

# H2020-GALILEO-GSA-2017

## Innovation Action



### Galileo-EGNOS as an Asset for UTM Safety and Security



*This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776293*

## Data Management Plan

<b>Report Identifier:</b>	D1.5		
<b>Work-package, Task:</b>	WP1, T1.1	<b>Status – Version:</b>	02.00
<b>Distribution Security:</b>	PU	<b>Deliverable Type:</b>	R
<b>Editor:</b>	Enric Oliveres (EVADS)		
<b>Contributors:</b>	Ramón Sánchez (EVADS), Andrea Mazzù (RINA-C), Ivan Tesfai (RINA-C), Ángel Rodríguez (USE)		
<b>Reviewers:</b>	Francisco Javier Pancorbo (EVADS)		
<b>Quality Reviewer:</b>	Elisabeth Pérez (EVADS)		
<b>Keywords:</b>	Data management plan; Galileo; EGNOS; RPAS; UTM; U-Space; ATM; traffic management; safety; security; Zenodo; metadata; Open Research Data Pilot (ORD)		
Project website: <a href="https://www.projectgauss.eu/">https://www.projectgauss.eu/</a>			

TITLE					
Data Management Plan					
ABSTRACT					
GAUSS is a R&D project and involves several type of data to develop the system. This document describes the type of data expected to be involved in the project and how it will be managed; the overall goal is to ensure an efficient data management and maximize the dissemination potential of the public project results.					
CHANGE CONTROL					
Version	Description	Created		Revised by	
		Author	Date	Responsible	Date
00	Internal consortium version	Enric Oliveres	06/08/2019	Francisco Javier Pancorbo	13/08/2019
01	First uploaded version	Enric Oliveres	28/08/2019	Francisco Javier Pancorbo	29/08/2019
02	Update according to reviewers feedback	Enric Oliveres	26/11/2019	Francisco Javier Pancorbo	28/11/2019

© 2019, GAUSS Consortium. All rights reserved.

# Content

Content .....	3
List of Figures .....	4
List of Tables .....	5
1. Introduction.....	6
1.1 Structure of the document.....	7
2. Data summary .....	8
2.1 Collected data.....	8
2.2 Working data .....	9
2.3 Generated data.....	9
3. FAIR data .....	11
3.1 Making data findable, including provisions for metadata .....	11
3.2 Making data openly accessible .....	11
3.2.1 Zenodo .....	12
3.3 Making data interoperable .....	13
3.4 Increase data re-use.....	13
4. Allocation of resources.....	14
5. Data security.....	15
6. Ethical aspects.....	16
Acronyms.....	17
Bibliography.....	18

---

## List of Figures

Figure 1: example of a telemetry csv file .....	10
Figure 2: Top 20 data providers for H2020 publications [20]. .....	13

---

# List of Tables

Table 1: potential standards for metadata [14]. ..... 11

# 1. Introduction

This document describes the data management life cycle for the data to be collected processed and/or generated by GAUSS, since it postulates as a beneficiary of the Open Research Data Pilot (ORD pilot); it has been developed following the H2020 guidelines and template [1].

It includes information on: the handling of research data during and after the end of the project; what data will be collected, processed and/or generated; which methodology will be applied; whether data will be shared/made open access; and how data will be curated and preserved (including after the end of the project).

The Data Management Plan (DMP) is intended to be a living document in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur; therefore this is a first version and it may be updated should the need arise [2].

The ORD pilot aims to improve and maximise access to and re-use of research data generated by Horizon 2020 projects and takes into account the need to balance openness and protection of scientific information, commercialisation and Intellectual Property Rights (IPR), privacy concerns, security as well as data management and preservation questions. According to European Commission figures [3] approximately an average of 67% of H2020 proposals opted in this initiative in the period 2014-2016.

Research data will be findable, accessible, interoperable and reusable (FAIR), to ensure it is soundly managed. Good research data management is not a goal in itself, but rather the key conduit leading to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse.

Several studies indicate that openness increases citations [4], which will be a key asset for dissemination of GAUSS' outcomes and "openness also improves reproducibility of your research results – and it might introduce new and perhaps unexpected audiences to your work" [5].

Nevertheless, sharing some information and data could be harmful not only for the consortium and its members but also for drone operations. This is why, although general guidelines are explained in this document, a consortium decision<sup>1</sup> will be made in an individual basis before sharing any specific information outside the consortium and/or making some information public, besides the deliverables that were stated as public in the Project Management Plan [6].

