



D7.2

Data Management Plan (DMP)

Project number:	700002
Project acronym:	ALFA
Project title:	Advanced Low Flying Aircrafts Detection and Tracking
Start date of the project:	1 st January, 2017
Duration:	36 months
Programme:	H2020-BES-2015

Deliverable type:	Report
Deliverable reference number:	BES-700002 / D7.2 / 1.0
Work package contributing to the deliverable:	WP7
Due date:	JUN 2017 – M06
Actual submission date:	3 rd July, 2017

Responsible organisation:	TUBS
Editor:	Jörg Schöbel
Dissemination level:	PU
Revision:	1.0

Abstract:	This document constitutes the Data Management Plan (DMP) of the ALFA project, explaining how the project plans to manage research data.
Keywords:	Data Management



The ALFA project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 700002.

Editor

Jörg Schöbel (TUBS)

Contributors (beneficiaries in alphabetical order)

Jaime Martín-Pérez, (ATOS)

Giuseppe Francaviglia, Massimiliano Camilli (ENG)

Paulo Chaves (INOV)

Barbara Gaggl, Martin Deutschmann, Oleksandr Tomashchuk (TEC)

Gilles Premel Cabic (TNL)

Rob van Heijster (TNO)

Markus Krückemeier, Fabian Schwartau (TUBS)

Disclaimer

The information in this document is provided “as is”, and no guarantee or warranty is given that the information is fit for any particular purpose. The content of this document reflects only the author’s view – the European Commission is not responsible for any use that may be made of the information it contains. The users thereof use the information at their sole risk and liability.

Executive Summary

This document constitutes the Data Management Plan (DMP) of the ALFA project, describing the data management life cycle for the data to be collected, processed and/or generated by a Horizon 2020 project. The overall aims are to make research data findable, accessible, interoperable and re-usable (FAIR).

This first version of the DMP mainly focuses on a synopsis of data management in ALFA in a top-down approach. After describing the ALFA system information flow and giving a general account of ALFA sensor data, the guidelines and provisions for making data FAIR are described. The final parts cover sharing, archiving and preservation of data, as well as data security.

This document has been compiled as a summary of a questionnaire-based survey, according to the *EC Guidelines on FAIR Data Management in Horizon 2020*. The survey was answered by all the partners of the ALFA consortium involved in technical or software R&D until May 2017.

The concept of FAIR research data finds its main application in certain fields of (basic) science, where the acquisition of research data may be the result of a scientific endeavour on its own, such as genome sequencing, generation of earth observation data etc. Therefore, the data resulting from such endeavours have a high value for other researchers. In contrast, in sensor and systems R&D as conducted in ALFA, the key result of the development is the sensor or the sensing system, but not the data acquired with this system. As a result, the overall approach of FAIR data does only fit partially to the nature of R&D conducted in ALFA, which is closely reflected in the ALFA partners' responses to the questionnaire. Notwithstanding, the ALFA partners will adapt concepts of FAIR data management wherever possible and appropriate.

Concerning research data that is used in scientific publications by ALFA partners, each partner owning such data is responsible for long-term storage and retrieval, as well as for making data accessible to the scientific community, preferentially on an individual basis ("by request"). As the topic of ALFA is a security system, data protection and security provisions may prevent sharing of data and of information thereof.

Extending the context of "research data", this document also conceptually describes the ALFA system data flow and the exchange of data between the sensors and data processing units of the system. However, at the current state of the project, definitions of the exact content of the to-be-exchanged data, data interfaces, structures, and formats have not been finalized.

Since it is early in the project, little is known in terms of sharing, volume and archiving of data. Data that is of interest to a larger number of consortium members may be shared and archived within the project SVN. However, more detailed concepts of information exchange and data preservation have not yet been defined. The project is aware of these aspects and will tackle them by updating the present document during the specifications and conceptual development of the sensors and data processing units of the system. Therefore, information in this document is subject to change and updates will be included in the periodic reports or during the project lifetime whenever necessary.

