Deliverable D 1.3
Data Management Plan

Project acronym:	LOCATE
Starting date:	01/11/2019
Duration (in months):	24
Call (part) identifier:	S2R-OC-IP5-01-2019
Grant agreement no:	881805
Due date of deliverable:	Month 02
Actual submission date:	10-03-2020
Responsible/Author:	Magno Santos - EVOLEO
Dissemination level:	PU
Status:	Final

Reviewed: yes

Document history		
Revision	*Date*	*Description*
0.1	10/01/2020	Draft version
0.2	28/02/2020	Revision by partners
0.3	10/03/2020	Final version

Report contributors		
Name	Beneficiary Short Name	Details of contribution
Magno Santos	EVOLEO	First Draft of Deliverable and final revision
Virgínia Infante	IST	Deliverable revision
Adam Bevan	HUD	Deliverable review
Loïc ANCIAN	VIB	Deliverable review

Table of Contents

1. Executive Summary

The overall objective of the LOCATE project is to develop a framework to replace, as much as realistically possible, the preventive scheduled maintenance of mechanical parts of the locomotive bogie by predictive maintenance.

The LOCATE project intends to achieve this overall objective through:

1. Development of optimized Condition-based Maintenance (CBM) techniques, locating and evaluating the impact of such approaches on maintenance planning and railway operations;
2. Developing intelligent tools integrated for the operation of the overall maintenance policy: supporting localization of faulty components, maintenance scheduling and integrating maintenance operations tasks into daily services;
3. Testing and validating in practice an open architecture able to carry asset management data to the operator beyond the locomotive bogie, locating events and defects on freight wagons, track condition, etc.;
4. Developing a minimal digital twin for the bogie system, based on vehicle dynamic simulations and post-processing, while considering the local requirements of the system;
5. Applying a cost-effective and reliability-based sensor design to locate defects and monitor critical components of the locomotive bogie.

LOCATE will address these challenges through a methodical diagnosis and specification process, based on specific use cases. The studies and assessments will be performed considering the critical components to be monitored and the current maintenance rules and procedures.

Based on this studies and assessments, LOCATE intends to gather information to assess the current condition of the critical subsystem of the locomotive boogie and the critical components failure modes which have the most impact on the overall bogie reliability and maintenance costs.

The data gathered could have three main origins:

- Utilising results from other projects , e.g. Innowag project
- LOCATE partners
- Advisory Board members, namely: FR8RAIL III and DB Cargo

The data used for these analyses have different types such as:

- measurements from previous interventions,
- signals from sensors,
- historical data of maintenance,
- data available in literature and databases

There will be three main formats of the data:

- content distributed by data bases
- semi-structured in CSV files
- unstructured text files: PDF and word

This data and information will be key for WP3 to WP5 work packages in the frame of LOCATE project structure as for possible scientific research articles support.

The LOCATE participants will submit papers in international conferences and events, which will be identified and reported on a regular basis to the Consortium partners, in order to properly organise participation in these events. Also, is expected to have publications in sectoral, education and scientific journals, the press, etc.

Regarding the scientific publications generated within the LOCATE project will be published as open access, in two possible modalities:

- "green", where the publication is self-archived after 6 months of the publishing date in a website that is widely available to the scientific community and public in general (free of charge);
- "gold", where the open access fee will be paid and the publication will be immediately accessible to the public/scientific community.

The modality used for each publication is decided by the LOCATE consortium along with the decision to create the respective scientific publication.

2. Abbreviations and acronyms

Abbreviation / Acronyms	Description
CBM	Condition-based Maintenance
DMP	Data Management Plan
ORD	Open Research Data Pilot
PA	Product Assurance
SVN	Versioning control software
TD	Technology Demonstrator
WA	Work Action
WP	Work Package

3. Background

The present document constitutes the Deliverable D1.3 "Data Management Plan" as part of the WP1 – Project Management.

It does not contribute any TD/WA.

4. Objective/Aim

The deliverable D1.3 – "Data Management Plan" describes the accessibility and exploitation of the data, detailing the following information:

- Data repository structure
- Data submission process
- Data accessibility
- Data preservation
- Data security and interoperability

Additionally, the information regarding the scientific publications access are described and the LOCATE management principles and tools are highlighted.

5. Data catalogue

The data catalogue is composed by a set of tools responsible for the data submission, archive and preservation and for dissemination and access.

The data catalogue is managed by an administrator that will guarantee that the data submitted will follow the policies defined in this document in terms of data representation, description and quality.