Some H2020 projects contemplate data management as one of the goals of the project, for example to create or populate a database ( [7], [8]) or to perform statistical studies and correlations ( [9] [10] [11]). On the other hand, in this project data is used as a mean to develop the GAUSS system and therefore the amount of data and information to publish and distribute at the end is reduced and it will probably be focused on the form of written reports and articles; nevertheless, other information is likely to be also published, as it is explained in this document.

---

<sup>1</sup> This "consortium decision" is not a single decision, it is a decision that will be made each time it is not sure whether some information could be shared outside the consortium or could be made public. More information on how consortium decisions are made is included in the Project Management Plan of GAUSS [12].

---

## 1.1 Structure of the document

The current Section 1 introduces the document and briefly describes the information included.

Section 2 summarises the main type of data that GAUSS will work with.

Section 3 describes the procedures that will be followed to ensure that the main research GAUSS data is as FAIR<sup>2</sup> as possible.

Section 4 describes how the costs related to data management will be assumed within the project.

Section 5 deals with data security issues.

Section 6 briefly describes potential ethical aspects

---

<sup>2</sup> “Findable, Accessible, Interoperable and Re-usable”

## 2. Data summary

GAUSS will both collect and generate information; the former is essential to have a proper view and knowledge of current technologies: UTM operation (WP2), Positioning (WP3), Integrity and Security (WP4) and UTM technologies (WP5). Furthermore, outputs from GAUSS will include technical reports (D3.2, D4.2, D5.2, etc.) and also operational definition and requirements (D2.1, D2.4, etc.).

This data generation will allow defining RPAS systems which will improve precision, safety and security and will optimize their usability by improving their coordination and enhancing their compatibility; these concepts are fully aligned with the project objectives.

The most common format will be reports in the form of .pdf files and datasheets, since they are an easy way to present post-processed data. GAUSS includes both unitary and laboratory tests together with field trials, these activities will generate huge amount of raw data in several formats: pictures and videos for visual proofs, .txt/.csv files for GNSS locations, .kml for wide use of geo-position information, etc.

Most data obtained during the unitary tests will be used to update and modify the different systems that will be developed in order to fulfil GAUSS objectives and furthermore successfully perform the final field trials.

Data will come from different sources depending on whether it is collected or generated:

- ❑ Collected data:
  - ✓ GAUSS partner knowledge
  - ✓ Proactive research
  - ✓ Feedback from AB and stakeholders.
  - ✓ Open source code
- ❑ Generated data:
  - ✓ GAUSS partner internal development
  - ✓ Tests
  - ✓ Field trials

Due to the complexity of the project and its duration collected and especially generated data will be of considerable size. However, most of this size will be due to raw experimental data (experiments and trials generate huge amounts of data that need to be processed to elaborate results) and an effort will be made to summarize such data to make it human-understandable and to eliminate irrelevant or repetitive data that do not add value to the project, in order to keep overall data size to a minimum.

### 2.1 Collected data

In order to develop GAUSS systems, each partner will gather data; this kind of information will be managed by each partner individually and it is not expected to be shared/published outside the consortium on behalf of GAUSS. Therefore, this report does not focus on these type of data.



Some examples and sources of these type of data are papers, journals/magazines, webpages information, other R&D projects<sup>3</sup>, etc. and the most common formats are: .pdf and other office formats.

## 2.2 Working data

This type of data will be generated by different partners within the consortium in order to design/develop the GAUSS systems they are responsible for. Similar to the previous type of data, working data is not expected to be shared/published outside the consortium on behalf of GAUSS. Therefore, this report does not focus on these type of data.

Some examples of these type of data are assessment reports, datasheets, minutes, etc. and the most common formats are: .pdf and other office formats.

## 2.3 Generated data

Finally, GAUSS will generate some data as outcome and results of the project, some of these will become relevant information which will be of public domain and therefore published, some information will only be shared among relevant stakeholders (advisory board and other H2020 projects) and other information will be confidential and will not be published; this report focuses on the former two.

