FLAIR	**H2020**	

TEKEVER	csem	DTU	Empa Materials Science and Technology
NKT Photonics the power of light	NIT New Infrared Technologies	Radboud University Nijmegen	SenseAir Sensors for Life

DATA MANAGEMENT PLAN

FLAIR - FLying ultrA-broadband single-shot InfraRed Sensor
GA732968

Deliverable Information	
Deliverable Number: D1.2	**Work Package:** WP1
Date of Issue: 26/07/2017	
Document Reference: 732968-FLAIR-D1.2-DataManagementPlan	
Version Number: 1.0	
Nature of Deliverable: R	**Dissemination Level of Deliverable:** PU
Author(s): TEK-AS (Responsible), CSEM	
Keywords: Data management, Open research data pilot, FAIR, DMP	

Abstract:
This document constitutes the first version of the FLAIR DMP as required by the Grant Agreement. In it the reader will be able to get an overview of the dataset that will be generated by the consortium during the project's execution and the strategy and plans the consortium will implement to ensure open access to relevant research data. Some data generated or used by the project will not be made available either because it is readily and, in some cases, freely available from the original sources or because the data has commercial potential and as such will be exploited by the project partners. This version of the DMP will be updated midway through the project and at the end of the contract.

Document History

Date	Version	Remarks
06/03/2017	0.1	Skeleton
27/04/2017	0.2	First draft
31/05/2017	0.3	Updates based on progress in identification of relevant data to be generated in the project
05/07/2017	0.4	Updates
26/07/2017	1.0	Revision and first issue

Document Authors

Entity	Contributors
TEK-AS	André Oliveira
CSEM	Gilles Buchs

Executive Summary

The FLAIR project addresses the challenge of performing air quality monitoring in complex urban environments by mounting a high-performance air sampling sensor based on photonic technology on an UAV for pervasive and large area coverage high-specificity and high-sensitivity air quality sensing. Operating in the two atmospheric windows of 2-5 μm and 8-12 μm wavelength, FLAIR can detect minute traces of molecules in complex gas mixtures from their characteristic IR absorption fingerprints and provide real time information to the operator of the drone. FLAIR can operate in remote or dangerous areas and outside of established monitoring networks.

FLAIR participates in the Pilot on Open Research Data (ORD) launched by the European Commission (EC) under H2020. The ORD pilot aims to improve and maximize access to and re-use of research data generated by H2020 projects. As such, the development and use of a Data Management Plan (DMP) is required for all projects participating in the ORD Pilot. This document constitutes the first version of the FLAIR DMP as required by the Grant Agreement.

This DMP has been created and developed following the Guidelines on FAIR Data Management in Horizon 2020 published by the EC, the UK's Digital Curation Center (DCC) guide on How to Develop a Data Management and Sharing Plan and the European Research Council's (ERC) Guidelines on Implementation of Open Access to Scientific Publications and Research Data. It describes the current picture of datasets to be generated or used in the project and provides answers to the questions of Annex 1 of the EC Guidelines on FAIR Data Management in Horizon 2020.

For each dataset to be generated in FLAIR, the following elements (when known or defined at the time of writing) are provided:

- What is the purpose of the data collection/generation
- Explanation on the relation of the data to the objectives of the project
- Specification of the types and formats of data generated/collected
- Explanation of whether existing data is being re-used
- Specification of the origin of the data
- Determination of the expected size of the data
- Identification of the data utility
- Consortium's plan to make the data findable
- Consortium's plan to make data openly accessible
- Consortium's plan to make data interoperable
- Consortium's plan to increase data re-use
- How the consortium will allocate resources to manage the data
- How the consortium will ensure data security
- Ethical aspects

The document concludes by identifying other data that will be used by FLAIR or generated by the project but for which no open access will be granted. This is either because the used data is readily and, in some cases, freely available from the original sources (e.g. HITRAN database and PNNL library spectral data) or because the data generated in the project has commercial potential and as such will be exploited by the project partners (e.g. UAV performance data and FLAIR system design data).