Input from WP(s):	not applicable
Output to WP(s):	not applicable
Deliverables:	D7.2
Duration:	January 1, 2017 to June 30, 2017 (for first issue of the DMP)
WP leader:	TUBS
Participants:	ATOS, ENG, INOV, TEC, TNL, TNO, TUBS

Contents

Chapter 1	Introduction	1
Chapter 2	Methodology	2
2.1	Structure of the Questionnaire	2
2.2	Information on Data Sets, Interfaces and Structures.....	3
2.2.1	ALFA Operative System Data.....	3
2.2.2	Research Data.....	3
Chapter 3	Synopsis of Data Management in ALFA.....	4
3.1	Data Summary	4
3.1.1	ALFA system information flow	4
3.1.2	ALFA sensor data.....	5
3.1.3	ALFA system data model.....	6
3.1.4	Ownership of data	6
3.2	FAIR Data	7
3.2.1	Making data findable, including provisions for metadata	7
3.2.2	Making data openly accessible	7
3.2.3	Making data interoperable	8
3.2.4	Increase data re-use (through clarifying licences).....	8
3.3	Accessibility – Data Sharing, Archiving and Preservation	8
3.3.1	Access to research data for validation of scientific publications	8
3.3.2	Sharing, archiving, and preservation of data	9
3.4	Data Security.....	9
3.4.1	ALFA system operation	10
3.4.2	ALFA research data.....	10
3.5	Ethical Aspects.....	10
3.6	Other Issues.....	10
Chapter 4	Summary and Conclusions	11
Chapter 5	List of Abbreviations.....	12
Chapter 6	Bibliography	13

List of Figures

- Figure 1: Information flow in the ALFA system. Arrows and elements depicted in grey colour are optional and have not been defined or decided upon by the time of writing of this document. .4
- Figure 2: Information flow and interdependencies in the ALFA sensors. Arrows and elements depicted in grey colour are optional and have not been defined or decided upon by the time of writing of this document..... 6

Chapter 1 Introduction

The H2020 programme has implemented a pilot action on open access to research data. Although ALFA is not formally participating in this pilot action, the consortium has pledged to develop a **Data Management Plan (DMP)** in order to document sources and content of data generated by the project partners, as well as the management, availability, handling and storage of this data. This DMP has been identified in the Description of Action (DoA) as ALFA deliverable D7.2. This document is drafted according to the *Horizon 2020 FAIR Data Management Plan template*. The major aim of the DMP is to ensure that our research data is FAIR – **F**indable, **A**ccessible, **I**nteroperable, and **R**e-usable.

The DMP is intended to be a **living document**. It will be periodically revised to reflect changes in the data that may be made available by the project, and to provide additional information on the datasets as this information is developed during the specifications of the experimental phases.

This first version of the DMP has been written at a rather early stage of the project. At the same time, partners are developing a mutual understanding of the ALFA system's details and functionalities and interaction of the subsystems and sensors in WP1. Therefore, an overview of the system and its data is given in this document, as well as general guidelines and processes of the handling of research data as well as operative system data in ALFA. However, detailed information on the actual format and content of data sets generated by sensors or transmitted between system units will be subject of later versions of the DMP.

All partners have contributed to the document, particularly through the use of a project-wide questionnaire. The methodology behind this approach is described in Chapter 2. A synopsis of these inputs with a description of the overall ALFA project perspective is provided in Chapter 3 with general conclusions in Chapter 4.

In future versions of the DMP, an index of ALFA data sets will be compiled that describes particular data generated in the project. Whereas the DMP's dissemination level is public, the index will be confidential with access limited to the consortium, EU project officer and reviewers.

Chapter 2 Methodology

In order to compile the data management plan, a questionnaire was first elaborated, covering the main questions that need to be answered in the template provided by the European Commission.

In a second phase, each project partner responded to the questionnaire, filling it with as much detail as possible at this stage of the project. Completed questionnaires were stored for analysis and traceability in the project's SVN repository.

In a third phase, the Data Management Plan was created as a synthesis of the questionnaire results, attempting to take advantage of communalities between responses in order to provide a concise view of data management procedures within the consortium.

Further revisions of the document will be based on updates to partner questionnaires. Therefore, the DMP will be updated at least by the mid-term and final review to be able to fine-tune it to the data generated and the uses identified by the consortium.

2.1 Structure of the Questionnaire

The questionnaire is structured following a **6-points structure** adapted from the *H2020 Guidelines on FAIR Data Management*:

1. **Data summary** gives a description of the data.
 - 1.1. **Data Management:** Collection/Generation/Documentation of Data & Metadata
2. **FAIR data** ensures that data are **F**indable, **A**ccessible, **I**nteroperable, and **R**e-usable.
 - 2.1. **Making data findable, including provisions for metadata** gives a description of the data.
 - 2.2. **Making data openly accessible** explains how the data are made openly accessible.
 - 2.3. **Making data interoperable** assesses the interoperability of the data.
 - 2.4. **Increase data re-use (through clarifying licences)** discusses how the re-usability of the data is ensured.
 - 2.4.1. Intellectual Property Rights
3. **Accessibility** – Data sharing, archiving and preservation
4. **Data security** addresses data recovery as well as secure storage and transfer of sensitive data.
5. **Ethical aspects** clarifies if there any ethical or legal issues that can have an impact on data sharing.
6. **Other issues** refer to other national/funder/sectorial/departmental procedures for data management and security are made use of.