Several type of files are expected for this purpose:

- ❑ Reports in the form of .pdf files. Some of GAUSS reports will be public, according to the GAUSS PMP (Project Management Plan) [6]: D2.1-Design of UTM Concept of Operations; D2.2-Definition of UTM Scenarios and use cases report; D4.1-Report on EGNSS security-enabling features relevant for RPAS; D6.4-First trials results report; D6.5-Second trials results report; D6.6-Performance results analysis and conclusions.
- ❑ Presentations in the form of .pdf files; for example for GAUSS workshops or public events.
- ❑ Telemetry files with information of location and parameters to be assessed for each flight, mainly flight plans and 4D trajectories, GNSS signal integrity, GNSS signal security and security of UTM communications to name a few)<sup>4</sup>. This is expected to be a .csv file in order to increase dissemination potential since they are easy to work with and manipulate [12]<sup>5</sup>. The main source of these type of files will be the field trials where it is expected that for each flight of each aircraft a series of three telemetry files will be generated:
  - ✓ Information coming from the current positioning system.

---

<sup>3</sup> Contact was established with a H2020 project (BLUEGNSS [19]) co-financed by GSA which gathered actual GNSS data (Galileo and EGNOS included)

<sup>4</sup> Leaders of each technical work package (WP3, WP4 and WP5) will be responsible for defining such parameters so attention is put on obtaining them during field trials.

<sup>5</sup> GAUSS is a R&D project and it was decided to use a widely extended format with high degree of customization to quickly change the structure when/if project conditions required it and that it was easy to work with even with partners with no expertise with the aviation industry.

Nevertheless, this telemetry files (.csv for example) could be converted to ASTERIX in post-processing. Furthermore, this would allow to easily convert this telemetry files to formats commonly used in aviation, such as EUROCONTROL's ASTERIX [24].

- ✓ Information coming from the GAUSS positioning system.
- ✓ True location information (coming from RTK or DGPS).

Telemetry files will also be generated by means of a simulator (Gazebo Simulator).

```
EU:ROLL;EU:PITCH;EU:YAW;POS:LAT;POS:LON;POS:ALT;VEL:IAS;VEL:GS;VEL:GPS;CAM: PAN;CAM:TILT;CAM:ROLL  
0,03116;-0,037753;0,063791;42,650065;-5,708916;1097,578369;18,645342;21,117432;21,12019;0;-0,02;0  
0,03116;-0,037753;0,063791;42,650065;-5,708916;1097,578369;18,645342;21,117432;21,12019;0;-0,02;0  
0,03116;-0,037753;0,063791;42,650065;-5,708916;1097,578369;18,645342;21,117432;21,12019;0;-0,02;0  
0,03116;-0,037753;0,063791;42,650065;-5,708916;1097,578369;18,645342;21,117432;21,12019;0;-0,02;0
```

Figure 1: example of a telemetry csv file<sup>6</sup>

Overall GAUSS data will be useful to RPAS industry and ATM/UTM/U-space community and low-level technical information will be useful for industries related to RPAS such as OEM and R&D centres that could build their developments on GAUSS information; other (present and future) R&D projects (including H2020) could use this information like “Collected data” to design and develop their systems.

---

<sup>6</sup> Formatted following Spanish coding.

## 3. FAIR data

### 3.1 Making data findable, including provisions for metadata

Files produced during the project will be tagged accordingly to ease their discovery and access; furthermore, since GAUSS involves a high level of experimental activities, raw data will be generated and their metadata associated will also be managed.

In order to maximise accessibility to GAUSS results a straightforward naming convention and keywords will be used together with a coherent and detailed versioning; such keywords will include "European Union (EU)", "GSA", "Galileo", "EGNOS", "GNSS", "UTM", "UAS", "U-Space", "satellite navigation system"; the name of the action, acronym and grant number; the publication date, and length of embargo period if applicable, and a persistent identifier (DOI). DOI identifiers will be used to cite reports and data sets, they are supported by most file repositories and its implementation will ease the process of identifying content in addition to provide a persistent link to its location on the Internet. The naming convention used for deliverables is the following (following the internal nomenclature of GAUSS):

GAUSS-D<Deliverable code>-<Deliverable name>-v<VV>

where: <VV> is the Version number of the delivery. The naming convention in place instructs to use underscores “\_” instead of blank spaces “ ” and avoid special characters such as “~!@#\$”. Where relevant, date designators will be used with the following format: YYYYMMDD. This format makes sure that all files can be sorted in chronological order.

