D1.2
Ethics Framework and Data Management Plan

Topic	H2020-SU-ICT-2018-2020
Project Title	Artificial Intelligence-based Cybersecurity for Connected and Automated Vehicles
Project Number	833611
Project Acronym	CARAMEL
Contractual Delivery Date	M06
Actual Delivery Date	M06
Contributing WP	WP1
Project Start Date	01/10/2019
Project Duration	30 Months
Dissemination Level	Public
Editor	8Bells
Contributors	8Bells, I2CAT

Document History		
Version	Date	Remarks
0.1	7/01/2020	ToC, Introduction
0.2	14/01/2020	Ethical Considerations
0.3	22/01/2020	DMP Legal Framework
0.4	11/02/2020	Data management Plan
0.5	02/03/2020	Restricted Deliverables Procedure
0.6	12/03/2020	Review, informed consent and annexes
0.7	17/03/2020	Final draft ready
0.8	19/03/2020	Ready for SAB review
0.9	26/03/20	SAB review received
1.0	26/03/20	Final version ready

DISCLAIMER OF WARRANTIES

This document has been prepared by CARAMEL project partners as an account of work carried out within the framework of the contract no 833611.

Neither Project Coordinator, nor any signatory party of CARAMEL Project Consortium Agreement, nor any person acting on behalf of any of them:

- makes any warranty or representation whatsoever, express or implied,
    - with respect to the use of any information, apparatus, method, process, or similar item disclosed in this document, including merchantability and fitness for a particular purpose, or
    - that such use does not infringe on or interfere with privately owned rights, including any party's intellectual property, or
- that this document is suitable to any particular user's circumstance; or
- assumes responsibility for any damages or other liability whatsoever (including any consequential damages, even if Project Coordinator or any representative of a signatory party of the CARAMEL Project Consortium Agreement, has been advised of the possibility of such damages) resulting from your selection or use of this document or any information, apparatus, method, process, or similar item disclosed in this document.

CARAMEL has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 833611. The content of this deliverable does not reflect the official opinion of the European Union. Responsibility for the information and views expressed in the deliverable lies entirely with the author(s).

DISCLOSURE STATEMENT

"The following document has been reviewed by the CARAMEL External Security Advisory Board as well as the Ethics and Data Management Committee of the project. Hereby, it is confirmed that it does not contain any sensitive security, ethical or data privacy issues."

Table of Contents

List of Figures

List of Tables

List of Acronyms

CA	Consortium Agreement
CMO	Collective Management Organisation
DMP	Data Management Plan
EDC	Ethics and Data Management Committee
EDM	Ethics and Data Manager
FAIR	Findable, Accessible, Interoperable and Re-usable
GA	Grant Agreement
IM	Innovation Manager
PSO	Project Security Officer
SAB	Security Advisory Board

Executive Summary

This deliverable consists of two main sections after the introduction. Chapter 2 describes procedures implemented for CARAMEL to ensure compliance with the Ethical Considerations in the project. Chapter 3 describes the management of data gathered in the scope of the project (Data Management Plan) provisions taken in CARAMEL for the data sharing, protection and exploitation. The Legal Framework is addressed from the perspective of Open Access, Open Data, and IPR restrictions. In addition, the Data Management Plan defines how the data that the project will generate will be stored, localized, preserved and in which cases will be publicly disclosed.

1 Introduction: Ethical Considerations and Data Management in CARAMEL

1.1 *Purpose and Scope*

The ethical aspects in the CARAMEL project are limited to potential interaction with data subjects as part of relevant research and in the use cases and in dealing with personal data in general.

The project addresses by:

- An ethical assessment and an assessment if any of the planned activities require ethical opinion, authorization or confirmation;
- An ethics report explaining the result of this exercise and providing all necessary documents;
- A data management plan, which includes best practice approaches and policies to be implemented by the consortium. The data management plan is to be updated in iterations;
- For any direct interaction with natural persons which require the collection of personal data, this will be justified, and any collection will be accompanied by a specific privacy notice and if based on consent a consent form, specifically drafted for the CARAMEL project by the partner 8BELLS.

Data Management Plan (DMP) is a written formal document that describes how data will be handled until the completion of the project and after it. The Guidelines on Findable, Accessible, Interoperable and Re-usable (FAIR). Data Management in Horizon 2020 [6] provide a set of principles and criteria that have to be addressed. Research data should become FAIR. The CARAMEL DMP will describe in detail the data that the project will collect/generate, the methodologies and standards that will be followed to make research data FAIR, the data that will be shared/made open, and how they will be curated and preserved during and after the lifetime of the project.