Table of Contents

List of Tables

List of Acronyms

Acronym	Meaning
AFCRL	Air Force Cambridge Research Laboratories
DCC	Digital Curation Centre
DMP	Data Management Plan
EC	European Commission
EMPA	Swiss Federal Laboratories for Materials Science and Technology
ERC	European Research Council
FAIR	Findable, Accessible, Interoperable and Re-usable
GA	Grant Agreement
GPS	Global Positioning System
HITRAN	HIgh resolution TRANsmission molecular absorption
HW	Hardware
IPR	Intellectual Property Rights
IR	InfraRed
IT	Information Technologies
KET	Key Enabling Technologies
ORD	Open Research Data
PNNL	Pacific Northwest National Laboratory
RIA	Research and Innovation Action
UAV	Unmanned Aerial Vehicle

Table 1 – List of acronyms.

1 Introduction

1.1 FLAIR Objectives

FLAIR (Flying ultra-broadband single-shot InfraRed sensor) is a Research and Innovation Action (RIA) funded by the European Union's H2020 programme under the Photonics Key Enabling Technologies (KET) topic.

Its core objective is to develop a high-performance air sampling sensor based on cutting-edge photonic technology capable of performing high-specificity and high-sensitivity (ppbv) air quality sensing in large areas as a result of its installation aboard an Unmanned Aerial Vehicle (UAV), also known as a drone.

Today, significant effort is devoted globally to improve air quality through e.g. land-use planning strategies, replacement of fossil fuels by clean energy sources and lower level of industrial emission. In order to be successful, these measures need be accompanied by air quality monitoring at large scale to ensure compliance with air quality legislation but also to provide information for political decision making regarding air quality and safety. This is particularly challenging outside the dense urban network of air quality monitoring stations. FLAIR addresses this challenge by mounting a high-performance air sampling sensor based on photonic technology on an UAV for pervasive and large area coverage high-specificity and high-sensitivity air quality sensing. Operating in the two atmospheric windows of 2-5 µm and 8-12 µm wavelength, FLAIR can detect minute traces of molecules in complex gas mixtures from their characteristic IR absorption fingerprints and provide real time information to the operator of the drone. FLAIR can operate in remote or dangerous areas and outside of established monitoring networks.

FLAIR applications include the monitoring of air around industrial infrastructure, maritime and land based traffic, landfills and agriculture facilities and the project contributes to a safer environment by providing detailed air quality data around current facilities and locations of interest or in the case of catastrophic events like wildfires, volcanic eruption or chemical accidents.

The advantages of using UAVs for this application are essentially related to the fact that these vehicles can rapidly access areas that are too dangerous or too difficult to reach by humans. Moreover, due to the local sampling FLAIR can provide data from inside optically dense clouds and plumes that are not accessible by ground based laser remote sensing methods.

Photonics technology is a promising approach to the challenge of air quality monitoring, as it can provide, in principle, accurate identification and concentration measurements of specific species in complex environments. Current solutions include several methods for air quality monitoring, among which are mass spectrometry, electronic noses and optical detection. While systems based on mass spectrometry are highly sensitive, they suffer from complexity and high footprint. Electronic noses are cheap but suffer from low accuracy.

Several systems based on light sources operating in the IR range such as quantum cascade lasers, diode laser, optical parametric oscillators or frequency combs have enabled highly sensitive and selective detection of molecules. Such high performance tools, however, typically remain confined to academic research laboratories due to their narrow spectral operating window (covering only very few molecules), their operational complexity and their prohibitively high cost. These are the technical challenges FLAIR is addressing.

The FLAIR project will generate trace gas absorption spectra from which information on the levels of pollutants can be derived through data processing. The datasets generated

throughout the project will be fundamental to validate the sensor prototype itself, to benchmark its performance against current standards and to assess the feasibility of using unmanned vehicles as a new instrument for deploying air quality sensors and implementing mobile dynamic measuring stations.

The FLAIR project participates in the Pilot on Open Research Data (ORD) launched by the European Commission (EC) under H2020. The ORD pilot aims to improve and maximize access to and re-use of research data generated by H2020 projects. As such, the development and use of a Data Management Plan (DMP) is required for all projects participating in the ORD Pilot.