The questionnaire is based on the *H2020 Guidelines on FAIR Data Management* and uses basically the same structure and numbering of sections. However, we have extended this structure to introduce new sections (1.1., 2.4.1. and 3.), in which information is compiled, which would otherwise be somewhat scattered over the other parts. This provides a concise description of Data and Metadata in section 1. In addition, IPR is specifically addressed in section 2.4.1.

The original section 3 of the template (*H2020 Guidelines on FAIR Data Management*) has also been adapted. In the new section 3, a more specific focus on data sharing and archiving is provided. Deviating from the template, it appears unnecessary to collect information on resource

allocation for making data FAIR, since ALFA does not take part in the Pilot on Open Research Data. Therefore, there will be no specific and coordinated efforts to prepare and annotate data for external re-use, although partners may individually take action to ensure FAIR data, wherever appropriate. Also, research data used in publications will be made available upon request, if no prohibitive circumstances (e.g. security provisions or IPR protection) apply.

In section 6 we specifically collect information on data management and security procedures, which may be recommended or obligatory, and may be defined e.g., by national regulations, funding agencies or within the partners' institutions.

2.2 Information on Data Sets, Interfaces and Structures

2.2.1 ALFA Operative System Data

When the project will have finished its phase of definition and conceptual work, a clear view of data sets, interfaces and data structures in the operative ALFA system will be formed. An **index of datasets** will be created and maintained in the project as a reference document for the ALFA partners. This information will be made available to the ALFA partners in conjunction with the DMP, as far as it is not subject of specific deliverables (in this case, the dataset index will not contain any detailed information but will refer to the deliverables).

The DMP itself is defined as a public document, because its main focus should be on FAIR research data availability and exchange. However, the information of the index on data structures and interfaces as well as on detailed communication structures within the ALFA system go far beyond availability and use of research data, therefore this index will be confidential and information about the datasets exchanged by the sensors and system units will only be provided to the EU and the reviewers.

2.2.2 Research Data

Research data will be acquired during the R&D process. These data do not belong to the ALFA system data flow, for example sensor raw data acquired during the testing and optimization phase of sensors. Information on research data will be included in the dataset index.

The ALFA project will consider open licenses and open availability for any research data, as long as no security provision or other reasons for data protection apply. The reasons for not offering open access will be documented in the partner questionnaires and in the dataset descriptions in the index of datasets.

In order to make any public research data findable, information on public datasets will also be made available publicly in the context of publications or in the DMP. Zenodo (<https://zenodo.org/>), which will be described in more detail below will be used as a suitable repository to grant open access.

Chapter 3 Synopsis of Data Management in ALFA

3.1 Data Summary

3.1.1 ALFA system information flow

The information flow in the ALFA system from a global perspective is depicted in Figure 1. Arrows and elements depicted in grey colour are optional and have not been defined or decided upon by the time of this writing.

Three sensor systems will deliver information about detected objects. This information may be complemented with annotations and metadata, which will be required for information processing in the sensor or which will be passed to the data fusion unit.

Annotations and metadata may comprise the following information:

- system-generated information, such as time stamps, GPS coordinates, etc.;
- user input necessary for signal and data processing, e.g. the orientation of the sensing unit with respect to some reference (global North) or information about the illuminator of opportunity (passive radar);
- user input which will only be used to annotate data to be saved for debugging, analysis or performance assessment of a sensor (e.g. specific descriptions on target, scenario etc.).

