  - EU Data Portal, https://www.europeandataportal.eu/de/homepage
  - EC Open Data Portal, http://data.europa.eu/euodp/en/home
  - Lighthouse Project: Transforming Transport, https://data.transformingtransport.eu/
  - Other Transport H2020 projects in place
  - EU Intelligence Transport Systems, e.g. safe & secure truck parking, https://ec.europa.eu/transport/themes/its/road/action_plan/intelligent-truck-parking_en
  - Standards et al (e.g. GS1, ISO, W3C…)
  - Weather and traffic information data from relevant countries: UK, Italy, Europe

The following types of data have been identified as being relevant for the LOGISTAR project. This list will be maintained and expanded over time along the LOGISTAR project.

Data Type	Data
Products ordered	Product Code
	Order Number
	Quantity

Data Type	Data
Order Data	Order Number
	Facility Picking Order (this could be Facility code)
	Date & time order placed
	Date & time order ready for despatch
	Date & time required for delivery
	Delivery location of order (This could also be a code)
	Special delivery requests
	Orders to take from A to B

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

Data Type	Data
Customer data	Customer Code
	Location of customer
	Vehicle access constraints
	Opening Hours
	Typical delivery drop times
Data Type	**Data**
Vehicle data	Tractor ID
	Trailer ID
	Vehicle departed from location
	Date & time of departure
	Current location
	Date & time expected at destination
	Order IDs on vehicle
	Truck type
	Truck features (characteristics)
	Position of vehicles

Data Type	Data
Tractor data	Tractor ID

Data Type	Data
Facility data	Facility Code
(These could be supplier/factory/warehouse)	Location of facility
	Vehicle access constraints
	Opening Hours
	Typical load/unload times

Data Type	Data
Product Profile	Product code
	Ambient / chill / frozen / hazardous
	Dimensions

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

	Weight
	Stackability
	Contamination data
	Danger classes
	Moving of goods (location of goods)
	Pallet type
	Cases or quantity per pallet
	Cases or quantity per layer

Data Type	Data
Costs	Transportation costs
	calculation of costs (metrics)
	Rates negotiated

Data Type	Data
Directives	Chemical directives
	EU mobility directives
	EU transport directives
	National directives
	EU Environmental Directives

Data Type	Data
Geo Information	Regions
	Countries
	Addresses
	Routes

Data Type	Data
Standards	Article codes
	GS1 for retail
	Global location numbers
	Industry sectors (e.g. NACE codes)
	Country Codes (e.g. ISO)
	Languages (e.g. ISO 2 or 3 digits)

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

	City Codes (e.g. IATA 3 digit)
	Existing logistics taxonomies and ontologies

Data Type	Data
Other data	Vehicle filled
	Empty miles
	Types of containers
	CO2 emission / carbon footprint calculation
	miles empty
	events of interest (to be specified)
	weather data
	traffic information
	news articles (relevant)
	pollution level / location

Data Type	Data
Facilities	Addresses
	Opening hours
	Vehicle access (restrictions)

Data Type	Data
Driver data	availability of drivers
	time already worked / allowed to work
	working on the day
	driver schedules

Data Type	Data
Rail data	Train & schedules
	Capacity
	Rail operator

Data Type	Data
Services	Schedules

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

	Real time information (to be specified)
	Service Level constraints

In regards of formats of data the project has to deal with a big variety of data – means data from several sources and in several formats. The data and information will be used in unstructured format (e.g. documents), as well as in semi-structured and structured format (e.g. tabular data like CSV files). Some data will be harvested via APIs and the format such data is being received needs to be specified in the course of the technical requirements specification and architecture work in WP6.

A preliminary list of data formats is as follows:

- API data (output in several formats available)
- XML
- RDF
- CSV
- Relational DBs
- Json (LD)
- Documents (MS Word, XLS, PDFs et al)
- Etc.

In regards to re-use of existing data the LOGISTAR project has a strong focus on making use of existing data like standards (ISO, W3C, et al) in the form of for example models, taxonomies, ontologies or controlled vocabularies (like code lists), furthermore the use of open and broad data (e.g. in the area of weather data, traffic information or environmental data) wherever possible.

