European
Global Navigation
Satellite Systems
Agency

D1.3 Data management plan

Project title:

Hybrid UAV-UGV for Efficient Relocation of Vessels (HUUVER)

Contribution by

CERVI ROBOTICS SP Z O O	CR
BRIMATECH SERVICES GMBH	BRI
RECTANGLE SP Z O O	REC
LAPPEENRANNAN-LAHDEN TEKNILLINEN YLIOPISTO	LUT
EVERIS AEROESPACIAL Y DEFENSA SL	EVADS
GINA SOFTWARE SRO	GINA
BLADESCAPE AIRBORNE SERVICES GMBH	BLDS

Signature for Submission and Approval

Cervi Robotics sp. z o. o. (as Leading Contractor)	Date 01/07/2020	.....
European GNSS Agency (GSA) (Customer)	Date	.....

Point of Contact

Name: Krzysztof Malik **Role:** Project Manager **Email:** Krzysztof@cervi.io **Phone** **Mobile:** +48 730 068 052	**Company:** Cervi Robotics sp. z o. o. **Site:** www.cervirobotics.com **Address: Jasionka 954E, 36-002 Jasionka**

Table of contents

MODIFICATION RECORD

Rev.	Modification No.	Date	Description	Author/s
01.00	N/A	03/07/2020	First issue	Krzysztof Malik, Javier Marinelli

1. Introduction

This document describes the data management life cycle for the data to be collected processed and/or generated by HUUVER, since it postulates as a beneficiary of the Open Research Data Pilot (ORD pilot).

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 [1].

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 [2] 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 [3], which will be a key asset for dissemination of HUUVER's outcomes and "openness also improves reproducibility of your research results – and it might introduce new and perhaps unexpected audiences to your work" [4].

Nevertheless, sharing some information and data could be harmful not only for the consortium and its members but also for society. This is why, although general guidelines are explained in this document, a consortium decision[1] 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.

Some H2020 projects contemplate data management as one of the goals of the project, for example to create or populate a database ([5], [6]) or to perform statistical studies and correlations ([7] [8] [9]). On the other hand, in this project data is used as a mean to develop the HUUVER 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.

2. Data Collection

The HUUVER solution will collect and manage the following data:

Localization data

The drone will be equipped with a high-precision navigation system, which will be composed from the information received from various instruments. The input information will deliver: GNSS/Galileo positioning data, inertial rotation and acceleration data, magnetic field vector, air pressure measurements, distance to ground and proximity to the near-field obstacles. The output of the navigation system will be processed information giving: absolute position of the drone, attitude of the drone, velocity and acceleration vectors, altitudes with relation to the ground and to the navigation frame.

Human detection data

The system will detect humans inside the field of view of both sensors on the considered mountain scenario. Depending on the field of view of the setup (drone height, camera tilt and camera zoom) the size of the human will be different. In order to achieve proper results from the IR/EO system, the minimum object size to detect a human with the IR/EO system should be 32x32 pixels.

For example: if the drone flight at height of 100 meters, with -45 degrees of camera tilt and 90 degrees of zoom, the system could cover an area of 250m on horizontal using the 4K EO sensor and 100m for the IR sensor (using the fixed zoom of 35 degrees) keeping a proper minimum object sizes.

The IR/EO system should align the EO and IR images regarding their positions, image size and zoom. This visualization will be provided at minimum 25 fps.

The IR/EO system should detect humans on the considered scenario using both sensors. For each detected human the system will provide:

- An object id: 0, 1, 2, etc.
- An object score that gives a probability confidence between 0-100. A higher object score means more confidence to be human. This score will be combined from the EO and IR detections.
- Image position in pixels where the human has been detected.A 2D bounding box will be provided with the upper left corner, width and height.
- Absolute global position. The system will provide an estimate of the center of the detected human on absolute global coordinates: latitude, longitude, and latitude in degrees.

