Ref. Ares(2019)5524865 - 02/09/2019

Horizon 2020
European Union Funding
for Research & Innovation

European
Global Navigation
Satellite Systems
Agency

SIA
SYSTEM FOR VEHICLE-INFRASTRUCTURE
INTERACTION ASSETS HEALTH STATUS MONITORING

SIA

Project Title:	System for vehicle-infrastructure Interaction Assets health status monitoring
Starting date:	01/03/2018
Duration in months:	36
Call (part) identifier:	H2020-GALILEO-GSA-2017-1
Grant agreement no:	776402

Deliverable D9.3
Data Management Plan (DMP)

Due date of deliverable	Month 12
Actual submission date	6/05/2019
Organization name of lead contractor for this deliverable	CEIT
Dissemination level	PU
Revision	3.0

1 Executive Summary

The present document constitutes the first issue of Deliverable D9.3 “Data Management Plan” in the framework of the Project titled “System for vehicle-infrastructure Interaction Assets health status monitoring” (Project Acronym: SIA; Grant Agreement No 776402).

The objective of the activities of WP9 for the exploitation, dissemination and communication of the results and outcomes of the Project is to ensure the appropriate measures to exploit, disseminate and communicate the main results derived from the project.

This deliverable is a document which provides an analysis of the main elements of the data management policy that will be used by the members of the consortium with regard to the data generated through the life of the project. This document is not a static document. Rather, the 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.

The DMP is released in compliance with the Horizon 2020 FAIR DMP template, provided by the European Commission in the Participant Portal.

2 Table of Contents

LIST OF FIGURES

3 Abbreviations and acronyms

Abbreviation / Acronyms	Description
ABA	Axle Box Acceleration
AES	Advanced Encryption Standard
CAT	Consortium Agreement
CERN	European Organization for Nuclear Research
DMP	Data Management Plan
DOI	Digital Object Identifier
EC	European Commission
EGNSS	European Global Navigation Satellite Systems
EOA	Exploitation and Ownership Agreement
EOS	Disk-based, low-latency storage service (CERN)
EU	European Union
FAIR	Findable, Accessible, Interoperable and Re-usable
FGC	Ferrocarrils de la Generalitat de Catalunya. Spanish regional train operator
FIPS	Federal Information Processing Standard
GIS	Geographic Information System
GNSS	Global Navigation Satellite System
GSA	European Global Navigation Satellite Systems Agency
HTTPS	Hypertext Transfer Protocol Secure
IF	Interface
IEEE	institute of Electrical and Electronics Engineers
IP	Intellectual Property
IPO	Intellectual Property Rights Officer
IPR	Intellectual Property Rights
ISSN	international Standard Serial Number
JSON	JavaScript Object Notation
KPI	Key Performance Indicator
LHC	Large Hadron Collider
OAI-PMH	Open Archive Initiative-Protocol for Metadata Harvesting
OBB	Österreichische Bundesbahnen- Austrian Federal Railways, national train operator
PMO	Project Management Office
SIA	System for vehicle-infrastructure Interaction Assets health status monitoring
SMC	Strategic Management Committee
SP	Single Products
SQL	Standardized Query Language
URL	Uniform Resource Locator
WP	Work Package

4 Introduction

SIA is a Horizon 2020 project co-funded by GSA. This project is part of the Open Access to Scientific Publications and Research Data Programme in H2020. The goal of the program is to foster access to data generated in H2020 projects.

Open Access refers to a practice of giving online access to all scholarly disciplines information that is free of charge to the end-user. In this way data becomes re-usable and the benefit of public investment in the research will be improved.

As described in Figure 2, the EC provided a document with guidelines for projects participants in the pilot. The guidelines address aspects like research data quality, sharing and security. According to the guidelines, projects participating will need to develop a Data Management Plan.

The purpose of the DMP is to provide an overview of the main elements of the data management policy that will be used by the consortium with regard to the project research data. The DMP is not a fixed document but will evolve during the lifespan of the project.

The DMP covers the complete research data life cycle of the SIA project. It describes the types of research data that will be generated during the project, the strategies on research data preservation and the provision on access rights. The research data should be “FAIR”, that is findable, accessible, interoperable and re-usable. These principles precede implementation choices and do not necessarily suggest any specific technology, standard or implementation solution.