For each data object (file or package of files) submission, there will be a form to capture the data object representation information and preservation description, additionality will be defined the access rights to this data object.

Once approved by the administrator for submission, the package composed by the data object and representation information along with preservation description will be committed to a repository through a versioning control software (SVN).

The Preservation Description Information is divided into four types of preserving information called Provenance, Context, Reference, and Fixity.

- Provenance describes the source of the content Information, who has had custody of it since its origination, and its history (including processing history);
- Context describes how the content Information relates to other information outside the information package. For example, it would describe why the content Information was produced, and it may include a description of how it relates to another content information object that is available;
- Reference provides one or more identifiers, or systems of identifiers, by which the content information may be uniquely identified or a set of attributes that distinguish one instance of content Information from another. Also includes descriptions of how data and files are named, physically structured, and stored as well as details about the experiments, analytical methods, and research context;
- Fixity provides a wrapper, or protective shield, that protects the Content Information from undocumented alteration. For example, it may involve a check sum over the content information.

The descriptive information is that information which is used to discover which package has the content information of interest. Depending on the setting, this may be no more than a descriptive title of the data object, or it may be a full set of attributes capable to be searchable.

A data quality policy for the data object is also included in the data catalogue to be able to trace and analyse the data objects when extracted and assess the intrinsic data quality.

The record of this information is done in a central excel file with dedicated fields also under versioning control.

The structure outline of the repository is as the depicted on Figure 1.

Figure 15-1 –LOCATE repo structure

At the root, the LOCATE repo.xlsx is responsible to the data catalogue in which will have described all the data packages and respective information. This file has the capability to guarantee and ensure that there are unique identifiers for each data object. Each field of the file is obligatory to fill with context information for the expected information for each field, and managed by the administrator. Also provides the search capability.

Evolutions and changes on the archived data is possible and traced via the versioning software where shall be recorded the changes description and respective reasoning for the change, also manged by the repo administrator.

The LOCATE project main effort in the data repository preservation is to document and make easily available the descriptive information for the data objects, keeping the original format and representation.

For each data object, the data access is defined for internal use in the LOCATE project or external parties and clearly discriminated the entities with access grants.

6. Data access

The data catalogue will be available for the LOCATE partners. The access of the information, as for the submission, will be through the repository administrator. The interested party will perform an access request to the administrator which will assess if the access could be directly granted and in the affirmative case will extract the data objects from repository and make available via a secure link to the data extract.

In case of not foreseen available access, the re-assessment of the data access will be performed at LOCATE project frame view and decided if the access denial will be kept of released.

The data access will return the data objects extract, along with the preservation and representation information, including the data changing log on the repository versioning software. The data objects will be available through dedicated secure FTPs, and for the project deliverables and documentation on a shared OneDrive folder. This OneDrive folder is a mirror of the checkout of the latest committed version to the SVN versioning tool.

7. Data interoperability

The data objects in the LOCATE project, intends to be kept untouchable as much as possible, and keep the original content and representation. The data objects extracts made available for access will be contained with preservation and representation information in self-contained formats to be able to trace the vocabulary and methodologies used to better facilitate the interoperability with other data usages.

8. Data security

The LOCATE data repository is based on a distributed versioning software running on a duplicated server storage to guarantee and ensure correct versioning and data recovery in case of failure of one storage site.

The access and transfer of the data objects will be granted through dedicated secure FTPs with protected user access.

9. LOCATE management principles & communication methods

1.1. Cooperation tool

All working documents are shared by Office 365, with an edition link, so that everyone involved can actively contribute.

1.2. LOCATE e-mail convention

The email will be the most common way of exchange between all consortium members.
For a better organisation of the internal communication, the following email was created:
locate@locate-project.eu.

1.3. LOCATE meeting convention

Discussions and brainstorming sessions in the consortium meetings will be essential to incorporate in LOCATE potential new research and technical developments, as well as to update, if necessary, the needs identified by the end user's partners.

There will be regular LOCATE Consortium meetings (expected every 3 months) which will involve all partners and enable them to get together to receive a briefing on the project's status and coordinate the project's management and technical work.

The foreseen regular LOCATE project meetings include:
- project work progress meeting every two weeks (conference call)
- work progress face to face meeting every three months

1.4. Teleconference/videoconference facilities

The regular LOCATE project work meetings (every two weeks) are conference calls.
LOCATE Consortium meetings are made through the Skype for Business software.
- Skype for Business - is reliable and versatile, with integrated audio, video, and content sharing solution.