1.2 Purpose of the Data Management Plan

A Data Management Plan consists of a living document that describes the data management life cycle for the data collected, processed and generated by a H2020 project. It is considered a key element of good data management. The plan outlines how data will be created, managed, shared and preserved throughout the project, providing arguments for any restrictions that apply to any of these steps or any of the data.

The EC encourages all projects to follow and apply principles that will allow all research data to be Findable, Accessible, Interoperable and Reusable (FAIR principles).

The research data generated or created under the projects may include statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images. Open access to research data is important to allow validation of the results presented by project researchers and consortia in scientific publications.

As the project evolves and the research progresses, datasets will be created and may be subject to changes or updates in terms of the types, formats and origins of the data. Furthermore, the way the data is named or made accessible may change according to consortium policy changes and/or identification of potential for exploitation by project partners. Therefore, it is expected that the DMP will not remain unaltered throughout the project's lifespan. This document constitutes the first version of the DMP and an official deliverable of the project as defined in the Grant Agreement (GA) of FLAIR.

The FLAIR DMP will be updated at months 18 (April 2018) and 36 (December 2019) in line with the mid-term and final reviews respectively, as recommended in the EC Guidelines on FAIR Data Management in Horizon 2020.

The obligations to disseminate results (Article 29.1 of the GA), provide open access to scientific publications (Article 29.2 of the GA) and open access to research data (Article 29.3 of the GA) do not, in any way, change the obligation of consortia to protect results, ensure confidentiality obligations and the security obligations or the obligations to protect personal data, all of which still apply. Consequently, consortia do not have to ensure open access to specific parts of their research data if the achievement of the action's main objective or the exploitation of results would be jeopardised. Whenever this is the case, the DMP will identify said data and explain the reasons for not giving access.

1.3 Approach

The FLAIR DMP has been created and developed following the Guidelines on FAIR Data Management in Horizon 2020 published by the EC in 2016 as well as the UK's Digital Curation Center (DCC) guide on How to Develop a Data Management and Sharing Plan from 2011 and

the European Research Council's (ERC) Guidelines on Implementation of Open Access to Scientific Publications and Research Data from 2017.

1.4 Maintenance of the FLAIR DMP

The FLAIR DMP will be maintained throughout the project's lifespan by the coordinator with support from the other partners. This activity falls under WP1 of the project (Project Management).

2 FLAIR data

Large amounts of data in different formats will be collected and generated during the FLAIR project. These will consist mainly of physical parameters, designs, blueprints for hardware, electronic circuit design, and technical data. These data will be shared and exploited according to the policies established under the Grant and Consortium Agreements of the project. All data collected during the project will be placed in the official FLAIR repository, where they will be available for authorized persons and will be properly secured against theft.

In the context of the Pilot on ORD in H2020, all data will be made accessible to the general public for verification and further re-use. However, it is necessary to take into account the possibility that some of them will be part of the Intellectual Property Rights (IPR) of individual partners and therefore will be protected. The following sections describe the current picture of datasets to be generated or used in the project and provide answers to the questions of Annex 1 of the EC Guidelines on FAIR Data Management in Horizon 2020.

2.1 FLAIR laboratory dataset

2.1.1 Summary

2.1.1.1 Purpose of the data collection/generation

The FLAIR laboratory dataset will comprise the set of measurements of trace gas absorption spectra accomplished with the FLAIR sensor in a controlled laboratory environment. The purpose of this dataset is to validate the FLAIR sensor design and components, and to characterize its performance against known quantities of the target trace gases.

2.1.1.2 Relation to the objectives of the project

The FLAIR laboratory dataset is fundamental for achieving the project core objectives. It will be through this dataset that the design of the FLAIR sensor and the prototype of the sensor and its components will be evaluated and characterized. The sensor's performance will be benchmarked against standard sets of parameters such as the HITRAN database and the PNNL spectral library described in sections 3.1 and 3.2.

