ReMAP

Real-time Condition-based Maintenance for Adaptive Aircraft Maintenance Planning

D9.3 Open Research Data Pilot – initial Data Management Plan

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769288

Grant Agreement No.	769288
Project Start Date	01-06-2018
Duration of the project	48 months
Deliverable Number	D9.3
Deliverable Leader	TU Delft
Dissemination Level (PU, CO, CI)	PU
Status	Final
Submission Date	30 November 2018
Author (Institution) Email	Bruno Santos (TUD) B.F.Santos@tudelft.nl Heather Andrews (TUD) H.E.AndrewsMancilla@tudelft.nl Marc Boonstra (TUD) m.h.boonstra@tudelft.nl
Project website	https://h2020-remap.eu

Document History

Revision Nr	**Description**	**Author**	**Review**	**Date**
v0.1	First Draft	Marc Boonstra (TUD)		18 10 26
v0.2	Updated Draft	Heather Andrews (TUD)		18 11 06
v0.3		Bruno Santos (TUD)	Data usage information	18 11 07
v0.4	Updated Draft	Marc Boonstra (TUD)		18 11 10
v0.5		Javier Garcia Hernandez	Internal Review	18 11 23
v0.6	Updated Draft	Heather Andrews, Bruno Santos, Marc Boonstra (TUD)		18 11 30

Acronym	Consortium
TUD	Technische Universiteit Delft (Coordinator)
ATOS	ATOS Spain SA
CTEC	CEDRAT Technologies SA
ENSAM	Ecole Nationale Superieure d'Arts et Metiers
EMB	Embraer Portugal SA
IPN	Instituto Pedro Nunes Associacao para a Inovacao e Desenvolvimento em Ciencia e Tecnologia
KLM	Koninklijke Luchtvaart Maatschappijnv
ONERA	Office National d'Etudes et de Recherches Aerospatiales
OPT	Optimal Structural Solutions LDA
STEC	Smartec SA
UTRCI	United Technologies Research Centre Ireland, Limited
UC	Universidade de Coimbra
UPAT	Panepistimio Patron

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 769288.

Index

1. Introduction

1.1. Project Summary

ReMAP "Real-time Condition-based Maintenance for adaptive Aircraft Maintenance Planning" (hereinafter also referred as "ReMAP" or "the project"), is a European project started on the 1st of June 2018 and has a duration of four years. The project addresses the specific challenge to take a step forward into the adoption of Condition-Based Maintenance in the aviation sector. In order to achieve this, a data-driven approach will be implemented, based on hybrid machine learning & physics-based algorithms for systems, and data-driven probabilistic algorithms for systems and structures. A similar approach will be followed to develop a maintenance management optimisation solution, capable of adapting to real-time health conditions of the aircraft fleet. These algorithms will run on an open-source IT platform, for adaptive fleet maintenance management. The proposed Condition-Based Maintenance solution will be evaluated according to a safety risk assessment, ensuring its reliable implementation and promoting an informed discussion on regulatory challenges and concrete actions towards the certification of Condition-Based Maintenance.

1.2. Purpose of this Document

Deliverable D9.3 Data Management Plan (DMP) addresses the way research data is managed in the ReMAP project within the Open Research Data Pilot (ORD Pilot). The ORD pilot aims to improve and maximise access and re-use of research data generated by Horizon 2020 projects, considering the need to balance openness and protection of sensitive information, commercialisation and Intellectual Property Rights (IPR), privacy concerns, as well as data management and preservation of questions.

DMPs are a key element for good data management, as they describe the management of the data to be collected, processed and published during a research project, creating awareness about research data management topics such data storage, backup, data access, data sharing, archiving and licensing. ReMAP hereby states the adherence to the FAIR[1] data principles, whereby research data is made Findable, Accessible, Interoperable and Re-usable for the community, responsibly considering possible data restrictions on public sharing.