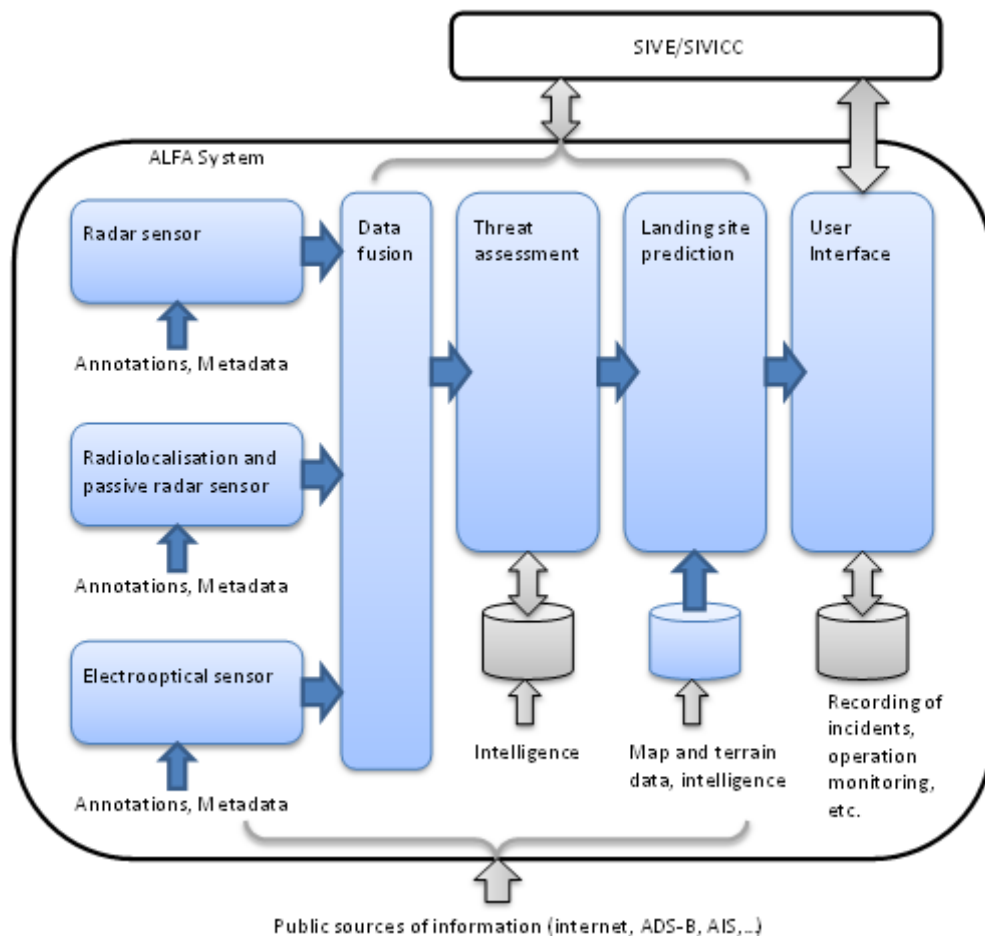


Figure 1: Information flow in the ALFA system. Arrows and elements depicted in grey colour are optional and have not been defined or decided upon by the time of writing of this document.

The information interfaces between the sensors, the data fusion, and between all successive units will be defined within the ALFA system data model. Work on architecture and the data model will be subject of task 2.3 of WP2 (Months 09-12) and results will be found in D2.2.

The ALFA system will detect incoming drug aircraft in time. The aircraft will be located, classified as drug smuggler and tracked. Behavioural analysis will predict its landing zone. Based on characteristics of the craft (size and the range it can expectedly fly), a number of most probable landing sites will be indicated.

As the architectural work is ongoing, it has not yet been defined, to which extent the information processing will require external data or will include their own repositories of data or other intelligence (depicted as cylinders in Figure 1). The landing site prediction unit will certainly use some database of map and terrain data.

The final unit in the information processing chain is the user interface. The operator is assisted in all phases of the operation by easy- to- read screens on his or her mobile device. Locations to move to, as well as the best route, are shown in a navigation pane. Target location and potential landing site are also shown. Recording of incidents and operation monitoring, as well as the specific conceptual links to the SIVE/SIVICC systems, are to be defined in WP1 and WP2.

ALFA system output in terms of information on (potential) incidents and system operations monitoring are considered confidential and will only be made available to the consortium and to end users (particularly, security and border protection forces).

3.1.2 ALFA sensor data

Figure 2 shows the information flow, interdependencies, and information generation in the ALFA sensors. The sensor raw data will be processed to acquire information on basic observations of the objects (e.g. velocity, range) and at the same time generated object hypotheses. Some of the basic processing, however, requires an initial object hypothesis (e.g. angle of arrival processing in a radar sensor), so that basic signal processing and object acquisition may be strongly linked. Finally, the detection results are represented by objects, which will be tracked within the sensor. Tracking and generation of object hypotheses are also strongly linked, given that an object hypothesis is typically confirmed when it has been tracked over a certain period.

The specification of sensor output in terms of information content and data format are subject of the system data model (task 2.3 of WP2). Annotations and metadata will be provided to support data fusion and further steps of information processing. In operation, the sensor will typically not retain or store any data once these data have been transferred to the information fusion unit.