2.1.1.3 Data types and formats

The data comprising the FLAIR laboratory dataset will consist of:

- Trace gas concentrations in ppbv (parts per billion volume).
- Gas absorption spectra in the 2-5 μm and 8-12 μm windows as a set of absorbance per wavelength measurements.

At the moment of writing, the consortium plans to generate the data above for the following gases (according to project deliverable D2.1 – Requirements for FLAIR Sensor System):

- CO_2 – Carbon Dioxide
- CH_4 – Methane
- CO – Carbon Monoxide
- O_3 – Ozone
- N_2O – Nitrous Oxide
- SO_2 – Sulfur Dioxide
- NH_3 – Ammonia
- HCl – Hydrogen Chloride

Comparison results between the results obtained and the standard values of benchmarks may also be included in the dataset as absolute and relative differences.

The consortium does not foresee the inclusion of the data processing algorithm details in the dataset as this can be commercially exploited by the project partners.

2.1.1.4 Re-use of existing data

No existing data will be re-used in the generation of the FLAIR laboratory dataset.

2.1.1.5 Data origin

The data for the FLAIR laboratory dataset will be generated at the laboratories of partners of the FLAIR consortium directly by the researchers involved in the project. The generation procedures will be described in detail in the appropriate deliverables and associated with the dataset.

2.1.1.6 Expected size of the data

At the moment the expected size of the dataset is not known.

2.1.1.7 Data utility

The FLAIR laboratory dataset will be useful to the entire consortium (namely in helping accomplish the project objectives as outline above) as well as for other researchers wishing to evolve the work of FLAIR and potential customers interested in new sensors based on the FLAIR technology.

2.1.2 FAIR

2.1.2.1 Making data findable

The FLAIR laboratory dataset will be made discoverable through the association of metadata to the dataset. At the moment of writing the type of metadata and identification mechanism to be applied is not yet defined. The process or standard to be used to create the metadata is not clear yet. However, the consortium expects to associate the following metadata to the dataset:

- Date of measurement
- Gases measured
- Absorption windows

Files in the dataset will be clearly named and their names will include the date of measurement, gas to which the file refers to and whether the file refers to the absorption spectrum or the detected concentration. The following is an example of such a name: 20180428-FLAIR-CO2-AbsSpectrum.ext.

All files in the dataset will allow the clear identification of the version. This may be achieved through the addition of a version suffix to the filename or by supporting versioning in the FLAIR repository. The chosen solution is not defined at the moment. In case the consortium opts for versioning through the filename, then the file itself will describe in its initial contents the major changes to the previous version (e.g. different laboratory test run).

2.1.2.2 Making data openly accessible

At the moment of writing, the consortium expects to make the entire FLAIR laboratory dataset openly available. The consortium expects to make the dataset available through the project's repository (which is foreseen to support versioning). The repository is maintained by the project coordinator and access to it is authenticated. Access to the repository will be enabled through a web interface that only allows download of the dataset (i.e. it will not be possible to delete, upload, check-out or commit other files).

Registration to the repository will be required and will consist in providing the name, entity and reason of interest or foreseen purpose of use for the FLAIR dataset.

2.1.2.3 Making data interoperable

FLAIR will provide the data in the FLAIR laboratory dataset in the standard units for detected concentration and those used in absorption spectra. This should be enough to ensure the interoperability of the data. Parts per billion volume (of air) is a commonly accepted unit for measuring concentrations of compounds in the air. The standard way of presenting an absorption spectrum is to represent the absorbance (unitless) against the wavelength in μm or nm depending on the spectrum window of interest.

2.1.2.4 Increasing data re-use

At the moment of writing the consortium has not yet addressed this issue. However, it is expected that the data will be made available for re-use after the conclusion of the project. It is not clear if licensing will be applied nor if an embargo period will be needed.

2.1.3 Resources

At the moment, and based on the assumptions and plan above, the costs for making the FLAIR laboratory dataset FAIR will be covered by the regular testing activities of the project (WP5 and WP6) and by project management (WP1, where the responsibility for data management lies). Data management responsibility lies with the project coordinator through WP1.
The consortium has not analysed or estimated the costs and/or potential benefits of long term preservation of the FLAIR laboratory dataset.