The size of data is still not specified at the moment of the creation of this deliverable but it can be said, that LOGISTAR data has the following 3 main attributes: (i) big data (volume) and (ii) velocity data (real time data / streaming data) and (iii) high variety (different sources, different formats) of data.

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

3. Management plans and policy

This section reflects the current status of the primary data envisioned in the project. Being in line with the EU's guidelines regarding the DMP (European Commission, 2016[45]), this document should address for each data set collected, processed and/or generated in the project the following elements:

1. Data set reference and name
2. Data set description
3. Standards and metadata
4. Data sharing
5. Archiving and preservation

To this end, the consortium develops a number of strategies that will be followed in order to address the above elements.

In this section, we provide a detailed description of these elements in order to ensure their understanding by the partners of the consortium. For each element, we also describe the strategy that will be used to address it.

Wherever possible the LOGISTAR project will follow the **EC Guidelines for Open Access**[6] as well as the **principles of FAIR data**[7], this means: **FAIR data** are data which meet standards of findability, accessibility, interoperability, and reusability.

Remark: as LOGISTAR is also working with sensitive industry data from partners and 3rd parties (e.g. data from transport management system from project partners) such data cannot be made publicly available – BUT the listed principles can be applied for data sharing between the partners and inside the consortium, where applicable and necessary)

3.1. Data set reference and name

Unique identification of datasets is ensured by following provisioned unique naming convention drafted for the purpose of the LOGISTAR project. The convention for the dataset naming is as follows:

1. Each data set name consists of 5 different parts separated with a ":", e.g. **PartnerName:EntityGroup:EnityType:VarcharId**,

[4] http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm
[5] http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm
[6] http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm
[7] https://en.wikipedia.org/wiki/FAIR_data

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

a. **PartnerName** represents the name (or the short name) of the organisation (e.g. data owner, data custodian) associated with the dataset:
    i. UDEUSTO - UNIVERSIDAD DE LA IGLESIA DE DEUSTO ENTIDAD RELIGIOSA
    ii. UCC - UNIVERSITY COLLEGE CORK - NATIONAL UNIVERSITY OF IRELAND, CORK
    iii. CSIC - AGENCIA ESTATAL CONSEJO SUPERIOR DEINVESTIGACIONES CIENTIFICAS
    iv. DNET - DRUSTVO ZA KONSALTING, RAZVOJ I IMPLEMENTACIJU INFORMACIONIH I KOMUNIKACIONIH TEHNOLOGIJA DUNAVNET DOO NOVI SAD
    v. SWC - SEMANTIC WEB COMPANY GMBH
    vi. PRESTON - PRESTON SOLUTIONS LIMITED
    vii. MDST - MDS TRANSMODAL LIMITED
    viii. SAG - SOFTWARE AG
    ix. DBH - dbh Logistics IT AG
    x. GENEGIS - GENEGIS GI SRL
    xi. AGLERS - AHLERS BELGIUM NV
    xii. ZAILOG - CONSORZIO ZAILOG
    xiii. NESTLE - NESTLE UK LTD
    xiv. PLADIS - UNITED BISCUITS (UK) LIMITED
    xv. CODOGNOTTO - CODOGNOTTO ITALIA SPA
b. **EntityGroup** – represents the category of data source, such as carrier name of the load
c. **EntityType** – represents the type of data source category
d. **VarcharId** – in systems there is a chance that context already have being assigned with the ID. In some cases, certain data context IDs in databases will be automatically iterated. For both, this suffix will be used as a final part of ID. It can be text and or numerical.

An example of one dataset name generated used above provide convention would be:

Nestle:Group1:VehicleSpeed:0001

3.2. Data set description, Standards and Metadata

Data collected, processed or generated within the project will have its description to explain dataset in more details (the metadata, MD). This data will be provided by the data owner/producer and/or other stakeholders. Information gathered during the LOGISTAR will be accompanied by context information (location, date, time) as well as publicly available information such as weather from the online local measurement stations.