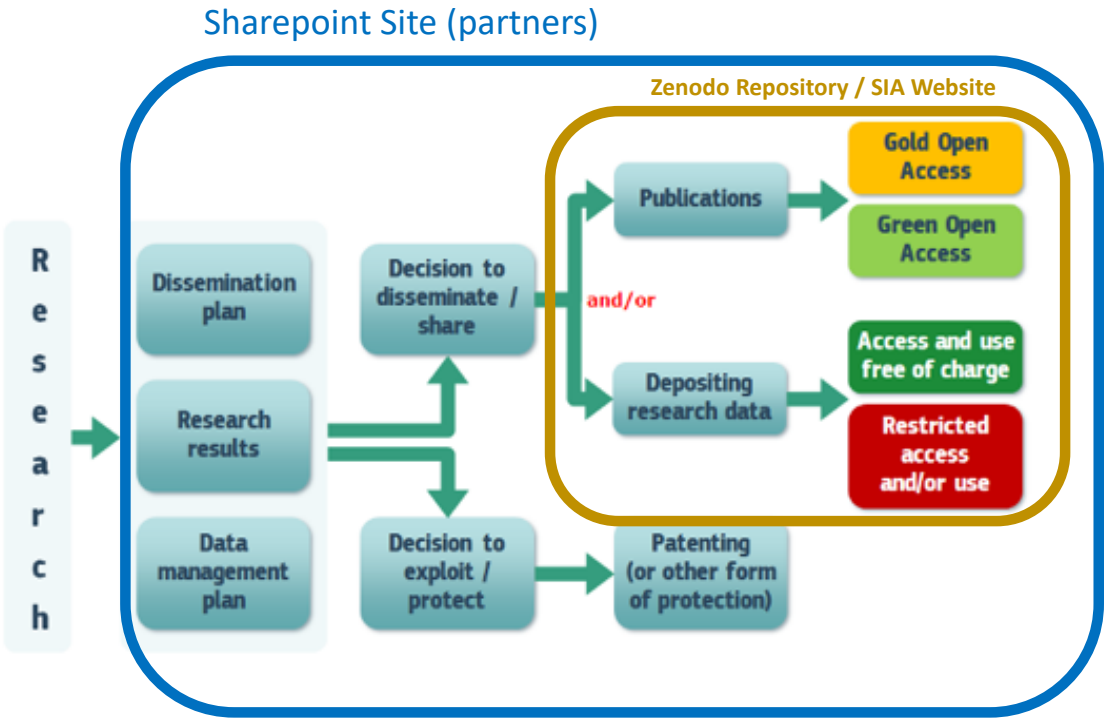

Figure 1. SIA data sharing policy

In the context of research funding, open access requirements do not imply an obligation to publish results. The decision to publish is entirely up to the grant beneficiaries. Open access becomes an issue only if publication is chosen as a means of dissemination. Moreover, open access does not affect the decision to exploit research results commercially, e.g. through patenting. The decision on whether to publish through open access must come after the more general decision on whether to publish directly or to first seek protection.

The policy for data sharing in SIA project is sketched in Figure 1, which includes different access levels for consortium members and for external users:

- Data Sharing and Access for Project Partners: all the information and sifgnificant research data related to the SIA project will be shared through a Sharepoint Private Site to which only project partners can access.
- Data Sharing and Access for External users (Open Access):
    - Project Website: will be a key in supporting the project communication to the general public and the project stakeholders during the project lifetime.
    - **Zenodo** Repository: this repository will host the Open Access Research Data both during the project lifetime and after the project ending.

More information concerning the project website and the **Sharepoint** can be found in the deliverable D1.1 of the project (Project Management and Quality of Assurance Plan) [1], which is published in the website.

29.3 Open access to research data

*Regarding the digital research data generated in the action ('**data**'), the beneficiaries must:*

(a) deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user — the following:

(i) the data, including associated metadata, needed to validate the results presented in scientific publications as soon as possible;

*(ii) other data, including associated metadata, as specified and within the deadlines laid down in the **'data management plan'** (see Annex 1);*

(b) provide information — via the repository — about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (and — where possible — provide the tools and instruments themselves).

This does not change the obligation to protect results in Article 27, the confidentiality obligations in Article 36, the security obligations in Article 37 or the obligations to protect personal data in Article 39, all of which still apply.

As an exception, the beneficiaries do not have to ensure open access to specific parts of their research data if the achievement of the action's main objective, as described in Annex 1, would be jeopardised by

making those specific parts of the research data openly accessible. In this case, the data management plan must contain the reasons for not giving access.

29.4 Information on EU funding — Obligation and right to use the Agency logo and EU emblem

Unless the *Agency* requests or agrees otherwise or unless it is impossible, any dissemination of results (in any form, including electronic) must:

(a) display the Agency logo;

(b) display the EU emblem, and

(c) include the following text:

"This project has received funding from the European GNSS Agency under the European Union's Horizon 2020 research and innovation programme under grant agreement No 776425".

When displayed together with another logo, the logo and emblem must have appropriate prominence.

For the purposes of their obligations under this Article, the beneficiaries may use the logo and emblem without first obtaining approval from the Agency or the Commission.

This does not however give them the right to exclusive use.

Moreover, they may not appropriate the logo and emblem (or any similar trademark or logo), either by registration or by any other means.

Figure 2. H2020 Open Data Access from H2020 rules

4.1 SharePoint as internal and private sharing tool

As detailed in the Quality Assurance Plan (Deliverable D1.1), in order to facilitate efficient internal communication among partners, a Microsoft SharePoint server has been set up. This site will be used to manage and coordinate the project; its access is restricted to the partners of the Consortium. This site can be accessed following the menu item that is provided in the web page entitled "Extranet". Additionally, a valid user name and password are needed. These are notified to users separately by the PMO.