The sensor may use specific intelligence, such as micro-Doppler, radar, or RF signatures, shapes and outlines of objects for optical identification, etc. to classify or identify a potential target. This intelligence will be stored securely and will only be accessible by the respective sensor. As the generation of such intelligence may require a considerable effort and its content may provide information on sensing and classification capabilities, such data is considered confidential and will not be made externally available.

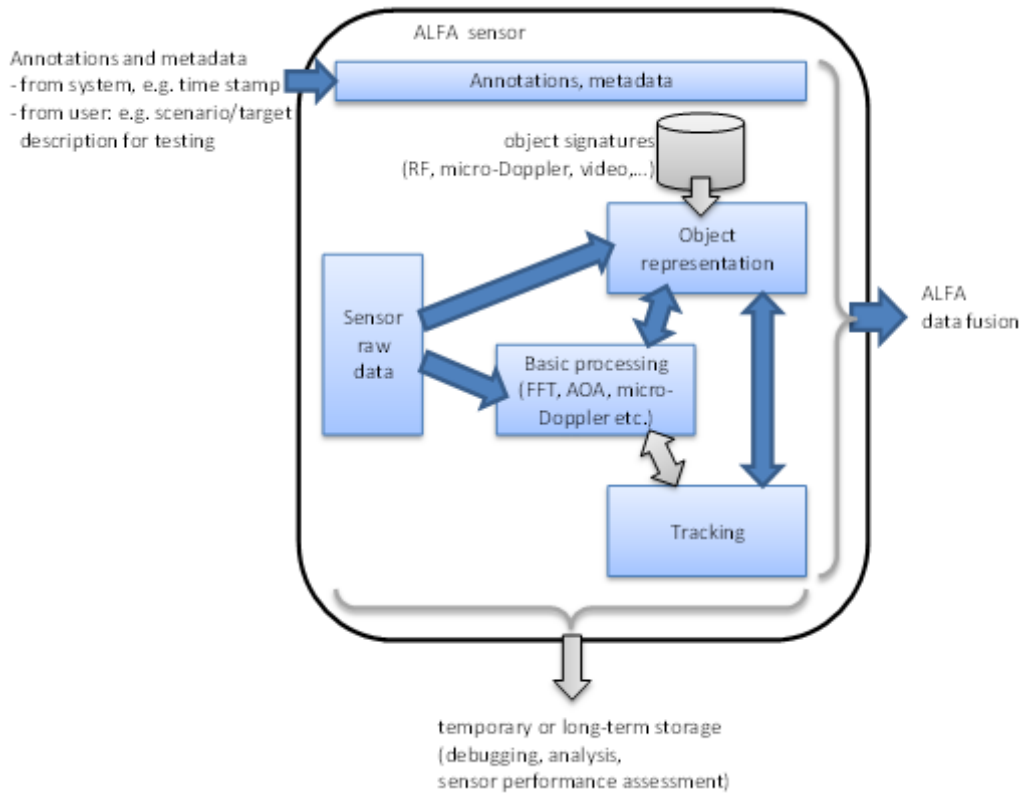


Figure 2: Information flow and interdependencies in the ALFA sensors. Arrows and elements depicted in grey colour are optional and have not been defined or decided upon by the time of writing of this document.

Metadata may contain additional information, which may only be used for testing and performance assessment (e.g. scenario and target descriptions).

During testing, debugging and performance assessment, a sensor may be configured to store data at any step of the signal processing or any level of abstraction (from raw data to object-level data). These data may in principle be useful for other researchers. As an example, radar or passive radar raw data may be used to test new concepts and algorithms for radar signal processing. To be re-useable such data has to be annotated with target and scenario description. In addition, a description of interfaces with measurement and digital processing setup has to be provided. Sharing of such data is foreseen in the context of publications, see section 3.3.1.

3.1.3 ALFA system data model

Work on architecture and the data model will be subject of task 2.3 (months 09-12) and results will be found in D2.2 “ALFA data model description”.

3.1.4 Ownership of data

Ownership of results and related rights are defined in the consortium agreement (CA). The ownership of data and rights to decide on the distribution of data can be classified into three categories:

- (1) Individual partners’ data: mainly sensor raw data and background information. The partners will own these data individually and decide on the distribution.
- (2) ALFA shared data: data that have been generated using input of several partners or data that are needed by several partners (mainly ALFA data model, database of landing sites, directory of

RF emissions and transmitters of opportunity). If necessary (i.e. cases not covered by the CA), distribution and access procedures will be defined individually by the partners contributing to the generation of such data, at the latest when these data become available in the ALFA system.