Data sharing in the open domain can be very beneficial to society, however, we need to balance openness on the one hand and protection of sensitive data on the other hand. As stated in the Guidelines on FAIR Data Management [6] data should be ‘as open as possible and as close as necessary’. All data providers that participate in the consortium should comply with all applicable data protection or similar laws regulating the processing of any personal data.

Status of the Document

This deliverable ensures all ethical issues and European Charter of Fundamental Rights are covered (especially in relation to data protection) and all necessary confirmations/authorisations have been obtained. Following this deliverable (D1.2) on ethics in M6 there will be updates in D1.4 in M12 and D1.5 in M21, but ethic issues will be further controlled and implemented throughout the project. Additionally, this deliverable (D1.2) focuses on proper data management (including data protection aspects) within the project outputting a data management plan.

2 Ethical Considerations

The following sections present an overview of the research ethical framework. The ethical aspects in the CARAMEL project are limited to potential interaction with data subjects as part of relevant research and in the use cases and in dealing with personal data in general. It also describes how ethical procedures are approached in H2020 projects, and more specifically in CARAMEL.

2.1 *Research Ethical Framework*

Ethics is an integral part of research and is given a high priority in EU funded research [1]. CARAMEL will comply with existing regulations and codes of conduct. Some of the most relevant documents are the following.

2.1.1 Charter of Fundamental Rights of the European Union

This document [2] gathers the fundamental rights to be shared, fostered and protected by every Member State of the European Union. The first draft was created by the European Convention in 2000 and was solemnly proclaimed by the European Parliament, the Council of Ministers and the European Commission during the same year. However, it was not legally binding until the entry into force of the Treaty of Lisbon, on 1st December 2009. The Charter contains 54 articles divided in seven titles: dignity, freedoms, equality, solidarity, citizens' rights, justice and general provisions governing the interpretation and application of the Charter. This Charter must be abode by Member States when applying European Union law.

The Charter sets the starting point for any research or action conducted within the context of the European Union. Every article needs to be taken into consideration in order to develop a study within an ethical framework, such is the case of any project supported and funded by the European Union. There are certain specific articles that are of high importance when developing the methodology to conduct a research in Social Science. For example, Article 8, Title II (European Parliament, Council and Commission, 2012), on Protection of personal data, which literally states that:

- Everyone has the right to the protection of personal data concerning him or her.
- Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified.
- Compliance with these rules shall be subject to control by an independent authority.

Any action taken within CARAMEL project needs to be compliant with all fundamental rights preserved in this Charter.

2.1.2 European Code of Conduct for Research Integrity

The Code of Conduct for Research Integrity was created by the European Federation of Academies of Sciences and Humanities and has been recently revised and republished in 2017 [3]. This document contains a set of rules to self-regulate academic research through European territories and it is designed to be used across all scientific fields, without distinction. It includes the principles to preserve research integrity, a list of good practices and some guidelines about violations of research integrity (the most serious being fabrication, falsification and plagiarism) and procedures to be followed in the event of those violations.

According to this Code of Conduct [3], the principles to preserve research integrity:

- Reliability in ensuring the quality of research, reflected in the design, the methodology, the analysis and the use of resources.
- Honesty in developing, undertaking, reviewing, reporting and communicating research in a transparent, fair, full and unbiased way.

- Respect for colleagues, research participants, society, ecosystems, cultural heritage and the environment.
- Accountability for the research from idea to publication, for its management and organisation, for training, supervision and mentoring, and for its wider impacts. The document describes good research practices in various contexts: research environments; training, supervision and mentoring; research procedures; safeguards; data practices and managements; collaborative working; publication and dissemination, and reviewing, evaluating and editing.

Among other good practices and recommendations, the document establishes:

“Researchers handle research subjects, be they human, animal, cultural, biological, environmental or physical, with respect and care, and in accordance with legal and ethical provisions.” [3: 6]

An explicit mention of ethical practices is also made under the section “Violations of Research Integrity”, where the document states: “It is of crucial importance that researchers master the knowledge, methodologies and ethical practices associated with their field. Failing to follow good research practices violates professional responsibilities. It damages the research processes, degrades relationships among researchers, undermines trust in and the credibility of research, wastes resources and may expose research subjects, users, society or the environment to unnecessary harm.” [3: 8] This document is especially important for all researchers participating in Horizon 2020 funded projects, since it has become a reference document.