2.1.4 Security

The FLAIR laboratory dataset will not include sensitive data. The dataset will be stored in the project repository which is hosted in a server of the project coordinator's IT infrastructure. The repository supports version control which should be enough to ensure data recovery in case

of accidental deletions. Data back-ups will be done according to the internal IT policy of the project coordinator as applicable to all other relevant digital data of the company. Access to the data will only be possible through authenticated access to the repository: one account per partner and one account per external individual (i.e. not belonging to the consortium) requesting access to the data – see 2.1.2.2 above.

2.1.5 Ethical aspects

Not applicable.

2.2 FLAIR field test dataset

2.2.1 Summary

2.2.1.1 Purpose of the data collection/generation

The FLAIR sensor system will be operated at the sensor test facility on the roof of the suburban air quality monitoring station on the premise of EMPA in Dübendorf, Switzerland. It will be run for several days next to the inlet of high precision reference instruments. These parallel measurements are used for the determination of the sensitivity and the measurement uncertainties of the FLAIR sensor system for the four target gas species (CO_2, CH_4, CO and O_3) under real world conditions.

2.2.1.2 Relation to the objectives of the project

The FLAIR field test dataset is fundamental for achieving the project core objectives. It will be through this dataset that the design of the FLAIR sensor and the prototype of the sensor and its components will be evaluated and characterized under real world conditions.

2.2.1.3 Data types and formats

The data comprising the FLAIR field test dataset will consist of:

- Trace gas concentrations measured by the FLAIR instrument in ppbv (parts per billion volume) for CO_2, CH_4, CO and O_3
- Trace gas concentrations measured by the Dübendorf air quality monitoring station in ppbv (parts per billion volume) for CO_2, CH_4, CO and O_3
- Absolute difference between the measurements of previous bullets
- Relative difference between the measurements of first two bullets

The consortium does not foresee the inclusion of the data processing algorithm details in the dataset as this can be commercially exploited by the project partners.

2.2.1.4 Re-use of existing data

No existing data will be re-used in the generation of the FLAIR field test dataset.

2.2.1.5 Data origin

The data for the FLAIR field test dataset will be generated at the Dübendorf air quality monitoring stations in Switzerland by researchers directly involved in the project. The measurement process will be described in detail in the appropriate deliverable and associated with the dataset.

2.2.1.6 Expected size of the data

At the moment the expected size of the dataset is not known.

2.2.1.7 Data utility

The FLAIR field test dataset will be useful to the entire consortium as well as for other researchers wishing to evolve the work of FLAIR and potential customers interested in new sensors based on the FLAIR technology.

2.2.2 FAIR

2.2.2.1 Making data findable

The FLAIR field test dataset will be made discoverable through the association of metadata to the dataset. At the moment the consortium expects to associate the following metadata to the dataset:

- Date of measurement
- Gases measured
- Time of measurement

Files in the dataset will be clearly named and their names will include the date of measurement and gas to which the file refers to.

Versioning may be achieved through the addition of a version suffix to the filename or by supporting versioning in the FLAIR repository. The chosen solution is not defined at the moment although it is already clear that the repository will support versioning.

2.2.2.2 Making data openly accessible

At the moment of writing, the consortium expects to make the entire FLAIR laboratory dataset openly available. The consortium expects to make the dataset available through the project's repository.

Registration to the repository will be required and will consist in providing the name, entity and reason of interest or foreseen purpose of use for the FLAIR dataset.

2.2.2.3 Making data interoperable

FLAIR will provide the data in the FLAIR field test dataset in the standard units for detected concentration. This should be enough to ensure the interoperability of the data. Differences will be provided in concentration units (ppbv) for absolute differences or percentages for relative differences.

2.2.2.4 Increasing data re-use

At the moment of writing the consortium has not yet addressed this issue.

2.2.3 Resources

The costs of making and maintaining the dataset FAIR will be covered by the regular testing activities of the project (WP5 and WP6). Data management responsibility lies with the project coordinator through WP1.
The consortium has not analysed the costs and/or potential benefits of long term preservation of this dataset.