Obstacle detection data

This system will detect obstacles inside its field of view on the considered industrial scenario. The selected LiDAR has 32 degrees of vertical field of view and there is no limit on the horizontal field of view because it is totally circular. The position of the LiDAR will have a great influence on the obstacles detected and should be defined by CER.

The LiDAR signal will also be aligned with the EO/IR images in order to provide a better environment visualization. The EO field of view will be employed to project the LiDAR points inside this image. This visualization will be provided at 10 Hz.

The LiDAR system should detect obstacles in the considered scenario. For each obstacle the system will provide:

- An obstacle id: 0, 1, 2, etc.
- Image position in pixels where the obstacle has been detected. A 2D projected bounding box with the upper left corner, width and height, and the 3D bounding box.

Video picture

The optical system will provide a Full HD video streaming composed (RTSP protocol) from the three sensors (EO, IR and LiDAR) via an Ethernet cable connection to an on-board. A configurable command (or fixed) will be implemented to compress more or less the compressed video sequence regarding the enabled bandwidth.

A control command to select the visualization mode will be provided. Three modes will be available:

- EO visualization. In this mode, the optical system will provide just the EO images.
- Combined EO and IR visualization. This mode combines the images from EO and IR sensors inside 1 image. The IR image is mapped to a colour map in order to highlight the warmer image areas. The coloured IR will be aligned regarding the EO imagen, and this image can be overlaid with a transparency over the EO imagen.
- Combined EO, IR and LiDAR visualization. In this mode, the LiDAR points are projected inside the EO image field of view.

Numerical data

The optical system will provide two types of numerical data that should be transported by the MAVLink protocol.

- Optical system status. The optical system will send the gimbal attitude, gimbal mode, and camera zoom. This data status information can be transferred at 25-50 Hz if it is required.
- Optical system results. The EO/IR and LiDAR systems will provide their numerical results, such as, human id, obstacle id, obstacle position and volume, human position and size, human absolute position, etc.
- velocity vector of the drone with respect to nadir
- gimbal attitude w.r.t. the drones body frame

The numerical data should contain three-dimensional information about position, attitude and velocity of the drone, also the camera modes and gimbal control signals. Another visualization mode control will be introduced to select between the three proposed modes. The data should be transported by the MAVLink protocol.

3. Documentation and Metadata

What documentation and metadata will accompany the data?

In order to ensure the data generated in the project is Findable, Accessible, Interoperable and Re-usable (FAIR), it should be accompanied by technical documentation and metadata.

Therefore, files produced during the project will be tagged accordingly to ease their discovery and access; furthermore, since HUUVER involves experimental activities, raw data will be generated and their metadata associated will also be managed.

In order to maximise accessibility to project 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”, “unmanned”; 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 HUUVER):

1. Working draft:

Each deliverable site will contain one single file as working draft. Only the deliverable responsible can update this file although it can be delegated to other partners if necessary. Only one partner will be responsible of updating the working version at a time. Deliverable working draft nomenclature will be as follows:

HUUVER-D<Deliverable code>-<Deliverable name>

2. Contributions:

Partners who contribute over this working draft can upload their contributions in separate documents or can use the online google editor. It is important to do so with Changes tracker enabled. The deliverable responsible shall integrate this contributions into the working draft. Contributions nomenclature will be as follows:

PARTNER- yyyymmdd -HUUVER-D<Deliverable code>-<Deliverable name>-r<RRR>

where: <RRR> is the Version number of the working draft which has been modified.

3. Deliveries:

Final versions submitted to H2020 portal will be kept as delivered versions of the document. Delivered versions of the deliverable nomenclature will be as follows:

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

where: <VV> is the Version number of the delivery.

4. Support:

Support documentation can be shared within the deliverable sharing site. Supporting documents for the deliverable nomenclature will be as follows:

S-HUUVER D<Deliverable code>-<Supporting document name>

Most information will be used internally by each partner in which case they shall follow the procedure they deem necessary. However, when data is to be used among the consortium, and even outside the consortium, it is important it is accompanied by relevant documentation and metadata.