In case metadata information is useful to be included in the files, several standards will be assessed focusing on generic standards since a specific standards for GAUSS metadata files has not been found; an initial assessment has been performed, see Table 1.

Metadata standard	Brief description
Data Package	Generic wrapper format for exchanging data.
ISA-Tab	General purpose to collect and communicate complex metadata
PREMIS	Defines a set of metadata that most repositories of digital objects work with
ISO 19115	Schema used for describing geographic information and services
Dublin Core	Basic, domain-agnostic standard and one of the most widely used as a metadata standard. It is included in the OpenAIRE guidelines [13].

*Table 1: potential standards for metadata [14].*

### 3.2 Making data openly accessible

Some deliverables will be publicly available (D2.1, D2.2, D4.1, D6.4, D6.5 and D6.6) and other project information might be made public if agreed by all partners and PO. In this sense, some GAUSS partners (especially research centres) will publish scientific/technical publications<sup>7</sup> regarding their field.

<sup>7</sup> More information on these publications may be found in periodic dissemination reports [15]

GAUSS results that are to be publicly available will be made accessible via Zenodo (see Section 3.2.1 for further information), an open source repository, although publication in other directories will also be encouraged. EVADS will manage the upload and submission of the files to Zenodo although such files are to be provided by the partner responsible of the task associated to each file within the scheduled deadlines.

Some software implementations, especially regarding WP5, will be made openly available by USE in Github.

Furthermore, in order to maximize project dissemination, some results are likely to be published in scientific journals and GAUSS website, see [15] for a potential list of scientific journals and public reports. EVADS will manage the upload and submission of the deliverables to the website although they are to be provided by the partner responsible of the task associated to the deliverable within the scheduled deadlines.

### 3.2.1 Zenodo

Zenodo [16] is an interdisciplinary open research data repository service built and operated by CERN and OpenAIRE that enables researchers from all disciplines to share and preserve their research outputs, regardless of size or format.

Other repositories were assessed [17] and Zenodo was chosen for its simplicity, its wide use (Figure 2 shows how Zenodo is widely used among H2020 projects) some partners' previous experience<sup>8</sup> and their relation with CERN and OpenAIRE. Furthermore, Zenodo also contemplates the use of API.

A persistent identifier (DOI) is issued to every published record on Zenodo. This is a top-level and a mandatory field in the metadata of each record which helps to make uploads easily citable.

In order to upload data to Zenodo GAUSS consortium has created a project account and all files to be made public will be submitted through this account to maximize dissemination and project awareness. Files may be downloaded directly without the need an account, which will increase dissemination reach.

Files uploaded to Zenodo have a size limit of 50 Gb (GAUSS files will be much lighter than this) and will be retained online for the lifetime of the repository, which is the same than for the CERN laboratory (the host) and it has an experimental program defined for the next 20 years at least [18]. Metadata is exported in several standard formats such as MARCXML, Dublin Core, and DataCite Metadata Schema [16].

Restricted access can be configured on Zenodo although this is not expected to be used since the repository will be used to upload public information.

Furthermore, Zenodo provides analytics of uploaded information which will be used for analysing dissemination impact, [19] and [15].