2.2 *Ethics in H2020 projects*

Ethical compliance is seen as fundamental in research projects funded by the European Union. As explained in Research Ethics [1], ethics is dealt with in the Horizon 2020 legislation at various levels. There are also a specific Ethical Appraisal Procedure in Horizon 2020 projects. The Horizon 2020 Rules for Participation [4] determine that proposals cannot contravene ethical principles and that the Commission shall systematically carry out ethics reviews for proposals (Article 14).

The Horizon 2020 Regulation of Establishment [5], establishes in Article 19 (Ethical principles) that:

“All the research and innovation activities carried out under Horizon 2020 shall comply with ethical principles and relevant national, Union and international legislation, including the Charter of Fundamental Rights of the European Union and the European Convention on Human Rights and its Supplementary Protocols. Particular attention shall be paid to the principle of proportionality, the right to privacy, the right to the protection of personal data, the right to the physical and mental integrity of a person, the right to non-discrimination and the need to ensure high levels of human health protection.”

More specific dispositions are found in the Grant Agreement, as described in the next section.

2.2.1 Grant agreement dispositions

Article 34 in section 4 of Grant Agreement no 833611 establishes the obligations concerning “Ethics and research integrity”. Its subsections discuss (1) obligation to comply with ethical and research integrity principles, (2) activities raising ethical issues, (3) activities involving human embryos or human embryonic stem cells, and (4) consequences of non-compliance (Grant Agreement, pp. 52-54). Summaries of the respective subsections are provided below, but project participants are encouraged to read the relevant sections in the Grant Agreement. Article 34.1. puts forward the obligation of the beneficiaries to carry out the action in compliance with (1) ethical principles, and (2) applicable international, EU and national law. It also stresses when funding will not be granted.

It addresses the question of research integrity, already discussed above, and enumerates principles that the beneficiaries must respect, namely: honesty, reliability, objectivity, impartiality, open communication, duty of care, fairness, responsibility for future science generations. It also discusses how these values should be implemented while conducting research activities (see Grant Agreement, page 53). Article 34.2 tackles the question of activities raising ethical issues. It states that they need to comply with ethical requirements considered as deliverables presented in the Annex 1 (description of the action) of the Grant Agreement. It also states obligations before beginning any activity raising an ethical issue, namely obtaining ethics committee opinion required under national law or obtaining notifications or authorisations required by national/European law (see Grant Agreement, page 53). These documents must be kept on file and submitted if requested, as indicated on Grant Agreement (page 54). Article 34.4 refers to the consequences of non-compliance, which may result in reduction, termination of the grant, or other measures (Grant Agreement, page 54). Article 41.2 also establishes that each beneficiary must submit to the coordinator in good time the "ethics committee opinions and notifications or authorisations for activities raising ethical issues" (Grant Agreement page 61).

2.3 *CARAMEL Ethical Procedures*

Ethics and Data Management Committee (EDC) To guarantee the respect and the constant monitoring of ethics and data management issues, a specific committee was appointed. The committee is composed of three (3) CARAMEL partners (8Bells, T-SYS, Panasonic). The CARAMEL EDC will help to ensure that potential data gathering procedures are done on the basis of consent forms that follow Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. The board will also monitor compliance with the requirements regarding ethical, privacy and data protection issues throughout the project lifetime and will assess the sensitivity of all deliverables before any publication and will review progress regularly assuring constantly an appropriate classification level. The EDC is chaired by the project coordinator. In accordance with the 'data minimisation 'principle, all of the data the Consortium intends to process is relevant and limited to the purposes of the research project.

2.3.1 Informed Consent Procedures

The corresponding national data protection legislation and the Regulation (EU) 2016/679 - General Data Protection Regulation should be taken under consideration, and all legal documents and certifications required for compliance with such legislation will be obtained. The Parties who provide or make available to any other Party shared information containing Personal Data must have: (i) the authority and/or the authorisation to disclose the aforementioned information; (ii) obtained appropriate informed consents from all the data subjects involved, or from any applicable institution and (iii) a confirmation that there is no restriction that would prevent any other Party from using the shared information. This deliverable includes to annexes with templates about informed consent.