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

The metadata schema (as follows) has been created by taking into account the DCAT vocabulary[8] (a W3C recommendation, that is used for e.g. metadata for open data across Europe and/or for the European Data Portal) but has been slightly adapted to the needs of LOGISTAR.

The description will provide information as given in the table below.

Title	Title of the dataset
Type of data	The type of data, e.g. driver data, vehicle data, geo information
Data provider	Provider of the data, not necessary owner
Dataset owner	Owner of the data not necessary provider
Description	Brief description of the data features and the purpose of the data
Format (Media type)	Doc, pdf, api, json, xml
License	The license and terms under data can be used
Language	ISO code of language
Metadata	Yes, no, if available provide URI reference schema
Static / Dynamic dataset	Information about the dataset if it is static (e.g. dataset as a cvs file) or dynamic data (e.g. real time data via API)
Data Type along the ODI Data Spectrum	Controlled vocabulary (CV) to describe the type of data: open, shared, closed
Classification of sensitivity of data	CV to describe of a dataset is sensible (attributes: Yes / No). If a dataset is marked as sensible: Yes then the Information & Data Security Checks and Guidelines (see below) need to be taken into account.

Potentially this LOGISTAR MD schema can be adapted and expanded over time if necessary, for example the following additional fields could be useful for metadata in LOGISTAR:

- Data purpose (what is this data used for?)
- Temporal Coverage (of Data; e.g. year 2017 or June 2016)
- Geographical Coverage of data (e.g. UK or Scandinavia)
- Language [use ISO like EN please]
- Size (this should be given approximately in MB. The goal is to know if the dataset is too large to plan how to handle it, e.g. if we need to import file couple of hundred of MB)

[8] https://www.w3.org/TR/vocab-dcat/

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

Finally, the MD schema will be used for data monitoring and identification and the LOGISTAR data catalogue (a catalogue of metadata of identified datasets for LOGISTAR).

In regard to the use of standards and the re-use of models (e.g. controlled vocabularies, taxonomies and / or ontologies) the relevant standards and models will be screened and identified along the requirements elicitation and the use case specification in the LOGISTAR project – sources are as follows:

- Standards
  - GS1
    - https://www.gs1.ch/en/home/topics/master-data-based-on-gs1
    - https://www.gs1.org/standards
    - https://www.gs1.org/standards/gdsn
  - ISO (logistic related standardsm but also terminology and metadata and information security related standards as follows)
    - ISO 28000:2007: Specification for security management systems for the supply chain
    - ISO 704, Terminology work
    - ISO 1087-1, Vocabularies
    - ISO 11179, MDR
    - ISO 20943-1 MDR (consistency)
    - ISO 25964-1 Thesauri & Interoperability
    - ISO 27001 - Information Security (Certified end 2018)
  - W3C, https://www.w3.org/
- Resource for Vocabularies: https://bartoc.org/
- EC ISA2 (Core) Vocabularies: https://ec.europa.eu/isa2/solutions/core-vocabularies_en
- EU Data Portal, https://www.europeandataportal.eu/de/homepage
- Lighthouse Project: Transforming Transport, https://data.transformingtransport.eu/

The principle approach is to make use of existing standards and wherever necessary to interlink and/or map, or to adapt such to the LOGISTAR requirements.

3.3. Secure Data sharing & Information Security Guidelines

The LOGISTAR project will define how data will be shared and more specifically the access procedures, the embargo periods, the necessary software and other tools for enabling re-use, for all datasets that will be collected, generated, or processed in the project.

In case the dataset cannot be shared, the reasons for this will be mentioned (e.g. ethical, rules of personal data, intellectual property, and commercial, privacy-related, security-related).

In addition, beneficiaries do not have to ensure open access to specific parts of the research data if the achievement of the action's main objective, as described in Annex 1 of the DoW, would be jeopardised by making those specific parts of the research data openly accessible. In this case, the data management will present the reasons for not giving access.

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

More concrete the following mechanisms has been specified:

Remark: such mechanisms will be adapted over time following the dynamic requirements of the LOGISTAR project.