1.3. Context

This Data Management Plan is closely linked to the report on D10.1 Ethics Requirements – POPD, submitted to the European Commission at the end of October 2018, whereby in section 4, a general overview of ReMAP data management strategy regarding interview data was described.

In the following chapters, we enfold the ReMAP DMP making use of the UK Digital Curation Centre template for Initial DMP. It is acknowledged that a DMP is a living document and, therefore, as the implementation of the project progresses and significant changes occur, we will update this plan accordingly on a finer level of granularity at the end of each project period (M18, M36 and M48) using the templates for the Detailed DMP and Final Review DMP.

It is important to mention that, at the outset of the project, we have engaged the Data Steward of the Faculty of Aerospace Engineering at Delft University of Technology (TUD), Dr. Heather Andrews (H.E.AndrewsMancilla@tudelft.nl). TUD has appointed dedicated Data Stewards[2] at every faculty with the goal of improving awareness of good research data management practices in a disciplinary manner. Data Stewards are then the first point of contact for research data management advice. The input and guidance of the faculty's Data Steward are at the basis of this plan.

[1] https://www.force11.org/group/fairgroup/fairprinciples
[2] https://www.tudelft.nl/en/library/current-topics/research-data-management/research-data-management/data-stewardship/ and https://openworking.wordpress.com/data-stewardship/

2. Data summary

Purpose

The aim of ReMAP is to develop a maintenance management optimization solution to monitor the real-time health conditions of aircrafts. The algorithms resulting from this project will run on an open-source IT platform built by the consortium.

The data collected, stored, protected and analysed throughout this project consist of:

- personal data from interviews and workshop participants (see D10.1 Ethics Requirements – POPD deliverable) and,

- technical data (KLM operations, aircraft sensors and structural laboratory tests).

Personal data from interviews and workshop participants will be used for both dissemination purposes and research purposes. Data collection, storage, protection and analysis procedures regarding personal data used for dissemination purposes has been already presented in the D10.1 Ethics Requirements – POPD deliverable. In this DMP, only the management of interviews data used for research purposes will be discussed.

Regarding the technical data, this consists of: i) data provided by KLM on aircraft operations, health monitoring and flight information; ii) laboratory test data on aircraft structural elements; iii) programming algorithms (code) to analyse and model different aspects of the maintenance management of aircrafts; iv) technology design and assessment results; and v) external data collected from multiple sources relevant for the project (e.g., EUROCONTROL Monoradar Weather Data or EASA Air quality data repository).

Table 1. Research activities to be done per partner lists a description of the research that will be carried out by each collaborating partner, and its purpose, focusing on the technical data mentioned above.

ATOS	IT Platform for Integrated Fleet Health Management solution (IFHM) Development of an IT platform to collect data from systems and sensors, and provide it to the different algorithms and decission support solutions. Collaborating partners: ENSAM , IPN, KLM, TUD
CTEC & STEC	Development of sensor technology for Structural Health Management (SHM) Procurement, development and integration of promising sensor technologies for damage monitoring in aeronautical composite structures. Collaborating partners: ENSAM, UPAT
ENSAM	Damage monitoring of complex aeronautic structures by means of Ultrasonic Lamb Waves Develop hardware and software systems able to monitor damages in composite structures by means of Lamb waves emitted and received by piezoelectric elements. Collaborating partners: CTEC
EMB	Development of predictive algorithms for aircraft systems Prognostics & Health Monitoring (PHM) Develop algorithms for predicting the Remaining Useful Life (RUL) of aircraft, based on data from aircraft models provided by KLM (KLC). Collaborating partners: ATOS, KLM, ONERA, UTRCI, UC
KLM	Development, verification and test of an IFHM KLM will provide the data to the rest of the partners. Data provided by KLM consist on operations data, aircraft health monitoring data, and flight data for different aircraft models. These datasets are commercially and safety-sensitive data. Thus, they are subject to strict institutional and national rules and protocols (see Annex A). Collaborating partners: ATOS, EMB, ONERA, TUD, UTRCI, UC
ONERA	Safety risk assessment of the IFHM Identification of hazards and safety barriers related with CBM technologies. Future CBM regulations and industrial processes discussion. Collaborating partners: EMB, KLM, ONERA, TUD
OPT	Design and manufacture of aircraft structure cupons and study of ReMAP's impact on aircraft weight Design, test and manufacture components for experimental tests (WP4). Study of the impact of a Condition-based Maintenance (CBM) approach in weight reduction of aircraft structures.