(3) The ALFA system data flow (corresponding to the system data model): sensor output, fused data output, ALFA system output to end users. Distribution and access procedures will be defined by the ALFA consortium as far as not already covered by the CA.

3.2 FAIR Data

Making data FAIR ensures that data are Findable, Accessible, Interoperable, and Re-usable.

3.2.1 *Making data findable, including provisions for metadata*

When data shall be made available to the scientific community, these data need to be discoverable and identifiable for potential users. If possible, standard identification mechanisms (e.g. persistent and unique identifiers such as Digital Object Identifiers) should be used.

Discoverability and identification also apply to data recorded by a single partner in ALFA for his own use and not intended for the public, e.g. in sensor performance assessment.

Partners will adopt naming conventions focusing on “telling names”, i.e. relevant information (date, location, reason or occasion of the measurement, etc.) will be coded in the file names and in the organisation and naming of directory structures.

Whenever research data are relevant for the consortium and will be used in cooperation between several partners, data storage in the project SVN will be considered. Data that are individually owned and individually used will preferentially be stored by the owning partner.

Metadata will be provided preferably directly together with the data (e.g. in the same directory or directory substructure). Metadata may comprise information on the specific sensor setup, measurement arrangement, measurement occasion, target and scenario description, specific remarks, etc. As there are no preferential standards for such metadata, free text in text or word processor files (.txt, .doc, .pdf) will be preferred. Such metadata will, at the same time, also provide sufficient documentation for any data used in scientific publications, therefore following the rules of good scientific practice [ALLEA, 2017, sec. 2.5].

3.2.2 *Making data openly accessible*

Open access to research data encourages not only transparency and quality assurance as they become reproducible, but it increases efficiency and saves costs by making it possible to conduct secondary analyses. It enhances a barrier-free knowledge transfer and contributes to build a research union.

At the time of writing, the ALFA coordinator plans to make quality data openly accessible. To guarantee open access to scientific publications and research data, Zenodo was selected as a suitable repository. It is convenient to access and also easy to use. This repository allows to easily share the long tail of small research results in a wide variety of formats including text, spreadsheets, audio, video, and images across all fields of science. Further, each uploaded publication and dataset receives a persistent identifier (DOI), which ensures the long term preservation. Besides that, Zenodo is an EU compliant repository.

In contrast, much of the ALFA sensor data and information generated by the ALFA system will be considered confidential for security reasons. It is not only that criminals could use data or intelligence generated or used by the ALFA system or its subsystems (e.g. target/non-target signatures) to assess the performance of the detection system and devise techniques of cloaking

or faking the appearance as a non-target. It is also a matter of protecting ALFA partners' individual know-how from being exploited by their direct competition.

Concerning the access to research data used in scientific publications see 3.3.1.

3.2.3 Making data interoperable

Within the ALFA system, interoperability of data is a key element of the information flow between the processing units (Figure 1). The architecture is subject of the ALFA system data model (see above and see D2.2 when available).

As far as data will be provided to internal or external partners (e.g. in the case of publications, see 3.3.1), no specific provisions will be made by the provider to make data interoperable, e.g. by converting them into specific formats etc. However, data format and interface descriptions will be provided, so that a user can read and interpret the data accordingly.

3.2.4 Increase data re-use (through clarifying licences)

At the time of writing, data re-use is devised to be mostly restricted to individual partners' own re-use or to re-use within the ALFA consortium. Sensor data will be re-used to optimize signal processing or other steps in the ALFA processing chain.

Referring to external re-use of data: see 3.2.2 and 3.3.1.

3.3 Accessibility – Data Sharing, Archiving and Preservation

3.3.1 Access to research data for validation of scientific publications

The concept of FAIR data in the context of the Open Research Data Pilot as stated in [EC, 2016] “applies primarily to the data needed to validate the results presented in scientific publications. Other data can also be provided by the beneficiaries on a voluntary basis, as stated in their Data Management Plans.”

Numerous research institutions (Global Research Council, European Science Foundation, etc.), funding agencies, academic foundations (e.g. German Academic Foundation, DFG) etc. have issued codes of conduct for good scientific practice. As these documents exhibit many similarities, we have chosen [ALLEA, 2017] as a representative reference.