- Datasets that are identified as 'sensible / closed' data in the course of the data monitoring activities of LOGISTAR will be part of a data security check and the specified LOGISTAR data / information security guidelines (see as follows)
- Between the partners a clear Non Disclosure Agreement / NDA will be executed (in addition to the Consortium Agreement) that includes mechanisms and agreements for secure data sharing between the parties
    - This NDA will be adapted to be used also for agreements for data sharing with 3rd parties in the course of the use case development in the project
- The NDA mentioned above will include – in the form of an Annex – specific Information and Data Security Guidelines that apply to LOGISTAR secure data sharing. The attributes for this are as follows:
    - Store data (into the LOGISTAR store) from each data provider separately (physically) & ensure secure data transfer
    - No unnecessary data transfer (e.g. as of federated systems)
    - Aggregation / anonymisation of data (decision on dataset by dataset basis) if necessary and useful
    - For data analytics / prediction etc data needs to be integrated, thereby the people making use of such data need to be specified and listed (remark: data sharing can increase data quality (e.g. address data))
    - Establish mechanisms of TRUST (LOGISTAR operator = data stewardship) need to be specified and implemented
    - Compliance in (i) GDPR and (ii) other data regulations, including:
        - NO export of any data outside of the EU
        - Any data breach, data loss or similar to be reported in e.g. 72 hours
        - Data will be deleted on request in specified duration (plus written confirmation)
        - Create and maintain a list of team members with access / data / organisation
        - Data processing agreements for PII (personal identifiable information) between partners to be executed
    - Any data processing, storage et al must follow Industry standards
    - Data sharing only on a 'need to know basis to data & any use is for LOGISTAR purpose only

To ensure a stable and efficient solution LOGISTAR will take into account best practises of other projects and initiatives for secure data sharing like for instance: NextTrust (http://www.nextrust-project.eu/) or iShare (https://cordis.europa.eu/project/rcn/208159_de.html).

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

4. Archiving and preservation

The data sharing procedures will be different across the datasets depending of license and will be in accordance with the Grant Agreement.

Raw data will be converted to non-talking identifiers (coded identification of the vehicle) with the use of one-way thickening algorithms using random values (different for each tracked data set – salt cryptography) while the original data will be discarded and the coded identification of the vehicle will be stored for maximum of 24 hours for the specific needs of the system.

Appropriate technical and organizational measures will be taken against unauthorised or unlawful processing of personal data in order to ensure that the individual cannot be identified from the captured data and furthermore we will ensure that data that could eventually lead to subsequent determination of the individual's path (where, when, how fast) will be stored for maximum of 24 hours. This way the possibility of identification of Individual is sufficiently minimalized

The system will aggregate the data collected in short intervals which will ensure the anonymity of the data.

5. Ethical aspects

As in LOGISTAR sensible data will be harvested, stored and processed as well as potentially also personal data will be processed the project has established a separate workpackage (WP10 - Ethics requirements) to tackle ethical issues.

The 'ethics requirements' that the project must comply with are included as deliverables in this work package, see as follows:

D10.1: NEC - Requirement No. 1 [7] (Month 1)
The applicants must ensure that the research conducted outside the EU is legal in at least one EU Member State.

D10.2: H - Requirement No. 2 [8] (Month 1)
The informed consent procedures that will be implemented for the participation of humans must be submitted as a deliverable. Templates of the informed consent/assent forms and information sheets (in language and terms intelligible to the participants) must be kept on file.

D10.3: POPD - Requirement No. 3 [9] (Month 1)
Detailed information on the procedures for data collection, storage, protection, retention, and destruction, and confirmation that they comply with national and EU legislation must be included in the Data Management Plan.
However, relevant information that pertain to the interviewing/surveying activities performed before the delivery of the Data Management Plan in M6 must also be provided by the start of these activities. In case personal data are transferred from/to a non-EU country or international organisation, confirmation that this complies with national and EU legislation, together with the necessary authorisations, must be kept on file. Detailed information on the informed consent procedures in regard to the collection, storage, and protection of personal data must be submitted as a deliverable. Templates of the informed consent forms and information sheets (in language and terms intelligible to the participants) must be kept on file. In case of further processing of previously collected personal data, relevant authorisations must be kept on file.