SharePoint is a web-based collaboration and document management platform from Microsoft. It will be used during the Project to share workspaces and documents. Important project files will be stored and maintained on SharePoint, furthermore a folders structure has been created.

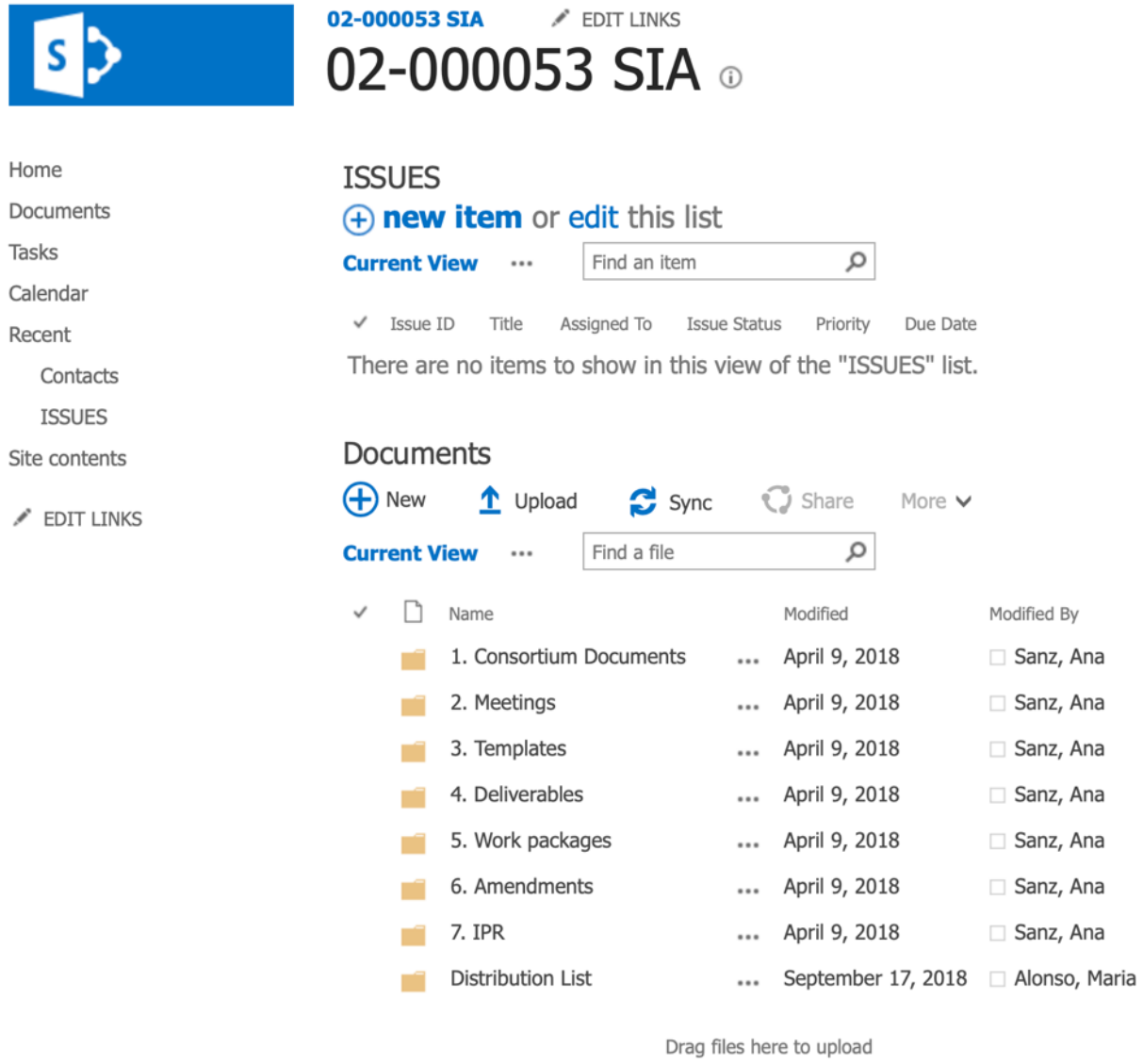

Figure 3. Home page of SIA private site

4.2 SIA Website

The SIA project website (www.siaproject.eu) is a reference point and anchor for SIA online content and outreach activities. It will explain the context, developments and ambitions of the project to our stakeholders and the general public. In order to keep a continuous and current information flow from the project to the public at large, the website will contain articles about SIA topics, interviews with experts from within and outside the project, and press releases about project highlights.

The SIA website will be a key in supporting the project communication to the general public and the project stakeholders.

Figure 4. SIA website

4.3 Zenodo repository

The repository Zenodo [2] has been chosen as the main repository to store, classify and provide open access to the stored data objects originated within the SIA project frame.