	Collaborating partners: -EMB, TUD, UPAT
TUD	Development of predictive algorithms for aircraft structures and systems & maintenance scheduling decision support tool & safety risk assessment Develop validated multi-disciplinary SHM system methodologies towards remaining useful life estimation (prognosis) in the presence of adverse conditions during flight. Several sensing technologies are going to be used along with an ambitious extended test campaign. This campaign will result in a massive SHM database upon which the diagnostic and prognostic methodologies are going to be developed and validated. Development of the adaptive aircraft fleet maintenance schedule solution, including the definition of an uncertainty mapping. Model development for safety assessment of CBM technologies included in the IFHM. Collaborating partners: CTEC, EMB, ENSAM, OPT, STEC, TUD
UTRCI	Development of system level Prognostics & Health Monitoring (PHM) & Condition-Based Monitoring (CBM) technologies Develop PHM models to predict and detect degradation and failures in aircraft systems and components by using data on aircrafts, weather conditions, component removals, among others. Collaborating partners: ATOS, EMB, KLM, ONERA, UC
UC	Development of system level Prognostics & Health Monitoring (PHM) & maintenance planning decision support tool Enabling edge computing and Actionable information extraction and visualization for optimal maintenance Development of efficient machine learning algorithms for optimal maintenance. Development of a user interface for the maintenance planning decision support tool. Development of an adaptive plan and uncertainty mapping. Collaborating partners: ATOS, EMB, IPN, KLM, ONERA, TUD
UPAT	Structural Health Management: Diagnostics & Remaining Useful Life (RUL) Prognostics Develop validated multi-disciplinary SHM system methodologies towards remaining useful life estimation (prognosis) in the presence of adverse conditions during flight. Several sensing technologies are going to be used along with an ambitious extended test campaign. This campaign will result in a massive SHM database upon which the diagnostic and prognostic methodologies are going to be developed and validated. Collaborating partners: CTEC, EMB, ENSAM, OPT, STEC, TUD

Table 1. Research activities to be done per partner

Data Types and Formats

As explained above, there are two main sets of data in this project: technical data (KLM data, laboratory data, programming algorithms, design and assessment data, external data) and personal data.

Technical data:

The IFHM solution resulting from ReMAP will be developed, validated and demonstrated based on KLM's operational data. KLM will follow internal protocols to anonymize all data before sharing it with the consortium partners, meaning that the research team will not have access to personal data.

The KLM operational data includes:

- Aircraft Health Monitoring data, used as input and/or validation data for to-be-developed model for assessing condition and prognosis of individual aircraft systems. This data is owned by KLM and it is restricted data under governmental and company regulations.
- Aircraft Maintenance data, used to support the development of the maintenance schedule solution and as validation data for to-be-developed model for assessing condition and prognosis of individual aircraft system. This data is owned by KLM and it is restricted data under governmental and company regulations.
- Risk assessment, used to map and mitigate operation, technical, commercial, economical and health & safety risks that are associated with aircraft maintenance. This data is owned by KLM and it is restricted data under governmental and company regulations.

Most of these datasets correspond to log files, reporting documents and tabular data. Health monitoring data and flight data from aircraft models are observational data from sensor measurements. KLM will provide the monitoring data to partners in .csv format, while log files

and reports will be provided in .pdf format. It is important to mention no personal information (e.g. ground staff, flight crew, etc.) will be disclosed by KLM. No information will be disclosed that can, directly or indirectly, be linked to KLM staff either.