1.5.   Confidentiality

All consortium shall consider the issues of confidentiality and the obligations of the Grant and Consortium Agreements when engaging with third parties. The requirements of the Article 36 (Grant Agreement) shall also be taken in account.

All data uploaded to the cooperation tool for the LOCATE remains confidential with access restrictions applied.

Confidential Information shall include, without limitation, technical information, marketing and business plans, databases, specifications, formulations, tooling, prototypes, sketches, models, drawings, specifications, procurement requirements, engineering information, samples, computer software (source and object codes), forecasts, identity of or details about actual or potential customers or projects, techniques, inventions, discoveries, know-how and trade secrets.

All information designated as "confidential" (either orally or in writing) shall be clearly indicated as such on both paper and electronic document copies.

1.6.   Word processing

Dedicated templates will be used for uniform internal and/or external presentations of the project. The office 365 (support by EVOLEO) should be used to elaborate the deliverables.

Documents will be edited in "Track changes" mode so that the leader of the deliverable can clearly identify updates performed on the document by other partners.

1.7.   Document naming conventions

All the documentation developed under the project, will be referenced according to the following table:

Document Control			
LOCATE	D n.m	Name	Ve.r
Project Name	Deliverable ID	Deliverable Name	V edition.revision
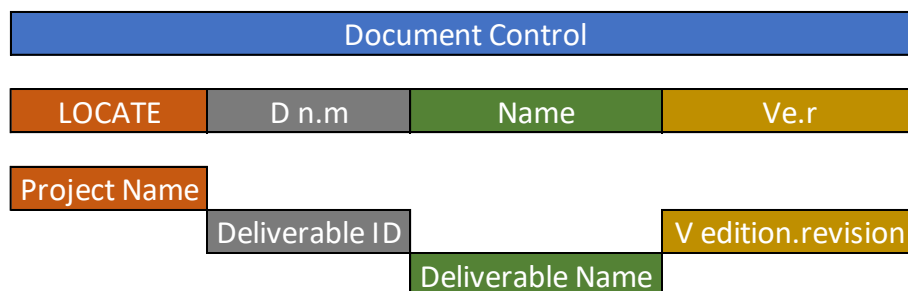
All documents are present in a list of documents and controlled by Product Assurance (PA) manager. This list allows the control and management of the documents associated with the project.

All the revisions are clearly identified within the document.

10. Roles & responsibilities

Table 1 shows the partner Contribution to LOCATE Project.

Table 1 10-1 - Partner Contribution to LOCATE Project.

Partner	Role Within LOCATE
EVOLEO	EVOLEO is the project coordinator and the WP leader for WP1 (Project Management) and WP2 (Requirements and Specifications). Besides being the Project Coordinator, it will focus on the development and implementation of critical electronic components, particularly for data collection and communication, as well as for data processing. It will also contribute to the development of a functional Software to showcase the project results.
IST	IST will lead WP4 (Reference Behaviour) and will participate in the other work packages (WP1, WP2, WP3, WP5, WP6 and WP7). Since IST is an Engineering University, it will pay special attention also to the dissemination part of the project (WP7).
HUD	HUD will lead WP5 on 'Operational Behaviour' and will support the overall delivery and management of the other work packages.
FGC	FGC is the end-user of the product developed within the project. Regarding the development FGC contributes with his expertise to define the requirements and facilitates the infrastructure required for the demos and pilots.
UIC	It assists FGC in the modification of the schemes and the organization of the maintenance of the locomotives to which the project relates by working on the operational use of the tools developed by the other members of this project. At the end of the project, the UIC guaranties the dissemination of the conclusions and the good practices defined on this occasion. Other railway companies of all sizes need to benefit from this experience.
Vibratec	It will lead the design of the measurement chain necessary for monitoring the selected components: sensor, acquisition hardware, communication to on-board unit, etc. It will implement the integration, test and verification of the demonstrator. It will support the construction of the required computer models.

11. Project planning

Each work package leader is responsible for each deliverable and responsible for ensuring that the deliverable meets the quality assurance requirements.