Zenodo is an open, dependable repository for all of scholarship, enabling researchers from all disciplines to share and preserve their research outputs, regardless of size or format. The main features of Zenodo that makes it a suitable tool for data sharing and preserving are:

- Zenodo is linked to Horizon2020 projects and all results are immediately linked to OpenAIRE and the EC portal.
- Share and link research: Zenodo provides a rich interface which enables linking research outputs to datasets and funding information. All open content is harvestable via OAI-PMH by third parties.
- Supports versioning: Via a top-level DOI you can support all the different versions of a file.
- Trusted, reliable, safe: Data is stored at CERN, which has considerable knowledge and experience operating large scale digital repositories. Data files and metadata are kept in multiple online and offline copies.
- Reviewing: Research materials can set to share with reviewers only, if needed.

5 Data summary

Figure 5 shows the architecture diagram of SIA, as described in deliverable D2.2 of the project [3]. The different interfaces through which data will circulate are depicted as follows:

- External interfaces (both input and output) that support the functionality described in D2.1. are depicted in red.
- Internal interfaces that are derived from the design of the architecture are depicted in blue.

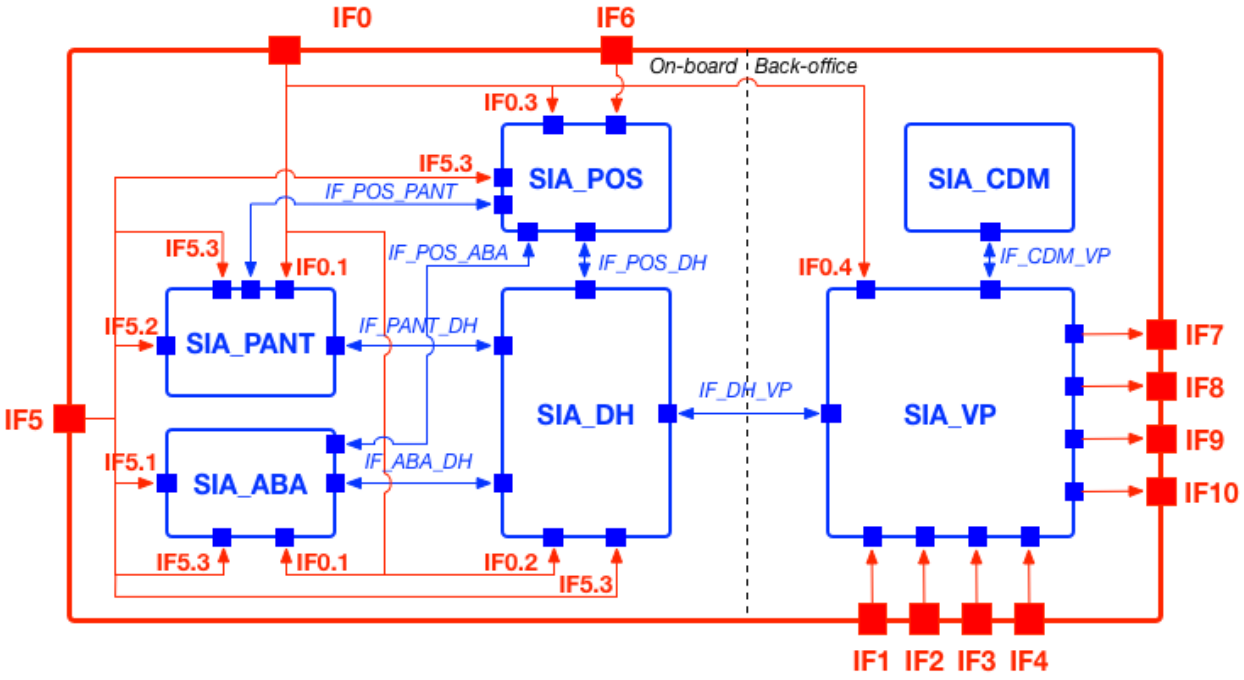

Figure 5. SIA system architecture

Relevant information on the input/output data hereby below described, shall be subjected to anonymization and/or confidentiality when necessary, especially when declaring IPR issues.

The input / output data that will circulate through the external interfaces is briefly summarized:

- **IF1: Operations data**. This information will be provided by the end-users of the consortium (FGC and OBB). It will include
  - Infrastructure characteristics (GIS map of the lines, composition of the infrastructure, location of assets, etc.)
  - Vehicle characteristics (constructive parameters, Multibody models, etc.)
  - Service and operational characteristics (list of vehicles and train compositions, timetables, speed profiles, loading curves, etc.)
- **IF2: Maintenance procedures**. This information will be provided by the end-users of the consortium (FGC, OBB, VIAS, TELICE). It will include
  - Catalogue of failures associated to the relevant components