Aside KLM's data, there will be experimental data from laboratory tests on aircraft structural elements and composite generic elements and subcomponents, typically found in modern commercial aircraft. This type of data will be generated during the project at UPAT and TUD premises. The data correspond to sensor recordings obtained during the tests, as well as Finite Element Analysis outputs from simulation endeavors. Various file formats will be involved depending on the monitoring technique and the associated software that is utilized to record the data. However, all data will be converted to .txt, .csv, or .dat files after some raw data processing to increase data interoperability.

The codes that will be generated throughout the project will be mainly in MATLAB, Python and R languages. For their development and validation, researchers will also make use of external data (e.g. on weather conditions, pollution information, etc.) collected from multiple sources, which shall include public repositories and other existing data gathering channels available to partners (e.g., EUROCONTROL Monoradar Weather Data or EASA Air quality data repository).

Finally, the information generated with the design and development of sensor technology and aircraft structures will be produced during the project, together with data related with the assessment of the performance of the multiple technologies developed for the IFHM solution proposed.

Personal Data:

As mentioned in Section 3 of D10.1 Ethics Requirements – POPD deliverable, the project will carry out interviews to team members and external workshop participants. In order to collect, store and use the personal data from interviews, the consortium shall seek the informed consent of each individual, following the policy of the EU for Data Protection (see D10.1 Ethics Requirements – POPD deliverable). The individual subjects will be informed about all aspects of the research in which they are being asked to participate and the future use of the data they might provide.

The interviews will be recorded as audio-visual footage. The recording of each interview will be stored in the work laptop/computer of the IPN researcher in charge of the interview. The data will be saved in a private password-protected folder accessed only by the respective IPN researcher. The IPN team or another partner will anonymise and transcribe the data (e.g., into .docx files). The transcriptions will then be shared with relevant consortium researchers via SURFdrive (https://www.surfdrive.nl), which is a password protected cloud storage service. The raw interview data will be then transferred to a private repository in the DataverseNL environment (https://dataverse.nl/). This environment is expected to be safe enough for the recorded material. In case the raw interview data contains highly sensitive information, then the data will be saved in a Project Data drive at TUD. This is a drive maintained by IT TUD and it is meant for confidential data. The folder containing the raw data will be managed by Dimitrios Zarouchas (d.zarouchas@tudelft.nl) from TUD, and accessible to the project coordinator Bruno Santos (b.f.santos@tudelft.nl) and to Mónica Ferreira (mferreira@ipn.pt), the IPN coordinator for WP9.

If it becomes pertinent for the research purpose of the project, the interview data might be released to the public. This will happen only if the respective interviewee agrees on it via email, reacting to a consent request sent by either IPN (WP9 leader) or TUD (project coordinator) explaining the context, purpose, content to be made public and the right to reject this request, which is assumed by default. If any of the interviewees request not to keep the recordings but for getting notes, they will be deleted from the repositories and only the notes will be circulated among the consortium.

Data Size

The estimated size of the data delivered by KLM to partners is about 2 TB in total, taking as a reference data from 30 aircraft operating during a time frame of 3 years. Included in these 2 TB is also the external data (e.g., weather conditions and pollution information) taken from public repositories and/or data gathering channels available to partners.

The estimated size of the laboratory data is expected to be in the order of 100 GB.

The processing of the data and its use for algorithm development might be on the order of 1 TB.

The data regarding the design and development of sensor technologies and aircraft technology, together with the data resulting from the technology assessment, should be in the order of few tens GB.

The estimated size of the interview data considering the audio-visual footage is undetermined, but it might be in the order of few hundred GBs (considering they will mainly be audio files and the transcribed files into .docx documents).