---

<sup>8</sup> USE has already used this repository in other projects

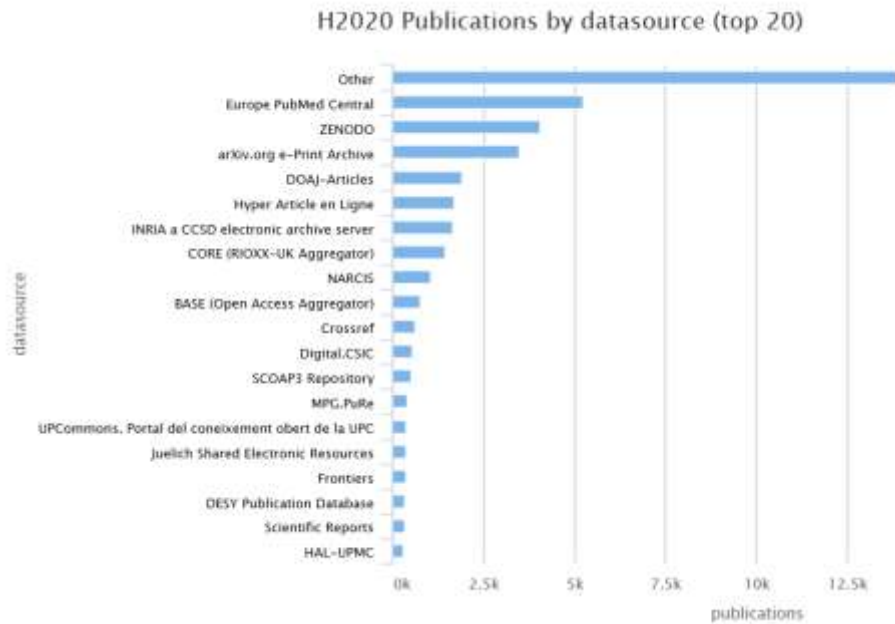


Figure 2: Top 20 data providers for H2020 publications [20].

### 3.3 Making data interoperable

As it has been explained in Section 2, data to be published will be in very common formats so it is expected that they will be easily used by external parties (researchers, institutions, organisations, etc.).

Standards could be used at some times for specific files and information summarised in reports will also include a detailed glossary of the specific notations used along the document to ensure concepts are clearly defined.

### 3.4 Increase data re-use

In order to maximize information reach and ease their access, GAUSS data publicly available will be published under “Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)”<sup>9</sup> as soon as possible and at the latest of publication.

One of the main goals of GAUSS’ is to help in defining future drone regulations so it is expected that most of GAUSS’ outcomes will be further used.

Data to be published will be subject to quality assurance procedures (further explained in [6] and [21]).

<sup>9</sup> “Attribution-ShareAlike 4.0 license lets others remix, tweak, and build upon the reported work even for commercial purposes, as long as they credit the owner and license their new creations under the identical terms. This license is often compared to “copyleft” free and open-source software licenses” [23]

---

## 4. Allocation of resources

Most of this FAIR initiative will not involve direct costs such as licenses, software or hardware since open source third parties servers and technologies will be used to host public data (see Section 3.2). Nevertheless, it will require spending time in managing such data (receiving, selecting, adapting, uploading, etc.) which has been contemplated within the management activities of both the project (WP1) and the dissemination activities (WP7). Although the final upload to the repository and/or website will be performed by EVADS, each partner is to deliver the required files to be uploaded within the agreed deadlines.

Some activities with a direct cost associated have been contemplated (such as publications in scientific journals under payment, attendance to specific events, etc.) in order to maximize GAUSS dissemination. This cost has already been contemplated within travel and other goods/services costs associated to each partner and it has been specified in the Grant Agreement.

## 5. Data security

As it has been explained in Section 2, GAUSS will work with different type of data and it may be classified into some groups:

- ❑ Internal data: this is the data that each partner works with individually within the framework of GAUSS (reports, papers, software algorithms, etc.). Each partner has measures and procedures implemented within their company (secure storage, periodic backups, malware detection, etc.) to ensure CIA triad (Confidentiality, Integrity and Availability)<sup>10</sup>.
- ❑ Transferred data: this is the data/information that is transferred among the consortium within the framework of GAUSS, for example deliverables preliminary versions, specification documents, field trials telemetry, etc. As it was defined in the Grant Agreement, the highest level of security of the information within GAUSS is “Confidential”, which is not as restricting as other levels (such as Classified).
- ❑ Published data: this is the data/information that is published in the website and/or the repository. As it was defined in [6] the former will be operative during the project duration and information on how the repository manages the information is included in Section 3.2. The repository allows for restricted access to specific files, but this feature is not considered relevant at this point since all the information published there will be public for dissemination purposes.

---

<sup>10</sup> Some of the most critical information is regarding jamming/spoofing. Real time flow of these data will only be used in laboratory environment and in the field trials always within GAUSS consortium and it will not be exchanged with external parties. Furthermore, during the field trials data flow among the different systems will be encrypted; a specific task is designed for this purpose (T5.4: Security in UTM-related communications), more information may be found in [6].

---

## 6. Ethical aspects

No ethical or legal issues are expected regarding data sharing since no private data is expected to be included in the information that will be published. Whenever such data is gathered (pilot data, serial numbers, etc.) it will be omitted or anonymized before submitting it.

Furthermore, information that has been obtained through external means (for example, AIS data and FIS data) will not be shared unless explicit agreement from their provider is obtained even when it does not contain private data.