  - KPIs that are associated to the health status of relevant components
  - Limits / Thresholds that assess the health status of relevant components
  - Maintenance actions associated to relevant components
- **IF3: Auscultation raw data**. This interface contains data coming from measurement devices that imply physical contact (e.g. auscultation train, etc.). The number of parameters, their nature and format of data are dependent on the asset under measurement (i.e. catenary, pantograph, wheelset, rail). Raw data will be stored if relevant in accordance to the Commission rules to use it (eventually and depending on IPR) in future potential related projects, and also to understand post-processed data or summarised information for analysis and post-analysis of the information leading to diagnosis.
- **IF4: Inspection raw data**. This interface contains data coming either from measurements (that do not imply physical contact) and/or inspection forms. The number of parameters, their nature and format of data are dependent on the asset under inspection (i.e. catenary, pantograph, wheelset, rail). Raw data will be stored in the same way as mentioned above.
- **IF5: Ambient**. This interface will contain different datasets that correspond to the physical magnitudes that will be measured by the on-board sensors (e.g. accelerations, displacements, forces, etc.).
- **IF6: EGNSS systems**. This interface computes data coming from EGNSS systems: GNSS signals, inertial signals and digital map.
- **IF7: Asset status**. SIA system will generate data related to the health status of the relevant components. It will include
  - Historic auscultation/inspection data
  - Current status of relevant components
  - Predicted status of relevant components
- **IF8: Early detection of component failure**. This interface will contain data that, based on the future (i.e. predicted) status of relevant components, and the data from IF2, will contain messages with information about early detected failures.
- **IF9: Maintenance recommendations**. This interface will contain data that, based on the future (i.e. predicted) status of relevant components, early detected failures (IF8), and the data from IF2, will contain messages with information about maintenance recommendations.
- **IF10: External interfaces**. This interface will include data that will summarize the current status of relevant components, the early detection of component failures and associated maintenance recommendations.

The input / output data that will circulate through the internal interfaces is briefly summarized:

- **IF_PANT_DH**: It is used for transmitting condition relevant features and/or raw data as well as status information on sensor performance from the Pant sub-system to SIA_DH.
- **IF_ABA_DH**: It is used for transmitting condition relevant features and/or raw data as well as status information on sensor performance from the SIA_ABA sub-system to SIA_DH.

- **IF_POS_PANT**: Used for synchronisation of SIA- PANT sensors with other sensing nodes within the vehicle.
- **IF_POS_ABA**: Used for synchronisation of ABA sensors with other sensing nodes within the vehicle.
- **IF_POS_DH**: It is used for outputting positioning related measurements or final positions to SIA_DH, depending on the approach taken as discussed above. It also is used to pass on configuration information to the GNSS receiver and/or positioning engine if applicable.
- **IF_DH_VP**: Receives the data in SIA_DH generated by the different subsystems and sends data introduced in the system through SIA_VP required by the rest of subsystems
- **IF_CDM_VP**: Sends input data to SIA_CDM and receives the calculated output data from SIA_CDM.

In every case, the template of datasets from partners acquiring and sharing these data will cover “Origin of the data”, “expected size”, “format”, “reusing of data”.

Every dataset and working document will be properly referenced to a version control number, where major versions are attained only when sending them for a milestone of the task.

The types and formats of data within the project frame include the following:

- **Laboratory data**: datasets (*.txt, *.doc, *.docx,*.xls, *.xlsx, hdf5, SensorML, etc.), multimodal measurements (*.txt, *.doc, *.docx,*.xls, *.xlsx), numerical data (*.XX), qualitative data (*.txt, *.doc, *.docx), data statistics (*.xls, *.xlsx), images (*.jpg, *.png, *.jpeg, *.tiff), videos (*.Mp4, *.mov), geographical information (*.kml, *.gpx).
- **Fusion data**: statistics (*.xls, *.xlsx), graphs (*.ogg, *xls, *.xlsx), bibliography (*.enl), code and executables (*.rpm, *.exe, *.c, *.cpp, *.py, *.java)
- **Scientific texts**: manuscripts and reports (*.doc, *.docx, *.pdf), publications (*.doc, *.docx, *.pdf), conference proceedings (*.doc, *.docx, *.pdf), conference presentations and posters (*.ppt, *.pptx, *.pdf), books and theses (*.doc, *.docx, *.pdf).
- **Operational data**: inspection forms, maintenance sheets and procedures, operations planning, timetables, vehicle-related data, infrastructure-related data, etc. All the documents will come in the form of *.doc, *.docx, *.pdf, *.jpg, *.png, *.jpeg, *.ppt, *.pptx, *.xls, *.csv.)
- **Dissemination material**: leaflets and fact-sheets (*.pdf), images (*.jpg, *.png, *.jpeg, *.tiff), animated images (*.gif), videos (*mp4), social network publications and website (*.html), presentations and templates (*.ppt, *.pptx, *.pdf).
- **Management documents**: deliverables (*.doc, *.docx, *.pdf), patents (*.doc, *.docx, *.pdf).

6 Fair data

6.1 Making data findable, including provisions for metadata

6.1.1 Metadata provision

Zenodo repository offers the possibility to assign several metadata to all uploads, in order to make the content findable. The tags Zenodo offers are:

- Publication type (journal article, presentation, book, thesis, etc.)
- Title, authors, affiliation
- Description of the content
- Communities that the data belong to
- Grants which have funded the research
- Identifiers (DOI, ISSN, PubMed ID, URLs, etc.)
- Contributors
- References

6.1.2 Identifiability of data

Zenodo assigns all publicly available uploads a digital object identifier (DOI) to make the upload easily and uniquely citable. If the upload already has a DOI assigned, it can be detailed in the metadata provision.