2.4 *Personal data protection: EU regulations*

Data protection regulations In CARAMEL are in line with GDPR. The project takes on board EU data protection policies following the Regulation (EU) 2016/679 - General Data Protection Regulation , and also national policies for the three countries where tests will be performed: The German Federal Data Protection Act (i.e. BDSG) or the Spanish Organic Data Protection Law 15/1999, and the Data Protection Act 1998 for the UK

2.5 *Video Recording rights*

The European Union adopted in February 2014 the Directive on collective management of copyright and related rights and multi-territorial licensing of rights in musical works for online uses in the internal market (CRM Directive). The CRM Directive is an essential part of Europe's copyright legislation. The CRM directive (2014/26/EU) aims at ensuring that right holders have a say in the management of their rights, and at improving the functioning and accountability of Collective Management Organisations (CMOs). CARAMEL will follow CRM directive using exclusively videos for testing produced by the project partners, as well as other open access audio/visual content.

In the framework of the PANASONIC demonstration, PANASONIC offer a test vehicle equipped with sonar sensors, LIDAR sensors, cameras, DGPS navigation system and embedded computers, some people can be recorded. PANASONIC will be in charge to collect the informed consent. This form will be stored for the duration of the project and five years after the end of the project.

Panasonic
AUTOMOTIVE

CESSION OF IMAGE RIGHTS

By the present, the signing ____, of legal age and holder of the identity document ________________ authorizes PANASONIC AUTOMOTIVE so that it can capture and reproduce by photography or video its image, for the internal use in the CARAMEL project framework and to share these images with the CARAMEL consortium.

For this purpose, I authorize the express and free exploitation of my image rights in favour of the CARAMEL consortium, for the use of them in the terms and conditions set forth above, expressly renouncing to request or claim anything from the CARAMEL consortium for this purpose.

This authorization and assignment of my image rights will remain in force, regardless of vicissitudes, until the termination of the CARAMEL project.

My authorization refers exclusively to the aforementioned object or purpose, allowing the use of the technical means known thus far and those that could be developed in the future. All that with the sole exception and limitation of those uses or applications that could violate the GDPR regulations.

And for that purpose, I hereby issue this AUTHORIZATION and FREE CESSION of my IMAGE RIGHTS on the site and date that appear in this document.

Signed:

3 Data management

A Data Management Plan (DMP) is a written formal document that describes how data will be handled until the completion of the project and after it. The Guidelines on FAIR Data Management in Horizon 2020 [6] provide a set of principles and criteria that have to be addressed. The CARAMEL DMP will describe in detail the data that the project will collect/generate, the methodologies and standards that will be followed to make research data FAIR, the data that will be shared/made open, and how they will be curated and preserved during and after the lifetime of the project.

3.1 *Legal framework*

The three points to consider when dealing with data in CARAMEL are: The Open Access of publications, Open Data, and the IPR restrictions included in the Grant Agreement and the Consortium Agreement. As a general rule partners must take into account that unless it goes against the legitimate interest of the beneficiaries the results must be disseminated by disclosing them. This means that the beneficiaries have the right to protect the results in case the institution plans to protect or exploit the results.

The dissemination and publication of visual data recorded by the vehicle in CARAMEL, is possible only after Panasonic provides written consensus.

3.1.1 Open Access to Scientist Publications

Open access (OA) refers to the practice of providing, free of charge to any user, online access to all peer-reviewed scientific information and all the research data. In consequence as indicated in article 29.2 of the Grant Agreement the members of the consortium must:

- As soon as possible and at the latest on publication date, deposit a machine-readable electronic copy of the published version or final peer-reviewed manuscript accepted for publication in a repository for scientific publications; Moreover, the beneficiary must aim to deposit at the same time the research data needed to validate the results presented in the deposited scientific publications.
- Ensure open access to the deposited publication — via the repository — at the latest:
  - On publication, if an electronic version is available for free via the publisher,
  - Or within six months of publication (twelve months for publications in the social sciences and humanities) in any other case.
- (c) Ensure open access — via the repository — to the bibliographic metadata that identify the deposited publication.

To ensure open access, it has been agreed to use OpenAIRE [7] to link different Zenodo [8] repositories.

It is worth noting that article 29.3 of the GA regarding Open access to research data is not applicable to CARAMEL project.

3.1.2 IPR commitments in CARAMEL

Apart from the commitments regarding the Open Access and the Open Data, the other main commitments linked with the IPR included in the GA and the CA are the following:

- The GA defines in article 26 the ownership of the project results. These results are any tangible or intangible output of the actions including data and information and are owned by the institution that generates them. In case the results are generated by two or more institution rules defined in article 26.2 of the GA and article 8.2 of the CA must be applied.