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

D10.4: GEN - Requirement No. 4 [10] (Month 6)
An ethics mentor must be appointed to advise the project participants on ethics issues relevant to the protection of personal data. A report on the activities of the ethics mentor must be submitted with the Data Management Plan.

5.1. Activities of the ethics mentor

The Ethics mentor appointed for LOGISTAR project is Dr Pedro Manuel Sasia who is lecturer and researcher at the University of Deusto.

During these 6 first months of the project (from June 2018 to November 2018), the activities of the ethics mentor have been focused on stablishing the adequate procedures to fulfil the ethical requirements that are relevant to the project, namely:

- Research out of Europe
  Definition of Procedures to be followed for analysing the collection and processing of personal data obtained or handled by LOGISTAR project in Serbia.
  Data transfer agreement to comply with GDPR requirements.
  (Detailed in Deliverable 10.1 [7])

- Informed consent
  Definition of Procedures implemented for the participation of humans in LOGISTAR's research activities both in interviewing activities and for the testing and validation of the system.
  Informed consent templates
  (Detailed in Deliverable 10.2 [8])

- Data management
  Definition of procedures in relation with data collection, storage, protection, retention and destruction in order to comply with the applicable national and EU legislation. Particularly, data handling procedures to be implemented related to interviewing activities that have taken place at the beginning of the project
  (Detailed in Deliverable 10.3 [9])

The ethics mentor has been in close contact with coordinator via direct mail, phone and regular meeting and has had access to all the information about the activities of the project that could imply ethically relevant aspects via email and accessing the common repository of LOGISTAR.

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

6. Conclusions

The LOGISTAR project makes use of data along the whole ODI Data Spectrum, means closed – shared – open data with main attributes: volume, velocity and variety. Data comes from different sources like open data sources, consortium members and also 3rd parties in the course of the use case realisation. Parts of the data are sensible data and potentially even personal data and thereby secure data management / sharing is an important issue to be tackled by the project. This will be taken into account from a technical as well as from an organisational viewpoint!

This Data Management Plan on hand is created as a living document that is maintained over time following the dynamic requirements of the LOGISTAR project and it acts as a guideline for the whole consortium in regards of any data management in the project.

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

List of abbreviations and acronyms

DMP: Data Management Plan

WP: Work Package

H2020: Horizon 2020 Programme

PIA: Privacy Impact Assessment

PIR: Privacy Information Retrieval

PPQ: Privacy Preserving Query

ODI: Open Data Institute

IoT: Internet of Things

TMS: Transport Management System

API: Application Programming Interface

XML: Extensible Markup Language

RDF: Resource Description Framework

CSV: Comma-separated values

DB: Database

Json: JavaScript Object Notation

LD: Linked Data

MS Word: Microsoft Word

XLS: Microsoft Excel

PDF: Portable Document Format

ISO: International Organization for Standardization

W3C: World Wide Web Consortium

MD: Metadata

DCAT: Data Catalog Vocabulary

NDA: Non Disclosure Agreement

GDPR: General Data Protection Regulation

PII: Personal Identifiable Information

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.

References

The following references are given in the document on hand:

[1] H2020 Programme, Guidelines on FAIR Data Management in Horizon 2020 http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[2] DCAT Vocabulary, https://www.w3.org/TR/vocab-dcat/

[3] ODI Data Spectrum, https://theodi.org/about-the-odi/the-data-spectrum/

[4] EC Guidelines for Data Management Plans, http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm

[5] EC Guidelines for Open Access, http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm

[6] FAIR Principles for data, https://en.wikipedia.org/wiki/FAIR_data

[7] LOGISTAR Consortium, D10.1 – NEC – Requirement No 1

[8] LOGISTAR Consortium, D10.2 – H – Requirement No 2

[9] LOGISTAR Consortium, D10.3 – POPD – Requirement No 3

[10] LOGISTAR Consortium, D10.4 – GEN – Requirement No 4

LOGISTAR project has received funding from European Union's Horizon 2020 research and innovation programme under grant agreement no. 769142.