All data generated under the SIA project will acknowledge the grant in the following way:

"*This project has received funding from the European Union's Horizon 2020 research and innovation programme and from the European Global Navigation Satellite Systems Agency under grant agreement No 776402*".

Plus, they will automatically be associated to the project via OpenAIRE portal.

6.1.3 Keywords

All uploads will include a group of relevant keywords in order to facilitate the identification of the results.

6.1.4 Clear versioning

Zenodo repository provides a new feature to handle versioning: DOI versioning allows to edit/update the record's files after they have been published, cite a specific version of a record and cite all of versions of a record.

When an upload is published on Zenodo for the first time, Zenodo registers two DOIs:

- a DOI representing the **specific version** of the record.
- a DOI representing **all versions** of the record.

Afterwards, a DOI is registered for every new version of the same upload.

6.2 Making data openly accessible

6.2.1 Which data will be made openly available?

Scientific Publications

Article 29.2 of the model grant agreement sets out detailed legal requirements on open access to scientific publications: under horizon 2020, each beneficiary must ensure open access to all peer-reviewed scientific publications relating to its results. Therefore, all the scientific publications originated by the SIA project will be made openly accessible; gold or green Open Access Publishing.

There is a list of journals with that status in “www.ieee.org/go/journals” following the description “IEEE Hybrid Open Access Journals”.

Other Research Data

Additionally, any other research data or information that might be publishable will also be made openly accessible. However, any dissemination data linked to exploitable results will not be put into the public domain if they compromise their commercialization or have inadequate protection.

6.2.2 How the data will be made available

The open access mandate comprises 2 steps: depositing publications in repositories and providing open access to them. The SIA project will fulfil these two steps by uploading the data to the Zenodo repository.

6.2.3 Methods or Software tools needed to access the data

As a general rule, the format of the data deposited in Zenodo repository will enable the access to them through standard software tools like Adobe Acrobat Reader or Microsoft Office Package.

For data formats that cannot be opened using standard software tools, reliable information on the tools required to validate the results will be provided with the data.

6.2.4 How access will be provided in case there are any restrictions

As detailed in previous sections, any dissemination data linked to exploitable results will not be put into the public domain if they compromise their commercialization or have inadequate protection. In this case, the scientific committee of SIA will individually analyse and decide on the particular access and time restrictions for each result.

6.3 Making data interoperable

SIA will encourage the use of standard vocabularies for all data types present in the data sets to allow inter-disciplinary interoperability. In case this is not possible for a specific data set, project partners will provide mappings to more commonly used vocabulary.

6.4 Data re-use

6.4.1 How the data will be licenced to permit the widest reuse possible

Data re-use is subjected to the license under which it is deposited on Zenodo. The Steering Management Committee (SMC) will decide on the specific license that applies to each data deposited, taking into account the exploitability of the results.

6.4.2 When the data will be made available for re-use

The data will be available for re-use immediately after deposition on Zenodo Repository.

Scientific publications will be uploaded to Zenodo as soon as they are published by the editorial and, at the latest, six months after publication.

Other research data not linked to scientific publication will be uploaded to Zenodo following the instructions of the SMC.

6.4.3 Third-parties and re-usability

The data uploaded to Zenodo, as they are deposited on a free-access base, can be re-used by third parties.

6.4.4 Data quality assurance process

Data quality assurance is performed by the Zenodo Repository. In particular, for each file uploaded, two independent MD5 checksums are stored. One checksum is stored by Invenio and used to detect changes to files made from outside of Invenio. The other checksum is stored by EOS and used for automatic detection and recovery of file corruption on disks.

6.4.5 Period for which the data will remain re-usable

All the files uploaded to Zenodo will remain re-usable for the lifetime of the repository, which is the lifetime of the CERN.

In case of closure of Zenodo, the best efforts will be made in order to preserve the data in an alternative repository.

7 Allocation of resources

7.1 Costs for making data FAIR

The costs for making data FAIR are mainly those related to the cost of Open Access to Scientific Publications, as the use of Zenodo Repository is free of charge.

Costs related to open access to research data in Horizon 2020 are eligible for reimbursement during the duration of the project under the conditions defined in the H2020 Grant Agreement.

7.2 Responsibilities for data management in SIA project

Any member of the Consortium can upload content in the repository taking into consideration committed data quality, naming conventions, etc. (see section 5).

7.3 Costs and potential value of long-term preservation

Long term preservation of SIA open access research data will be based on Zenodo Repository, which is free of charge.

The Steering Management Committee of SIA will decide on long-term preservation of research data associated to exploitable results. This will be done during the project lifetime and based on the protection strategy followed by the consortium.

8 Data security

8.1 Data storage

8.1.1 SharePoint

All the data of the project will be stored in the SharePoint site of the project.

Microsoft Sharepoint uses, according to Microsoft Corporation, some of the strongest, most secure encryption protocols in the industry to provide a barrier against unauthorized access to data. When data is at rest two types of encryption are used: disk encryption and file encryption. On disk encryption level, a BitLocker is used to secure data and on file encryption level every file is secured with its own key that uses Advanced Encryption Standard (AES) with 256-bit keys, which is a Federal Information Processing Standard (FIPS) 140-2 compliant.