2.2.4 Security

The FLAIR field test dataset will not include sensitive data. The dataset will be stored in the project repository which is hosted in a server of the project coordinator's IT infrastructure. The repository supports version control. Data back-ups will be done according to the internal IT policy of the project coordinator. Access to the data will only be possible through authenticated access to the repository.

2.2.5 Ethical aspects

Not applicable

2.3 FLAIR UAV dataset

2.3.1 Summary

2.3.1.1 Purpose of the data collection/generation

The FLAIR UAV dataset will comprise the following:

- Set of measurements of trace gas absorption spectra and minimum detected concentrations performed by the FLAIR sensor equipped UAV with profiling in
  - The vertical direction
  - The radius direction
- Measurements of concentration of gas obtained from the tall tower atmospheric research site in Beromünster (Switzerland);
- Number of particles measured near emission sources (e.g. the motorway in the area of the NABEL site Härkingen in Switzerland)
- Particle concentration measured near emission sources (e.g. the motorway in the area of the NABEL site Härkingen in Switzerland)

The purpose of this dataset is to prove the suitability of the airborne FLAIR sensor system for vertical profiling of atmospheric trace gases and for airborne atmospheric measurements in general. It will also serve the purpose of helping to interpret the measured spatial variation of gases that are emitted by road traffic emissions (or influenced by road traffic emissions like O_3) and will help to demonstrate the usefulness of the FLAIR system for other applications.

2.3.1.2 Relation to the objectives of the project

The FLAIR UAV dataset contributes directly to the achievement of the project core objectives. The vertical mapping of CO_2, CH_4 and CO at the tall tower atmospheric research site in Beromünster (Switzerland), where precise measurements of these atmospheric trace gases are available at different heights will be compared to the vertical profiles of CO_2, CH_4 and CO obtained by the UAV flying the FLAIR sensor by averaging the measurements obtained at different altitudes. This comparison will confirm or reject the suitability of the airborne FLAIR

sensor system for vertical profiling of atmospheric trace gases and for airborne atmospheric measurements in general.
The particle sensor measurements will be compared with the concentration measurements obtained from the FLAIR sensor at a different location and will contribute to demonstrate the usefulness of the FLAIR system for applications where toxic gases and particle emission occur simultaneously (such as in road traffic).

2.3.1.3 Data types and formats

The data comprising the FLAIR UAV dataset will consist of:

- FLAIR sensor trace gas concentrations in ppbv (parts per billion volume)
  - against height in meters
  - against horizontal distance from tower in meters
- FLAIR sensor gas absorption spectra in the 2-5 μm and 8-12 μm windows as a set of absorbance per wavelength measurements
  - against height in meters
  - against horizontal distance from tower in meters
- Beromünster tower trace gas concentrations in ppbv (parts per billion volume) against height in meters.
- Total number of particles measured in the UAV at Härkingen (site to be confirmed) against height in meters and horizontal distance in meters.
- Particle concentration measured in the UAV at Härkingen (site to be confirmed) in total particles per ccm against height in meters and horizontal distance in meters

At the moment of writing, the consortium plans to generate the data above for the following gases:

- CO_2 – Carbon Dioxide
- CH_4 – Methane
- CO – Carbon Monoxide

The consortium does not foresee the inclusion of the data processing algorithm details in the dataset as this can be commercially exploited by the project partners.

2.3.1.4 Re-use of existing data

No existing data will be re-used in the generation of the FLAIR UAV dataset.

2.3.1.5 Data origin

The data for the FLAIR UAV dataset will be generated at the locations of the flights (to be confirmed later during the project) directly by the FLAIR consortium and the operation of the UAV. The generation procedures will be described in detail in the appropriate deliverables and associated with the dataset.

2.3.1.6 Expected size of the data

At the moment the expected size of the dataset is not known.

2.3.1.7 Data utility

The FLAIR UAV dataset will be useful to the entire consortium (namely in helping accomplish the project objectives as outline above) as well as for other researchers wishing to evolve the

work of FLAIR and potential customers interested in new sensors based on the FLAIR technology.