Documentation associated to all kind of data may be similar since it will be in the form of reports (pdf) or presentation (pdf, ppt, etc.) and this where a project homogeneous look is important so the documents are easily associated and classified; this can be achieved by templates, common procedures, document exchange tools, repositories, etc.

On the other hand, metadata is very tailored to the data it complements and in order to homogenize and ease their transfer it is useful to use standards.

Although there are some generic standards (see Table 1) each type of data shall use a specific one[EOM1] ; for example, ISO 19115 is a schema used for describing geographic information and services[EOM2] .

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
Dublin Core	Basic, domain-agnostic standard and one of the most widely used as a metadata standard. It is included in the OpenAIRE guidelines [10].

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

4. Ethics and Legal Compliance

How will you manage any ethical issues?

Some of the data collected during the project might present some privacy issues, this is why a careful analysis of user requirements will ensure that any ethical considerations are taken into account by the project consortium. At the same time the existing and planned legal framework in each of the regions and at European level will be analysed, being GDPR the most important one. In this way all legal considerations will be identified and complied with.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

A first study was performed during the proposal where several consortium decisions were made:

- The partners agree that access rights to background knowledge is granted on a non-exclusive non-transferable basis, solely for the purposes of the carrying out of this project. Save as in exceptional circumstances, no costs shall be charged for the granting of these access rights.
- As a basic principal, IPR created during the HUUVER project shall be the property of the partner generating it. In case of a joint invention, design or work made by two or more project partners, the partners concerned agree that

they may jointly apply to obtain and/or maintain the relevant rights and shall strive to set up amongst themselves appropriate agreements in order to do so. Contributors shall be entitled to use and to license such rights.

- A partner will not publish any knowledge provided by another partner and identified as confidential, without the other Partner's prior written approval. However, if open source software licenses apply, the open source software license rules will apply for publishing knowledge.

Nevertheless, the consortium is aware of the importance of copyright and IPR issues due to type of technology and data developed and obtained in this project and a specific task will be performed to further study this issues (T7.4: Deployment and IPR management) which will summarise its conclusions in a deliverable (D7.4: Deployment plan and IPR management report), more information on those may be found in the proposal and/or in the project management report.

5. Storage and Backup

How will the data be stored and backed up during the research?

Each partner shall be responsible for the information they need to perform their activities (gathering, storing and back up).

Project HUUVER will use Google drive as a shared work-space where the partners will exchange, store and backup all the data.

How will you manage access and security?

Each partner has in place different measures to protect and manage the access to their data (log in credentials, firewalls, etc.). The project will use a toll to easily and quickly manage (grant and deny) access to different information/data to different people (internal and external) to ensure data is only available to the correct people.

6. Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

What is the long-term preservation plan for the dataset?

Data is to be preserved in the long term and will be uploaded to Zenodo where it 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 [12].

7. Data Sharing

How will you share the data?

Some deliverables will be publicly available and other project information might be made public if agreed by all partners and PO. In this sense, some HUUVER partners will publish scientific/technical publications[1] regarding their field.

HUUVER results that are to be publicly available will be made accessible via Zenodo (this tool is explained in more detailed below), an open source repository, although publication in other directories will also be encouraged. The project coordinator 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 could be made openly available in Github.

One of the main goals of the project is to help in defining future drone regulations so it is expected that most of the outcomes will be further used; so, in order to maximize information reach and ease their access, project data publicly available will be published under "Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)" as soon as possible and at the latest of publication.

Zenodo summary[EOM1]

Zenodo [13] 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 [14] and Zenodo was chosen for its simplicity and its wide use (Figure 1 shows how Zenodo is widely used among H2020 projects).

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.

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 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 [12]. Metadata is exported in several standard formats such as MARCXML, Dublin Core, and DataCite Metadata Schema [13].

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.