8.1.2 Zenodo

8.1.2.1 Data storage

All files uploaded to Zenodo are stored in CERN's EOS service [4] in an 18 petabytes disk cluster. Each file copy has two replicas located on different disk servers.

For each file two independent MD5 checksums are stored. One checksum is stored by Invenio [5] and used to detect changes to files made from outside of Invenio. The other checksum is stored by EOS and used for automatic detection and recovery of file corruption on disks.

Zenodo may, depending on access patterns in the future, move the archival and/or the online copy to CERN's offline long-term tape storage system CASTOR in order to minimize long-term storage costs.

EOS is the primary low latency storage infrastructure for physics data from the Large Hadron Collider (LHC) and CERN currently operates multiple instances totalling 150+ petabytes of data with expected growth rates of 30-50 petabytes per year. CERN's CASTOR system currently manages 100+ petabytes of LHC data which are regularly checked for data corruption.

Invenio provides an object store like file management layer on top of EOS which is in charge of e.g. version changes to files.

8.1.2.2 Metadata storage

Metadata and persistent identifiers in Zenodo are stored in a PostgreSQL instance operated on CERN's Database on Demand infrastructure with 12-hourly backup cycle with one backup sent to tape storage once a week. Metadata is in addition indexed in an Elasticsearch cluster for fast and powerful searching. Metadata is stored in JSON format in PostgreSQL in a structure described by versioned JSONSchemas. All changes to metadata records on Zenodo are versioned and happening inside database transactions.

8.2 Data transfer

All data exchanges will be performed throughout web services based on HTTPS protocol (both SharePoint and Zenodo fulfil this condition).

9 Intellectual Property Rights

9.1 IPR in the Consortium Agreement

The Consortium Agreement (CA) was signed by all project partners and has come into force on the date of its signature by the Parties and shall continue in full force and effect until complete fulfillment of all obligations undertaken by the Parties under the EC-GA and the Consortium Agreement. The purpose of the Consortium Agreement (CA) is to establish a legal framework for the project in order to provide clear regulations for issues within the consortium related to Results and Intellectual Property (IP), Ownership, Confidential Information, Open Source issues, Standard contributions, and Access Rights to Background and Foreground along the duration of the project and any other matters of the consortium's interest.

The CA also covers full rights and responsibilities of participants during the project in respect of the confidentiality of information disclosed by the partners, as well as the publication and communication of information. Moreover, the CA provides additional rules to ensure effective dissemination of the results. Settlements of internal disputes and of course Intellectual Property (IP) arrangements are part of the CA as well.

Any knowledge, information, data and/or IPR generated before the effective date of the CA (i.e. background) shall remain with the respective party providing such background to the project. Any result generated by a party after the said date, during and within the scope of the project (i.e. Result), whether or not it qualifies for Intellectual Property Right (IPR) protection, shall vest in the party that generated such Result. Any jointly generated Result will be jointly owned. The rights and obligations associated to such jointly generated Result will be regulated in the CA, but in any event each joint owner contributing to the cost of such jointly generated Result shall enjoy an unrestricted right to freely use and exploit such jointly generated Result. Throughout the execution of the project, all partners will continuously contribute to the identification of Results that may qualify for IPR protection and will act with the aim of achieving a meaningful outcome for the community following completion of the project.

In case certain results are identified to be essential for the future business opportunities of the involved partners, the necessary steps will be taken to protect such results accordingly. The patenting and other protective measure procedures will proceed along the regulations set forth in the CA.

The IP terms and conditions during the cooperation of SIA will be, a priori, based on a royalty free basis. After completion of the project (i.e. during exploitation) access rights to background and to Results may require fair and reasonable compensation with non-discriminatory conditions, subject to agreement amongst the parties and reflected in the CA.

All access rights needed for the execution of the project and following completion of the project will be granted on a non-exclusive basis, will be worldwide and in principle, will not contain the right to grant sub-license(s). The CA will further regulate rights and obligations for affiliated entities of a party, where those shall enjoy the same access rights conditions as the party participating in

the project, and where such affiliated entities will need to grant requested access rights to other parties if those are needed during execution and/or following completion of SIA.

The CA also provides additional rules on the introduction, pursuant to notification, of background that has been made available under controlled license terms, e.g. so-called open source licenses. To the extent required for proper use of software results, sub-licensing rights on software results are regulated by the CA if it is in the best interest of the project dissemination, where such sub-licensing rights shall not be in a manner where the so licensed software results would be subject to controlled license terms. Means to make software results available to the other parties or to the public are part of the CA.

9.2 IPR Models

Three possible models have been considered during the proposal preparation phase and will be estimated during the project execution phases. The current consensus of the consortium is the retention of title, ownership and exploitation rights of the results and IPR generated on an individual per partner basis as the preferred option, although other no binding options will be explored.