ALFA partners will adhere to good scientific practice. In the context of data management the guidelines in sec. 2.5 of [ALLEA, 2017] apply. For easy reference, these guidelines are cited in the following:

- *Researchers, research institutions and organisations ensure appropriate stewardship and curation of all data and research materials, including unpublished ones, with secure preservation for a reasonable period.*
- *Researchers, research institutions and organisations ensure access to data is as open as possible, as closed as necessary, and where appropriate in line with the FAIR Principles (Findable, Accessible, Interoperable and Re-usable) for data management.*
- *Researchers, research institutions and organisations provide transparency about how to access or make use of their data and research materials.*
- *Researchers, research institutions and organisations acknowledge data as legitimate and citable products of research.*
- *Researchers, research institutions and organisations ensure that any contracts or agreements relating to research outputs include equitable and fair provision for the*

management of their use, ownership, and/or their protection under intellectual property rights.

In contrast to scientific communities focusing on more basic research, research data do not play an as important role in many engineering disciplines, as in sensor and systems engineering. Disciplines, where FAIR data are of essential importance are, for example, genetics, where whole genome sequencing may be considered a scientific undertaking on its own. Other examples are earth observation data or astronomical observation data, which may be essential resources for scientific investigations seeking for specific information in these data.

The R&D procured in ALFA has a different character. Data are typically not a foundation of research in terms of analysing data but the research focuses on the design of data acquisition units (sensors) and directly connected “higher-level” data processing. The data themselves are typically not considered of high value for a larger scientific audience. Therefore, in contrast to genetics etc., there are no well-established public repositories for research data in electrical engineering and similar disciplines.

Within ALFA, research data will be preserved and provided to interested researchers on an individual basis: If a reader of a publication is interested in data, he would normally approach the authors directly via the contact given in the publication. The ALFA partner owning the data will consider providing these data – together with appropriate description of data format and/or data interfaces – as long as no conflict with security provisions, protection of IPR or other conflicting circumstances apply.

3.3.2 Sharing, archiving, and preservation of data

Whenever research data are relevant for the consortium and will be shared among several partners, data storage in the project subversion repository (SVN) will be considered, to which all partners have access. The SVN allows easy synchronization of documents between the server (hosted at the premises of the coordinator TEC) and a participant’s local file storage for sharing documents and data. The system includes tools for retrieving older versions of a particular file, which means that all former versions of a file are available and reproducible. Therefore, no conventional backup system is necessary. Regarding the preservation of the data, there is no expiration date until which the data is available. Even years after the project end the data will still be available. Hence, long term preservation of the data is secured. Due to the fact that the system is set up and maintained by the coordinator in their IT infrastructure in Austria, we have full control over the data at any time, which is a significant advantage over cloud based solutions.

In case of a small number of partners having interest in certain data and/or data volumes being excessive, sharing of data by other means will be considered (e.g. physical exchange of disk drives or solid state drives if the data volumes are in the region of Terabytes).

Data that are individually owned and individually used will preferentially be stored by the owning partner. All partners have individual data storage / server and backup solutions within their institutions.

3.4 Data Security

Data security addresses two different areas. The first area covers secure storage and transfer of sensitive data during the operation of the ALFA system, the second area covers security of research data owned by partners or submitted to the ALFA SVN.

3.4.1 ALFA system operation

During task 2.3 we will define the measures to ensure protection of sensitive data during collection and storage (data model). It will also require the definition of applicable destruction processes in order to observe the principle of retaining data for the minimal necessary period. Generally, during final system operation, ALFA sensors will not (or only minimally) store any data permanently but transfer their detection results to the data fusion unit.

The defined measures will be subject to the EEA (External Ethics Advisor) and SSC (Security Sensitivity Committee) evaluation. Tasks 3.5 and 3.6 of WP3 will enforce the necessary measures for protection of data in rest and in transit in order to minimize the risk of affecting either privacy or loss of sensitive data.

To ensure the protection and efficacy of LEA operations, data from ALFA systems to mobile devices and vice versa will be transmitted using secure communication. Encryption/decryption is implemented by utilizing the software package OpenSSL (<http://www.openssl.org/>). OpenSSL is an open source SSL/TSL toolkit implementation. Specific cryptographic measures will be undertaken to ensure that the data collected and handled by the project is not tampered with during transmission or storage.

Taking into consideration that ALFA may have the necessity to share data with previously unregistered third parties in an efficient way, implementation of techniques such as Attribute Based Encryption and Identity Based Encryption will be implemented in task 3.5. Special care will be taken on logging data accesses in a cryptographic tamper-resistant way, in order to later determine who has accessed to data.