Table 2 11-1 – Project Deliverables planning

ID	Deliverable name	WP	Lead Part.	Type	Dissem. level	Delivery date
D1.1	Gender Strategy Plan	1	EVO	R	PU	M2
D1.2	Quality Assurance Plan	1	EVO	R	PU	M3
D1.3	Data Management Plan	1	EVO	R	PU	M3
D2.1	Use Cases Description	2	UIC	R	PU	M3
D2.2	Report on Standard and Regulations	2	UIC	R	PU	M3
D2.3	FMECA Analysis	2	EVO	R	PU	M5
D2.4	Requirements and Architecture Specification	2	VIB	R	PU	M6
D3.1	Available technologies assessment report	3	VIB	R	PU	M6
D3.2	Impact on safety levels assessment report	3	UIC	R	PU	M15
D3.3	List of Selected Sensors and Devices	3	VIB	R	CO	M12
D3.4	LOCATE OBU Specification	3	EVO	R	CO	M18
D3.5	Datasets Specification	3	IST	R	CO	M21
D3.6	Alert Limits Specification	3	EVO	R	CO	M21
D4.1	Available Models Assessment Report	4	IST	R	PU	M6
D4.2	Computational Models Specification	4	EVO	R	CO	M9
D4.3	Simulation and Post Processing Results Report	4	IST	R	PU	M21
D4.4	Behavior Prediction Report	4	VIB	R	PU	M24
D5.1	Operational Constraints Identification Report	5	UIC	R	PU	M6
D5.2	Monitoring and Thresholds Rules Specification	5	UIC	R	PU	M8
D5.3	Scheduling Flowchart	5	IST	R	CO	M12
D5.1	Operational Constraints Identification Report	5	UIC	R	PU	M6
D5.4	Assessment report of unplanned events on planned maintenance	5	FGC	R	PU	M24
D5.5	LOCATE Software user manual	5	EVO	R	CO	M24
D6.1	System Integration Report	6	FGC	R	CO	M21
D6.2	RAMS Analysis Report	6	EVO	R	CO	M21
D6.3	Tests Report	6	VIB	R	CO	M22
D6.4	Predictive Maintenance Program Implementation and Results Report	6.5	UIC	R	PU	M24
D7.1	Dissemination and Communication Plan	7	UIC	R	PU	M4
D7.2	Report on dissemination activities	7	UIC	R	PU	M24
D7.3	Exploitation Plan	7	EVO	R	PU	M24
D7.4	Recommendations Brochure	7	UIC	R	PU	M24

12.   Reporting

The WP leaders, collects, compiles and submits consolidated technical and financial details in the form of 'Periodic Reports', every 12 months.

A repository (extranet) was created to share documentation, among partners. Each partner user will have a unique ID and Password to ensure confidentiality of the deliverables. The extranet link is the following: http://extranet.uic.org/

13.   Ethics

The LOCATE consortium partners will conform to Horizon 2020 ethical guidelines, including "Data protection and privacy ethics guidelines" and the "Guidance for Applicants on Informed Consent". The right to privacy and personal data protection will be assured and these regulations will be in line with data protection rules defined on Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Therefore, LOCATE project members will make sure that data collected shall be limited to the data necessary to perform the research and development activities of the project.

The project's planned activities do not have any significant negative impacts on the environment. Moreover, in vivo animal or human studies will not be performed. It is also clear that no experiments with human embryos/foetuses, human cells/tissues or with human patients will be undertaken. There are no foreseen dual use or misuse applications from the research to be done. The project will involve interaction with stakeholders. Although the discussions in the project activities will not touch on sensitive information of the individuals, and their organisations, any data gathered from participants during the events will be treated as confidential. Any data will be anonymised and will be adequately secured against misuse. The data controller nominated in the project will ensure that all personal data is destroyed at the end of the project.

All participants in LOCATE consultations and events will be presented with an informed consent form and be required to sign this form before participating in the activities. Copies of informed consent forms will be made available to the REA prior to the relevant work packages. No children under the age of 18 will be involved in this research and only participants capable of giving their consent will be allowed to participate.

Any national ethical requirements in countries where the research will be conducted and implemented will be respected. The appropriate local ethical authorisations will be attained, copies of which will be presented to the REA prior to commencing with the relevant research.

While no relevant ethical issues have been identified, there is a possibility of collection and storage of personal data in the frame of activities involving dissemination and validation. The beneficiaries will ensure full compliance with the General Data Protection Regulation 2016/679. The data controllers and processors are fully accountable for the data processing operations.   The beneficiaries will ensure that - if applicable - all the data intended to be processed are relevant
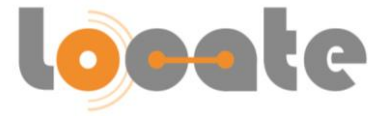
and limited to the purposes of the research project (in accordance with the 'data minimisation' principle).

END OF DOCUMENT