Results - Foreground and IP shall be owned by the project partner carrying out the work leading to such Foreground and IP. If any Foreground and IP is created jointly by at least two project partners and it is not possible to distinguish between the contributions of each of the project partners, such work will be jointly owned by the contributing project partners. The same shall apply if, in the course of carrying out work on the project, an invention is made having two or more contributing parties contributing to it, and it is not possible to separate the individual contributions. Such joint inventions and all related patent applications shall be jointly owned by the contributing parties. Alternative models to be explored, on the basis of unanimous agreement, are:

- *Joint Ownership and Exploitation*: The SIA partners will register title and jointly share the exploitation rights of the project foreground based on the relative share of Person Month effort dedicated to the project.
- *Per Work Package or Tasks Ownership and Exploitation*: In this case, each partner or sub group of partners involved in individual tasks will register title and exploit the results of their own Work Package.
- *Combined Approach to Title and Exploitation:* In this case the SIA partners will register title in the foreground and IPR in line with their own Foreground and IPR policy.

9.3 IPR Officer and the Auditing and Management of Generated Data, Results, and IPR

The SIA proposal has appointed a specialist on IPR. The role of the IPR Officer (IPO) will be to act as an honest broker with the Research and Management staff of the project and provide an objective audit and reporting on the title and ownership of the IPR generated during the project. This will be done via periodic surveys and reports based on the content of deliverables and

partners assigned to the associated tasks. The IPO nominated by the project will conduct Interim IPR Audit will identify the Results generated by the project, its dependencies on and External IP, or Background knowledge, and recommend actions to be taken by the consortium for its protection. CEIT has allocated a specific budget to appoint the IPO. The results of the Interim IPR Audit carried out will be reported internally and the results of the Final IPR Audit carried out at M36 will be reported in the corresponding deliverable. The IPR Officer is Dr. Isabel Hernando, Professor of Civil Law at University of the Basque Country and IP Lawyer.

9.4 Final IP and Generated Foreground Control Report

The Final Report will summarize the results - foreground generated by each partner during the SIA Project. During the periodic Reports all partners will be requested by the Intellectual Property Officer to identify:

- The **access rights** granted to another partner of the consortium of their background, which was needed for the implementation of the project.
- The **access rights** granted to another partner of the consortium of their foreground, which was needed for the implementation of the project.
- The **background** used in the implementation of the project.
- The **foreground** generated in the project.
- The **Party exploitable foreground** generated in the project where are identified the Consortium **Single Products** (SP) per Parties and their Contributions – components.

Moreover, the partners will also identify in these Interim Reports the commercial and open source software as well as the hardware used in the implementation of the project. The Intellectual Property Officer will review all this information and will provide advice about IPR when needed to the consortium. All this information will be requested using the “Result and IP Control Report Template” provided by the IPO and the Coordinator.

Once all the Foreground and IP information has been gathered from the partners, the IPO will carry out an objective audit and will report on the title and ownership of the IPR generated during the project.

Tables will be used to clearly identify the main outcomes of the IPR analysis. Among the foreground generated in SIA during the Project, a **Table 1** will identify those that are exploitable foreground of the consortium classified in five groups (1) for Further Research; (2) for developing, creating and marketing a product/process; (3) for creating and providing a service; (4) for Standardization activities; and (5) For others (as Joint ventures, Spin-off, licensing, etc.). A **Table 2** will provide the possible and recommended IPR qualifications for the identified SIA Exploitable Foreground. This Table will present the list of IPR registration applied or recommended**. A Table 3** will identify the Ownership percentages to the components – contributions to the Single Products and the Intermediate Single products developed in the SIA Project and finally, for the Exploitable

Foreground, a **Table 4** will identify the Parties ownership percentages to the Final Project Single products.

9.5 Exploitation and Ownership Agreement

Title and ownership of results is considered to be, from a legal perspective, a matter of undisputable fact. The IPR audits will be presented by the IPO to each partner as a proposal of ownership of results according to data provided by the Partners and derived from the control of technical outputs and deliverables for revision and acceptance by their organizations.

At the end of the SIA project, these audits will form the basis for a potential Exploitation and Ownership Agreement (EOA) if any. This agreement will not be a formal project deliverable and is a private contract between partners in the same level as the Consortium Agreement. The Exploitation and Ownership Agreement will be a binding legal contract between the partners which will be negotiated and approved by authorized representatives within the partners' organizations.

10 Conclusions

This document presents the internal guidelines that will be followed for the appropriate data and privacy management of the SIA project. Some of the sections in this document will be updated throughout the lifetime of the project, as previously indicated, in order to appropriately address the practical requirements of the project. The overall data and privacy management plan of the project described in this deliverable is aligned with the information already provided in the Description of Action for SIA (as per Grant Agreement number 776402).

11 References

[1] SIA Project Deliverable D1.1. Project Management and Quality of Assurance Plan
[2] Zenodo repository. https://zenodo.org
[3] SIA Project Deliverable D2.2. SIA Architecture and preliminary validation plan
[4] CERN's EOS service. http://information-technology.web.cern.ch/services/eos-service
[5] INVENIO Software. https://invenio-software.org