- As defined in the article 8.4 of the CA 45 calendar days prior to any publication notice must be given to the other parties. Objections must be raised in writing at least 30 days after the receipt of the notice.

3.2 *Data Management Plan*

The Data Management Plan (DMP) defines how the data that the project will generate will be stored, localized, preserved and in which cases will be publicly disclosed. This DMP follow the guidelines of the Digital Curation Center (www.dcc.ac.uk) on how to implement DMP and the recommendations of the EC as published in the Guidelines on FAIR [6].

3.2.1 Data Summary

During the lifetime of CARAMEL project several datasets form various consortium members, representing different domains, will be produced. CARAMEL does not foresee participation to the Open Data Pilot. For every public dataset used, the consortium will provide a valid license or proof of ownership. The Innovation Manager (IM) and the Ethics and Data Manager (EDM) will be responsible to maintain research integrity in terms of Privacy and Data Protection. CARAMEL does not parse personal or sensitive data and deep content inspection is not foreseen. CARAMEL will remain compliant with GDPR, and the e-Privacy and NIS directives in its design. The consortium, however, recognises that there might arise a need to openly publish other data generated by the project. In such a case, the consortium will prioritise the use of standardised and interoperable data and file formats when possible. In cases of unstructured, non-standardised data, the consortium will implement a data format to be published as a CARAMEL specification. Data sets will be described by: (i) Set Reference, (ii) Set Name, (iii) Set Ownership, (iv) Set visibility and sharing, (v) Set description, (vi) Standard/Information Format, (vii) Set archiving, curation & preservation mechanism. The Consortium will try to make some data publicly available in order to promote a broader commercial and scientific exploitation. We will support the publication of data in open data portals, i.e., portals that operate in national, regional or municipal level (e.g., http://open-data.okfn.gr), EU data portals (e.g., https://open-data.europa.eu/en/data). The data, however, will not be made publicly available if there are issues with Security, Confidentiality, Privacy and/or Data Protection.

The datasets to be generated are listed below in Table 1. It is possible, as the project evolves, this table to be modified with additions or removals of datasets. The potentially updated table will be presented in the next versions of the DMP.

Name of the Dataset	Responsible Partner	Historical data or Generated though the project	Accessibility
Threat Analysis and Cyber-Threats	ATOS	Generated	Confidential
GPS Data	UCY, UPAT	GPS DATA GENERATED FROM CARLA SIMULATOR	Public
Charge Detail Records	GFX	Historical data	Confidential
Meter Values	GFX	Historical data	Confidential
Lyft	UPAT	Historical data	Public
Image Data	All partners	Image data generated from CARLA simulator	Public

Table 1: Datasets to be generated during CARAMEL lifetime

As we can see in the Table 1, the aforementioned datasets are divided into two categories regarding their accessibility, namely Confidential and Public. Confidential datasets will be only shared or become accessible after a proper agreement is signed.

A more detailed description of each dataset follows.

Dataset Tables

Name of the dataset	GPS Data
Reference	http://carla.org/
Ownership	CARLA: An Open Urban Driving Simulator, Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, Vladlen Koltun; PMLR 78:1-16
Visibility and Sharing	PUBLIC
Description	CARLA Simulator will generate trajectories of vehicles. On a set of specific simulated paths, a dataset of GPS locations in normal and abnormal conditions will be generated.

Table 2: GPS data

Name of the dataset	Image Data
Reference	http://carla.org/
Ownership	CARLA: An Open Urban Driving Simulator, Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, Vladlen Koltun; PMLR 78:1-16
Visibility and Sharing	PUBLIC
Description	CARLA Simulator will generate labelled image data from virtual sensors located on virtual cars. The virtual landscape contains traffic signs, some them being modified to attack image sensors. Additionally, there might be maliciously modified lane markers.
Data types	Image data
Data formats	Common image data formats such as JPEG, PNG
Archiving, curation & preservation mechanism	CARAMEL project web site

Table 3: Image data

Name of the dataset	Charge Detail Records
Ownership	GreenFlux
Visibility and Sharing	Confidential
Description	CDRs give a summary of an EV charging session. It contains information such as start and stop time, chargepoint ID, user ID and transferred energy.
Data types	char, int, datetime
Data formats	CSV