Data Utility

The final outcome of ReMAP will be useful for aircraft manufacturers (OEMs), maintenance service providers and airlines around the world. The data generated throughout the project will be useful for:

- The data from the laboratory tests will be useful to structural scientists, OEMs and airline researchers, for the future development and training of SHM prognostics and diagnosis data-driven algorithms.
- The sensor technology design and reliability performance data will be useful to sensor companies and OEMs, for the analysis and development of sensor solutions for SHM in future aircraft.
- The safety risk assessment data will be essential for the development of the common roadmap towards the implementation of CBM in practice. In particular, it will be useful for the discussion with EASA about continuing airworthiness regulations part-M and part-145, and with the Maintenance Steering Group-3 (or MSG-3) regarding aircraft maintenance procedures and they could be adapted to CBM.
- The IFHM validation and test data will also be essential for the development of the common roadmap. In this case, regarding the involvement of maintenance service providers and airlines in a common solution to the implementation of CBM. The results will test the overall capability of the multiple CBM technologies developed, including the adaptive aircraft maintenance schedule approach proposed.

3. FAIR data

3.1. Making data findable, including provisions for metadata

Metadata

During the research project, KLM will deliver the data to partners in a structured way with proper documentation (README file) indicating (at least):

1. Data origin and collection methodology
2. Structure and organisation of data files
3. Data manipulations applied prior to sharing
4. Variable names and descriptors (if applicable)
5. Definitions of terminology and acronyms

At the moment there are no metadata standards defined, but KLM might select a metadata standard at a later stage for partners to use when describing the data. Even though KLM data will not be disclosed to the public (because of safety and commercial reasons; see Annex A), such standard would be the one used for other publishable datasets for consistency within the project. If no discipline specific standard is used, then Dublin Core (www.dublincore.org) metadata standard will be adopted (and further information on the data will be delivered in file headers or in separate documentation files). This last statement also applies to documentation files (e.g. reports) and other types of data including experimental data and input/output tabular data for code training, testing and validation.

Regarding codes, these will be developed and managed mainly via Gitlab (partners may use their account or use an account provided by ATOS) and Subversion (SVN), where both allow easy metadata attachment. The metadata standard to be adopted for all technology codes will be discussed as part of the IT platform requirements and specifications (WP2). As stated above, if no discipline specific standard is used, then Dublin Core metadata standard will be adopted (and further information on the data will be delivered in file headers or in separate documentation files).

Once scientific journal publications are published (in Open Access), publishable data (according to the Consortium Aggrement) will be publicly archived for the long term via the 4TU.Centre for Research Data archive (documentation, experimental data and tabular data; https://researchdata.4tu.nl/en/home/), or similar online archives, following their metadata standards (Dublin Core). TUD researchers can, nowadays, upload up to 1 TB of data per year free of charge. This should suffice for the data that will be archived for the long term. The 4TU.Centre for Research Data Archive ensures data will be well-preserved and findable in the long term (each uploaded dataset is given a unique persistent digital identifier). In order to allow for responsible public reuse of the data, datasets will be publicly released under an open content license (CC-BY). More specific, metadata of the files in the dataset will be given in XML format (if necessary) following the standards agreed upon during the project.

Version Control and File Naming Conventions

Partners working on programming algorithms will use Gitlab repositories to collaboratively work with their research team members. Gitlab allows for clear code version management, and it is already available to the respective partners.

Subversion tool will also be used by some partners to keep track of versioning for documents related to other technical data files.

Reports and other types of data will be managed manually for which file naming conventions will be followed as the project progresses. This will described in D1.1 Project Handbook.

3.2. Making data openly accessible:

The data provided by KLM cannot be open to the public for safety and commercial reasons (see Annex A for the provisions for KLM Operational Data use). Processed and/or analysed data might be released to the public via the 4TU.Centre for Research Data Archive after proper discussion with KLM (KLC). Codes and auxiliary scripts built upon the processed data provided by KLM might be released via GitLab after proper discussion with all partners.