# Acronyms

AIS	Automatic Identification System
DGPS	Differential GPS
DMP	Data Management Plan
DOI	Digital Object Identifier
EC	European Commission
FIS	Flight Information Service
GNSS	Global Navigation Satellite System
GSA	European GNSS Agency
OEM	Original Equipment Manufacturer
ORD	Open Research Data
PMP	Project Management Plan
PO	Project Officer
R&D	Research and Development
RPAS	Remotely Piloted Aircraft System
RTK	Real-Time Kinematic
UAS	Unmanned Aircraft System
UTM	UAS Traffic Management

---

## Bibliography

- [1] European Commission, "H2020 Programme: Guidelines on FAIR Data Management in Horizon 2020," 2016.
- [2] European Commission, "Data management," [Online]. Available: [https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management\\_en.htm](https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm). [Accessed August 2019].
- [3] European Commission, "Open access: the uptake of the Open Research Data Pilot in Horizon 2020. Explanatory note to the finalised data for 2014-2016," 2018.
- [4] SPARC Europe, "Setting the Default to Open," [Online]. Available: <https://sparceurope.org/what-we-do/open-access/sparc-europe-open-access-resources/open-access-citation-advantage-service-oaca/oaca-list/>. [Accessed April 2018].
- [5] OpenAIRE, "OpenAIRE FAQ," [Online]. Available: <https://www.openaire.eu/support/faq#article-id-234>.
- [6] GAUSS, "D1.1: Project Management Plan," 2018.
- [7] ArchAIDE, "ArchAIDE," [Online]. Available: <http://www.archaide.eu/>. [Accessed August 2019].
- [8] ArchAIDE, "Initial DMP," 2019.
- [9] European Commission, "Emerging population issues in sub-Saharan Africa: Cross-checking and promoting demographic data for better action," [Online]. Available: <https://cordis.europa.eu/project/rcn/199932/factsheet/en>. [Accessed August 2019].
- [10] European Commission, "Publishing and Enriching Linked Open Statistical Data for the Development of Data Analytics and Enhanced Visualization Services," [Online]. Available: <https://cordis.europa.eu/project/rcn/110723/factsheet/en>. [Accessed August 2019].
- [11] European Commission, "Big Data Analytics for Time Critical Mobility Forecasting," [Online]. Available: <https://cordis.europa.eu/project/rcn/199835/factsheet/en>. [Accessed August 2019].
- [12] R. Gupta, "Advantages of CSV file," [Online]. Available: <https://github.com/prakhargvp/csvWithCplusplus/wiki/Advantages-of-CSV-file>. [Accessed August 2019].
- [13] OpenAIRE, "OpenAIRE Guidelines," [Online]. Available: <https://guidelines.openaire.eu/en/latest/>. [Accessed August 2019].

- 
- [14] Metadata Standards Directory Working Group, "RDA Metadata Standards Directory," [Online]. Available: <http://rd-alliance.github.io/metadata-directory/>. [Accessed August 2019].
- [15] GAUSS, "D7.2 Periodic report on dissemination and exploitation activities," 2019.
- [16] Zenodo, "Zenodo," [Online]. Available: <https://www.zenodo.org/>. [Accessed August 2019].
- [17] EUDAT, "B2Share," [Online]. Available: <https://b2share.eudat.eu/>. [Accessed August 2019].
- [18] Zenodo, "Zenodo Policies," [Online]. Available: <http://about.zenodo.org/policies/>. [Accessed August 2019].
- [19] GAUSS, "D1.2 Periodic Project Report," 2019.
- [20] OpenAIRE, "H2020 Monitoring," [Online]. Available: <https://www.openaire.eu/h2020-stats>. [Accessed August 2019].
- [21] GAUSS, "D1.4 Quality Risk management report".
- [22] BLUE MED, "BLUEGNSS Project," [Online]. Available: <http://www.bluedmed.aero/nodo.php?id=326>. [Accessed August 2019].
- [23] Creative Commons, "Attribution ShareAlike," [Online]. Available: <https://creativecommons.org/licenses/by-sa/4.0/>. [Accessed 08 2019].
- [24] EUROCONTROL, "ASTERIX. All-purpose structured EUROCONTROL surveillance information exchange," [Online]. Available: <https://www.eurocontrol.int/asterix>. [Accessed November 2019].