Table 4: Charge detail records

Name of the dataset	Meter Values
Ownership	GreenFlux
Visibility and Sharing	Confidential
Description	MVs are in session meter-readings of consumed energy and/or amperage
Data types	int, datetime, float
Data formats	CSV

Table 5: Meter values

Name of the dataset	Lyft
Reference	https://level5.lyft.com/dataset/
Ownership	Lyft Level 5
Visibility and Sharing	Public
Description	It includes over 55,000 human-labeled 3D annotated frames, a drivable surface map, and an underlying HD spatial semantic map to contextualize the data.
Data types	jpeg, bin
Data formats	nuScenes format from https://www.nuscenes.org/data-format

Table 6: Lyft dataset

3.2.2 Purpose of the Data Collection & Relation to the Objectives of the Project.

The data generated in the project will be created through pilots, workshops and publications, and is related with the project objectives in the following way:

- Objective 1: To identify cybersecurity threats and vulnerabilities in the context of cooperative, connected and automated mobility vehicles (including electrical plug-in vehicles).

  CARAMEL identifies all critical assets, targeting and exploiting CCAM, and classifying attacks for facilitating the risk analysis. The classification will consider the attackers' security certification, their motivation, the range of the attack, etc.

- Objective 2: To design and develop a successful extensible, scalable and market-oriented cybersecurity architecture for the provisioning of situational awareness in CCAM vehicles.

  CARAMEL analyses the cyberthreats to provide considerably greater trust and reliability, as a necessary step for the cooperative exchange of data.

- Objective 3: To model cyberthreats, detect cyberattacks and to identify appropriate responses for each modern vehicle category considered in CARAMEL.

  CARAMEL will deploy a variety of automotive threat modelling activities tailored to the threats affecting the cyber security posture of connected vehicles.

- Objective 4: To consider and fuse different data sources of information in order to achieve contextual and situational awareness and to facilitate the decision-making process.

The provision of an adequate level of reliability permits the identification of conflicting information and reduces uncertainty. CARAMEL will combine existing tools and modules offered by the partners to develop different protection strategies.

- Objective 5: To create an anti-hacking device that will be able to disable higher level functions in case of a cyberattack.

  Vehicle networks enable connectivity, which makes them more vulnerable and prone to attacks. For this reason, CARAMEL will develop an embedded antihacking device, which together with a backend cloud, will collect in a secure and privacy-aware manner, and process large amounts of data through ML methods running on a big data cluster. On-board networks will be checked for anomalies, and then the platform will update itself step-by-step with modern ML methods for detecting intrusions.

- Objective 6: To perform penetration testing, validate and demonstrate the CARAMEL solution.

  Penetration testing in CARAMEL will be designed to assess the CCAM vehicle cybersecurity and to use testing tools to simulate real-world attack scenarios in order to discover security gaps that have not been addressed at a satisfactory level, potentially leading to injection of fake messages, map database poisoning, denial of services, vehicle’s location tracking, and so on.

- Objective 7: To ensure the long-term success of the project through standardisation and dissemination in commercial and industrial fora and by exploring synergies with other EU initiatives and projects.

  CARAMEL will focus on using all available means to disseminate and communicate the achievements, the results and the knowledge acquired during the project. In this direction.

- Objective 8: To design and develop a successful market-oriented, AI-driven and extensible cybersecurity framework for modern vehicles / Business Models to exploit findings in partners’ future product/service portfolios.

  CARAMEL will explore the means to deliver its achievements to the market, with emphasis in the quantification of benefits, especially in terms of total cost of ownership, revenues and profits.

3.2.3 Type and format of the data

The “Deliverables” are expected to consist of text, image and presentation data; the “Content” includes text, datasets, images, presentations, publications and software. As far as possible CARAMEL partners will use non-proprietary and open formats with documented standards. Format selection will favour the formats used by the project partners as well as the research community interested in the results.

Some of the expected datasets are:

- Deliverable documents, expected size: small
- Software components source code, expected size: large
- Simulation data

3.3 *Data localization and metadata*

3.3.1 Platforms for Data Storing

The project will use Zenodo [7] as platform for storing and managing the data generated and OpenAire [8] for linking the databases and publications. In addition, the project will disseminate through its website (https://www.h2020caramel.eu/) and the social media the public data.