2.3.2 FAIR

2.3.2.1 Making data findable

The FLAIR UAV dataset will be made discoverable through the association of metadata to the dataset. At the moment of writing the type of metadata and identification mechanism to be applied is not yet defined. The process or standard to be used to create the metadata is not clear yet. However, the consortium expects to associate the following metadata to the dataset:

- Date of measurement
- Gases measured
- Absorption windows
- Time of measurement
- Location of measurement

Files in the dataset will be clearly named. All files in the dataset will allow the clear identification of the version. This may be achieved through the addition of a version suffix to the filename or by supporting versioning in the FLAIR repository. The chosen solution is not defined at the moment.

2.3.2.2 Making data openly accessible

At the moment of writing, the consortium expects to make the entire FLAIR UAV dataset openly available. The consortium expects to make the dataset available through the project's repository (which is foreseen to support versioning). The repository is maintained by the project coordinator and access to it is authenticated. Access to the repository will be enabled through a web interface that only allows download of the dataset (i.e. it will not be possible to delete, upload, check-out or commit other files).

Registration to the repository will be required and will consist in providing the name, entity and reason of interest or foreseen purpose of use for the FLAIR UAV dataset.

2.3.2.3 Making data interoperable

FLAIR will provide the data in the FLAIR UAV dataset in the standard units as described above in section 2.3.1.1.

2.3.2.4 Increasing data re-use

At the moment of writing the consortium has not yet addressed this issue.

2.3.3 Resources

The costs of making and maintaining the FLAIR UAV dataset FAIR will be covered by the regular testing activities of the project (WP5 and WP6). Data management responsibility lies with the project coordinator through WP1.
The consortium has not analysed or estimated the costs and/or potential benefits of long term preservation of the FLAIR UAV dataset.

2.3.4 Security

The FLAIR UAV dataset will not include sensitive data. The dataset will be stored in the project repository which is hosted in a server of the project coordinator's IT infrastructure. The repository supports version control. Data back-ups will be done according to the internal IT policy of the project coordinator. Access to the data will only be possible through authenticated access to the repository.

2.3.5 Ethical aspects

Not applicable

3 Other data

3.1 HITRAN database

3.1.1 Summary

The FLAIR consortium will collect data from the HITRAN (High resolution TRANsmission molecular absorption) database that will be used to benchmark and characterize the FLAIR laboratory performance. HITRAN is a compilation of spectroscopic parameters that a variety of computer codes use to predict and simulate the transmission and emission of light in the atmosphere. The goal of HITRAN is to have a self-consistent set of parameters. The database is a long-running project started by the Air Force Cambridge Research Laboratories (AFCRL) in the late 1960s in response to the need for detailed knowledge of the infrared properties of the atmosphere.

The initial HITRAN database included only the basic parameters necessary to solve the Lambert-Beers law of transmission. In addition, the air-broadened Lorentz width was included as well as the unique quantum identifications of the upper and lower states of each transition.

The parameters stored in HITRAN are a mixture of calculated and experimental. HITRAN provides the sources for the key parameters within each transition record whereby the user can determine from where the value came. The experimental data that enter HITRAN often come from the results of analysis of Fourier transform spectrometer laboratory experiments. Many other experimental data also are inputted, including lab results from tuneable-diode lasers, cavity-ring down spectroscopy, heterodyne lasers, etc. The results usually go through elaborate fitting procedures. The theoretical inputs include standard solutions of Hamiltonians, ab initio calculations, and semi-empirical fits.

The HITRAN parameters will be compared to the data obtained from the FLAIR sensor in a lab environment. This way, the consortium will be able to characterize the performance of the FLAIR sensor and this will help establish the validity of the FLAIR design and the prototype.

HITRAN parameters can be obtained directly from the HITRAN database website (http://hitran.org) and are output in records of 160 characters which include parameters in integer, real and text data types. The data types and formats can be consulted at https://www.cfa.harvard.edu/hitran/formats.html.