During the project, laboratory data will be accessible only to consortium partners. Some processed data might be subject to Intellectual Property Rights of the respective partner(s) that generate the data, and thus, will be restricted for use within the consortium only. Whenever a journal article is ready to be published (in Open Access), the laboratory data related to the journal article will be released openly to the public via the 4TU.Centre for Research Data Archive. This data includes all data necessary for the re-use of the results, as well as the data needed to validate them.

The final technological outcome of ReMAP will be an open IT cloud platform where the finalized algorithms can work in an interoperable manner. This IT platform will be open source. The platform will be built in a modular way, following an architecture that will allow the integration of third-party data analytics and maintenance management solutions or the exploitation of solutions developed by the ReMAP consortium (these ones, not necessarily open).

Regarding the interviews to team members and workshop participants, the raw interview data (audio visual footage) will not be released to the public. It will only be accessible to relevant partners for a maximum period of 10 years and erased afterwards. The data will be kept in the long term for recording and auditing of the project and so it can be used for future research and learning, unless the interviewee does not grant such a permission. Nonetheless, as described, the participants have the “Right to be forgotten”. They can request at any time the elimination of their personal data (such as names, emails, contact details, information from interviews) stored in the project data storage services and IPN servers.

If requested by the participants, when given consent for the interview, only anonymized transcripts of the interviews might be used for research purposes. The anonymized transcripts will be available only to relevant consortium partners during the project. If the transcripts are used as material for a journal article, then anonymized transcripts will be published via the 4TU.Centre for Research Data Archive at the same time the journal publication is released.

The public outcomes of workshops that can be made public will be shared via DataverseNL, with reference to it via the project’s website. Unless an informed consent is given, the public outcomes from these workshops will be free of any personal reference. The workshop coordinator will be responsible for preparing these results and anonymize the information when necessary.

Documentation and Software for Data Access

As all openly publishable material will be made available via the 4TU.Centre for Research Data Archive, DataverseNL or GitLab, the datasets’ models will be easily downloadable from these platforms with the respective metadata.

Open and standard formats will be preferred for archived data files (e.g., .csv, .txt) and code (e.g., python). Proper documentation files will be delivered together with the datasets in order to facilitate reuse of data.

In some cases, MATLAB files (.mat) will also be released, as MATLAB will be used to record, process and visualize data in some research lines of ReMAP. MATLAB is a licensed software, widely used in the engineering community, and it is already available to all partners. CATIA software will also be used for computer-aided engineering. This is a licensed software already available to partners. Output design files of CATIA software will be converted to other formats whenever the data can be open to the public, to facilitate reuse.

3.3. Making data interoperable:

As mentioned above, all publishable data will be delivered in open and standard data formats. Discipline specific metadata is currently under discussion. If applicable, metadata will be delivered in XML/JSON format together with the data (depending on the chosen format).

Proper documentation (README) files will be delivered accordingly. Tabular data and codes (auxiliary scripts) will be archived with informative and explanatory headers to facilitate data re-use and interoperability.

All code will be managed via Gitlab and/or SVN, which are interoperable with one another, and are platforms that encourage interoperability between different workflows.

The final IFHM solution will gather different algorithms that will be interoperable with each other. The open IT platform over which the algorithms will be implemented, will have proper documentation and manuals for researchers to use.

3.4. Increase data re-use (through clarifying licenses):

The data provided by KLM may never be disclosed because of safety and commercial reasons, not after the end of the Project and not after the end of the 4-year non-disclosure term stated in the Consortium Agreement (see Annex A). All other data that cannot be disclosed (except KLM data) will be kept at the respective institutional server for the long term (at least 4 years after the end of the project); accessed only by team members within the institution, for auditing and validation purposes. It is also acknowledged that, for some of the outcomes, copyright and IPR rules as stated in the Consortium Agreement may apply. This will include some of the Prognostics & Health Monitoring (PHM) solutions to be developed. The following release of this data will be determined based on an internal review and decision from the Steering Committee, supported by the Project Board.