3.4.2 ALFA research data

Part of the research data may be sensitive, as for example data on the detection performance of the ALFA system with respect to certain threats. These data will be considered confidential and appropriate security measures will be taken to protect them. Obviously, these data will not be provided to external researchers, even if operational performance of ALFA were subject of a publication. Other data, which are less sensitive, will not need special protection.

Research data submitted to the ALFA SVN is protected by means of mandatory authorization. Partners, who wish to have access to the SVN (both read only and read/write access), need to obtain user credentials from TEC. A username will be generated and the user has to define a password. Only with these credentials, the user will get access to any documents on the SVN. All communication between the clients and the SVN server use SSL encryption.

To provide sufficient security to centrally stored data, partner institutions operate server rooms or data centres with limited access, climate control, power backup, and appropriate data backup mechanisms. If considered necessary, research data may be encrypted before storing.

3.5 Ethical Aspects

Ethical and legal aspects of the ALFA system and ALFA system operation will be subject of D7.4 “Interim report on societal impact” and D7.7 “Final report on societal impact”.

3.6 Other Issues

The questionnaire also asked for other national/funder/sectorial/departmental procedures for data management that are made use of and that have not been mentioned in the former chapters. By the time of writing there were no such procedures mentioned.

Chapter 4 Summary and Conclusions

The Data Management Plan of ALFA describes the activities of the partners related to datasets and is a key element of good data management. It contains a summary of all the information available as of 1st June, 2017. All partners intend to create or process data and make the results available within the consortium.

The DMP needs to be updated in the course of the project whenever significant changes arise, such as new data, changes in consortium policies (e.g. new innovation potential, decision to file a patent), changes in consortium composition, and external factors (e.g. new consortium members joining or old members leaving). This will be done within the periodic reports in M18 and M36.

This first version of the DMP mainly focuses on a synopsis of data management in ALFA in a top-down approach. After describing the ALFA system information flow and giving a general account of ALFA sensor data, the guidelines and provisions for making data FAIR are described. The final parts cover sharing, archiving and preservation of data, as well as data security.

The concept of FAIR data has a certain focus on genetics, biomedicine and astro and space physics, and similar scientific communities, as can be observed from the introductory sections of [Wilkinson, 2016]. In these communities of basic sciences the acquisition of research data may be the result of a scientific endeavour on its own, such as genome sequencing, generation of earth observation data etc., therefore the data resulting from such endeavours have a high value for other researchers. Other scientists make use of such data in their own analyses, for which such data are the key resource. Therefore, it is essential that these data are easy to find (if possible with automated search in specialized open data repositories), accessible, interoperable and reusable (FAIR).

In contrast, in sensor and systems R&D as conducted in ALFA, the key result of the development is the sensor or the sensing system, but not the data acquired with this system. As a result, the overall approach of FAIR data does only fit partially to the nature of R&D conducted in ALFA. Wherever possible and appropriate, the ALFA partners will adapt concepts of FAIR data management.

In the context of research data that is used in scientific publications by ALFA partners, each partner owning such data is responsible for long-term storage and retrieval, as well as for making data accessible to the scientific community, preferentially on an individual basis (“by request”). If, for example, an ALFA partner is contacted by a reader of his publication with a request for data, research data will be made available for validation of results or other purposes, such as testing of signal processing algorithms – as long as not prohibited by security provisions. Therefore, ALFA will follow rules of good scientific practice, as laid out e.g. in [ALLEA; 2017].

Chapter 5 List of Abbreviations

Abbreviation	Translation
CA	Consortium Agreement
DMP	Data Management Plan
DoA	Description of Action
DOI	Digital Object Identifier
EEA	External Ethical Advisor
FAIR	Findable, Accessible, Interoperable, Re-useable
LEA	Law-Enforcement Agency
R&D	Research and Development
SSC	Security Sensitivity Committee
SVN	Apache Subversion (often abbreviated SVN) is a software versioning and revision control system distributed as open source
WP	Work Package

Chapter 6 Bibliography

ALLEA, 2017. ALLEA, the European Federation of Academies of Sciences and Humanities, *The European Code of Conduct for 'Research Integrity*, Revised Edition, online:
<http://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf>

EC, 2016. European Commission, Directorate-General for Research & Innovation, *H2020 Programme Guidelines on FAIR Data Management in Horizon 2020*, V3.0, 26 July 2016

Wilkinson, M.D., 2016. Comment: The FAIR Guiding Principles for scientific data management and stewardship, *Scientific Data (Nature.com)*, Vol. 3, Article No. 160018, online:
<http://dx.doi.org/10.1038/sdata.2016.18>