At the moment of writing the size of the data to be used by FLAIR is not known. FLAIR will not store HITRAN data nor will it make the data available as it can be freely obtained directly from HITRAN. It is mentioned here as it will have an important role in the validation of FLAIR results.

Whenever HITRAN data is used the research results will clearly identify which parameters have been used.

3.2 PNNL spectral library

3.2.1 Summary

The Pacific Northwest National Laboratory (PNNL) has created a quantitative database containing the vapour-phase infrared spectra of pure chemicals. The digital database has been created with both laboratory and remote-sensing applications in mind.

This unique library is the "gold standard" of vapour phase infrared reference spectra and is uniquely adapted to both remote and point sensing. The library is a unique asset to the development of infrared sensors and methods. Data from the PNNL spectral library will be used in FLAIR as benchmark against which the FLAIR sensor's performance will be measured. This way, the consortium will be able to characterize the performance of the FLAIR sensor and this will help establish the validity of the FLAIR design and the prototype.

PNNL parameters can be obtained directly from the PNNL library website (https://secure2.pnl.gov/nsd/nsd.nsf/Welcome) via registration.

At the moment of writing the size of the data to be used by FLAIR is not known. FLAIR will not make PNNL data available as it can be freely obtained directly from the PNNL library. Whenever PNNL data is used the research results will clearly identify which parameters have been used.

3.3 UAV performance data

3.3.1 Summary

The FLAIR sensor will be installed on a modified Unmanned Aerial System based on an existing product from TEKEVER Autonomous Systems. While the consortium will gather and collect performance data for the FLAIR-equipped UAV for the purpose of characterizing the performance of the system while in flight and to validate the sensor application, specific UAV performance data will not be made available outside the consortium.

Specific UAV sub-systems data is proprietary and can be used to infer the solutions and specific components used in the product. As the UAV is a commercial product, open access to this data could result in the loss of competitive advantage in the commercial exploitation of the product. Any data which is considered relevant for the characterization and validation of the FLAIR sensor itself while installed in the UAV (e.g. altitude, GPS position, etc.) will be made available through the FLAIR UAV dataset.

3.4 FLAIR system design data

3.4.1 Summary

The FLAIR sensor system is described in detail in deliverable D2.3 – List of specifications doe all system (UAV and assembled sensor) and sub-system parameters. Designs, blueprints for hardware, electronic circuit design and technical data for the sub-systems will be generated throughout the project and will be made available in corresponding deliverables, most of which are of public dissemination. However, it must be noted that the consortium may decide, as the

project progresses that some specific design details be kept confidential as they provide the added value or the competitive advantage that will allow the generating partner(s) to protect and exploit the results. Examples of these include design and test data related to:

- Data processing algorithm due to its commercial potential;
- Subsystem development;
- UAV adaptation;
- And Sensor/UAV fitting.

Furthermore, while all of the physical components of the FLAIR sensor are public, open access to the HW and prototypes will not be ensured by the consortium. Rather, a description of the units will be published but the unit itself will not be public as it is needed in the FLAIR drone system.

References

FLAIR consortium, *Grant Agreement Number: 732968 – FLAIR*, Annex 1, Version of 14 of March 2017.

FLAIR consortium, *D2.1 – Requirements for FLAIR Sensor System*, Version 1.1, March 2017.

European Commission, *Guidelines on FAIR Data Management in Horizon 2020*, Version 3.0, July 2016.

Jones, S., *How to Develop a Data Management and Sharing Plan*, A Digital Curation Centre (DCC) 'working level' guide, 2011.

European Research Council, *Guidelines on Implementation of Open Access to Scientific Publications and Research Data in projects supported by the European Research Council under Horizon 2020*, Version 1.1, April 2017.

Atomic and Molecular Physics Division, Harvard-Smithsonian Center for Astrophysics, *The HITRAN Database*, http://hitran.org/about/, last visited on 9 May 2017.

Pacific Northwest National Laboratory, *Northwest-Infrared Spectral Library*, https://secure2.pnl.gov/nsd/nsd.nsf/Welcome, last visited on 9 May 2017.