Since the results from this project will make a strong impact in the aviation sector (airlines, manufactures, maintenance service providers, etc.), we find it is extremely important to share the data responsibly. Hence datasets that will be open to the public will be released along the journal scientific publications after proper discussion with partners. The datasets will be published via repositories such as GitLab (algorithms, code, auxiliary scripts) and the 4TU.Centre for Research Data Archive (documentation, images, tabular data, etc.) under open content licenses in order to increase data re-use (e.g., CC-BY license for documentation and MIT license for software and code). In the same way, and in order to motivate re-use of data, the journal articles associated to these datasets will be published in open access and/or self-archived in ReMAP's website and subject repositories, following the publisher's self-archiving policies.

Regarding the final IT platform, this will be distributed, deployed, and explored according to an open source license. The platform itself, and the users may integrate their data analytics solutions or use ReMAP's proposed solutions (both the ones made available in the public domain, and the ones copyrighted subjected to usage fees).

4. Allocation of resources

Costs

In principle no costs are expected for the archiving of the publishable data via the 4TU.Centre for Research Data Archive nor via Github. TUD researchers can upload nowadays up to 1 TB of data to the 4TU.Centre for Research Data Archive (per year) free of charge. TUD researchers also have free of charge access to the DataverseNL environment. Also, most of the software used for version control and data processing are already available at each institution, as well as the storage capacity and privately accessed drives managed by each institution's IT department.

The IT platform will be developed and implemented in ATOS infrastructure, using Cloud services with rental costs included in the budget of the project (ATOS 'other goods and services' budget).

Responsibilities

The following table specifies the team members who will be in charge of the management of the data, within each research line of study. It is important to mention that each institution has support staff that can provide advice whenever data management issues arise, and who will be contacted if it is necessary.

Partner	Name	Email address
TUD	Dimitrios Zarouchas	d.zarouchas@tudelft.nl
ATOS	Javier García Hernández	javier.garcia@atos.net
ENSAM	Nazih Mechbal	nazih.mechbal@ensam.eu
CTEC	Frank Claeyssen	frank.claeyssen@cedrat-tec.com
EMB	Rúben Menezes	ruben.menezes@pt.embraer.com
KLM	Floris Freeman	Floris.Freeman@klm.com
OPT	Nicole Cruz	nicole.cruz@optimal.pt
UTRCI	Anarta Ghosh	ghosha@utrc.utc.com
UC	Bernardete Ribeiro	bribeiro@dei.uc.pt
UPAT	Theodoros Loutas	thloutas@upatras.gr

In case any of the above team members is unavailable, other research team members (within one institution) will have access to the data, as all data will be stored on the respective institutional servers, with provided access only to team members.

5. Data security

KLM will take care of anonymizing the data before sharing it with partners. Servers that are set-up by KLM will have a redundancy scheme (e.g. RAID5 or alternative) or backup plan to mitigate risks of disk failure. KLM internal servers will keep a copy of the original raw data, in case the processed data is no longer available. Also, KLM will monitor ReMAP's process through an internal knowledge management system (e.g. Confluence). All data that KLM staff generated for ReMAP will be stored on an Office365 or similar environment, such that other KLM staff can pick up work in case of long-term illness or unforeseen staff changes.

Some data will be processed in work laptops of research team members only when allowed (given the sensitivity of the data). Master copies will be kept at the drives of each respective institution. The IT departments of each institution will maintain the data regarding backups (redundancy) and secure storage (protected access to only team members). Only team members within each institution will have access to the data during the research project. Such data access will be set up by the respective IT departments of each institution. The data that will remain close to the public will be archived at each partner's servers for at least 4 years after the end of the project.

Surfdrive will be used for temporal data storage and for data sharing among different partners coordinated by TUD (coordinator). Google Drive might also be used for the sharing and temporal storage of non-sensitive data.