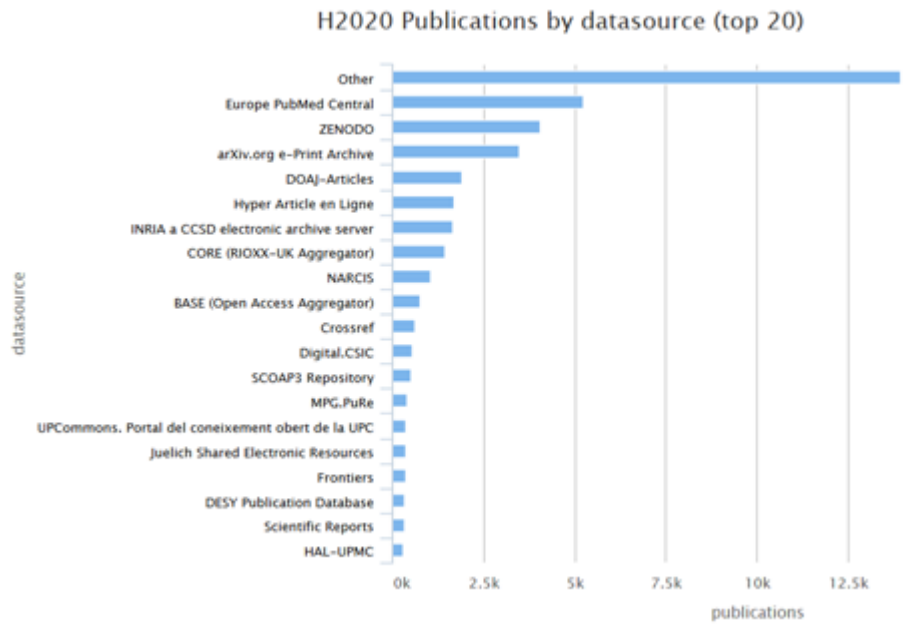

Figure 1: Top 20 data providers for H2020 publications [15].

Are any restrictions on data sharing required?

Data will only be transferred among the consortium when it is needed for the project development and will only be made public according to the project management report. Some further information could be made public or distributed to specific stakeholders if it is relevant, prior a consortium decision, and always keeping the ethics and legal compliance and IPR.

8. Responsibilities and Resources

Who will be responsible for data management?

The project coordinator will manage the upload, submission and distribution of all the data although such files are to be provided by the partner responsible of the task associated to each file within the scheduled deadlines.

What resources will you require to deliver your plan?

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. 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 and the dissemination activities.

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 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.

9. References

[1] 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. [Accessed August 2019].

[2] 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.

[3] 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].

[4] OpenAIRE, “OpenAIRE FAQ,” [Online]. Available: https://www.openaire.eu/support/faq#article-id-234.

[5] ArchAIDE, “ArchAIDE,” [Online]. Available: http://www.archaide.eu/. [Accessed August 2019].

[6] ArchAIDE, “Initial DMP,” 2019.

[7] 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].

[8] 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].

[9] European Comission, “Big Data Analytics for Time Critical Mobility Forecasting,” [Online]. Available: https://cordis.europa.eu/project/rcn/199835/factsheet/en. [Accessed August 2019].

[10] OpenAIRE, “OpenAIRE Guidelines,” [Online]. Available: https://guidelines.openaire.eu/en/latest/. [Accessed August 2019].

[11] Metadata Standards Directory Working Group, “RDA Metadata Standards Directory,” [Online]. Available: http://rd-alliance.github.io/metadata-directory/. [Accessed August 2019].

[12] Zenodo, “Zenodo Policies,” [Online]. Available: http://about.zenodo.org/policies/. [Accessed August 2019].

[13] Zenodo, “Zenodo,” [Online]. Available: https://www.zenodo.org/. [Accessed August 2019].

[14] EUDAT, “B2Share,” [Online]. Available: https://b2share.eudat.eu/. [Accessed August 2019].

[15] OpenAIRE, “H2020 Monitoring,” [Online]. Available: https://www.openaire.eu/h2020-stats. [Accessed August 2019].