6. Ethical aspects

Please refer to D10.1 Ethics Requirements – Protection of Personal Data (POPD) deliverable for the management of personal data information for both communicational and research purposes.

It is important to mention, in case there are ethics-related questions or issues arising throughout the project, these will be reported to Bruno Santos (b.f.santos@tudelft.nl) and will be discussed accordingly among team members. Extra advise can be discussed with the Human Research Ethics Committee of TUD (at HREC@tudelft.nl).

7. Other

ReMAP will make use of the TUD Research Data Framework Policy which can be found following: https://www.tudelft.nl/en/2018/library/researchdatamanagement/tu-delft-research-data-framework-policy-published/

Annex A: KLM Operational Data provisions

[relating to Section 9.9 of the Consortium Agreement]

Introduction: the ReMAP program aims to research Condition Based Maintenance, making use of real airline industry data for development, demonstration and validation purposes. KLM operational aircraft and maintenance data is, therefore, considered an important asset to the research program.

Because of the sensitive nature and value of KLM's Operational data, KLM imposes several constraints to the availability of its data and to the use of its data.

Definition of Operational data: Data generated by KLM Aircraft, KLM Aircraft operation and Maintenance operation to support the purpose of the research, such as, but not limited to:

- Aircraft Health Monitoring data: generated by KLM aircraft during flight (e.g. system sensors).
- Aircraft Operational data: network planning, maintenance planning, etc.
- Operational Maintenance events: activities and events related to aircraft maintenance, etc.
- Maintenance findings and records: maintenance logs, technical findings, repair data, etc.

Provisions concerning KLM operational data:

- Access rights: Access Rights apply to the KLM Operational Data.
- No rights can be derived from the KLM Operational Data.
- No Access Rights to KLM Operational data will be given to any party for the Exploitation of any Results as described in article 9.4 of the Consortium Agreement.
- KLM Operational Data may never be disclosed, not after the end of the Project and not after the end of the 4 year non-disclosure term as described in article 10.3 of the Consortium Agreement.
- KLM Operational data cannot be disclosed to any Affiliated Entity or Third party without prior written approval by KLM.
- KLM Operational Data cannot become (part of) a Result or be included in (Joint) Ownership of Results.
- Joint ownership of Results: For any Results generated by any Party, where KLM Operational data was used to generated the Results, KLM is entitled to Joint ownership of these Results, provided that the KLM Operational data does provide added value for the creation of a new product or improvement of an existing one. Results that only transport or visualize KLM Operational data cannot be considered to be part of joint ownership with KLM.
- Data availability: The KLM Operational data will be made available to the ReMAP consortium partners by KLM through a KLM deployed solution.
- Data confidentiality & limitations: The following rules apply when making use of KLM Operational data:

Req.#	Requirement description
	Compliance
1	The KLM Operational data may only be stored and hosted in the Netherlands.
2	The KLM Operational data cannot be locally stored or downloaded.
3	Access to the KLM Operational data has to be arranged through authorization management and logins.
4	The KLM Operational data may not be coupled or related to other data which makes the relation between the data the flight crew/pilot of the flight. Flight crew anonymity must be guaranteed at all times.
5	Any given moment, a list must be produced to KLM on which persons have access and authorization to the KLM Operational data.
6	Persons that have access to the KLM Operational data are informed on the confidentiality and limitations of the KLM data. The limitations is that this KLM Operational data may only be used for the purpose of ReMAP. The KLM Operational data shall not and cannot be used to investigate on the behavior and acting of (flight) crew in any sense and for any purpose.
7	Persons that have access to the KLM Operational data will be reminded on a yearly basis on the confidentiality and limitation of the KLM Operational data.
8	The KLM Operational data shall not be shared with any other parties outside the ReMAP consortium.
9	On request of the KLM, all KLM Operational data can be permanently removed.
10	On request of the KLM, all KLM Operational data sharing can be stopped.