3.3.2 Naming Convention

Each set of data produced will be named in a uniform way and will include a table with a version control. For deliverables:

Dx.y - [Name of the deliverable as described in the GA]

- "x" - Work package assigned to the deliverable
- "y" - Number of deliverable within the work package e.g.: D1.2 Ethics Framework and Data Management Plan

For datasets:

WP [Work Package number] P [Use Case number; Use Case activity number] - [description of the activity] i.e.: WP2 UC2.1 - Report on Threat Analysis and Cyber-Threats.

3.4 *Identification of the Data Stakeholders*

Data stakeholders within the project can come from many different backgrounds and with different intentions. Depending on their importance to the project's execution, decisions will be made by the GA of CARAMEL in order to increase ease of access. The most important and easily identifiable stakeholders are:

- Project partners
- European Commission
- Research community
- Cybersecurity AI, ML experts
- Platform as a Service Providers
- Open Source Communities
- Standardisation organisations

3.5 *Making data openly accessible*

CARAMEL has the objective of making data available publicly as much as possible. In this sense an initial array of datasets that will become available has been identified (see below). Its publication will be evaluated taking into account the possibility of exploiting the data, and the ethical considerations described in the first part of this document.

In this regard, the project will progressively distinguish between:

- **Internal datasets** that will not go public and that either can shared in the consortium or not. This decision will be taken on a case by case basis.
- **Open datasets** that will be provided with full access to the project results, allowing academia to reuse them.

3.5.1 Access to restricted / private datasets.

The security scrutiny process performed by EC experts during the proposal evaluation lead to the classification of a document in CARAMEL project – D2.2 Report on Threat Analysis and Cyber-Threats due at M5 (February 2020) as RESTREINT UE/EU RESTRICTED. This deliverable will report in detail on foreseen cyber-threats aligned with the proposed use cases. Thus, this deliverable will define in detail the cyber-security situation accompanied with more probable attack vectors.

The COMMISSION DECISION (EU, Euratom) ref 2015/444 [9] of 13 March 2015 establishes the general security rules for protecting European Union Classified Information (EUCI)

This guideline introduces the basic rules, modus operandi and established procedures in place in CARAMEL project to deal with the elaboration of the D2.2 deliverable (EU-RESTRICTED).

3.5.1.1 *SAB and PSO*

Following the recommendation to Security Scrutiny, CARAMEL project has established a Security Advisory Board (SAB) in order to manage the security dimension and ensure the proper management of the EU classified information according to the security rules.

The project has also appointed Mr. Pedro Soria (ATOS) as Project Security Officer (PSO). He has wide experience in cybersecurity thread analysis and detection, privacy aspects and dealing with EU-classified information. His short CV is included in the CARAMEL Grant Agreement.

The designed SAB together with the PSO will provide the consortium with expert advice in the security treatment according to EU directives and under the WP1 Project Management. The SAB may review the content of the document before its distribution avoiding a potential misuse of the information generated by the project. The SAB is originally composed for 3 experts whose short CV is included in the Grant Agreement.

- Mr. José María Blanco Navarro (Spain)
- Dr. Charalambos Sergiou (Cyprus)
- Mr. Michal Choras (Poland)

3.5.1.2 *Contributors registry*

The classified EU-RESTRICTED Deliverable D2.2 on Threat Analysis and Cyber-Threats is due at M9 (June 2020). The elaboration of this deliverable is framed on task 2.2 where the contributors are ATOS (lead editor), I2CAT, 8BELLS, ALTRAN, FICOSA, SID, CLS, 0INF, UBIWR, AVL, PANA, UPAT

The first necessary measure is to clearly identify the concrete people that will contribute to this task. A contributor registry will be updated according to project needs.

Partner	Position
ATOS	Head of Cybersecurity Lab at R&D department
ATOS	Project engineer and ECSO representative
ATOS	Senior project engineer
ATOS	Cybersecurity researcher
I2CAT	Project coordinator
8BELLS	Principal analyst
ALTRAN	IT Security consultant
FICOSA	Software project manager
SID	Software engineer
CLS	Senior researcher
0INF	Senior developer
0INF	Project manager

UBIWR	NVF/SDN researcher
AVL	Security engineer
PANA	Senior researcher
UPAT	Senior researcher
T-Sys	IT security consultant

Table 7: List of D2.2 Contributors

These selected team members will be the only authorised people to work on the document. ATOS will distribute the necessary encryption keys to exchange and work on this D2.2 classified deliverable. Part of the security key will be sent of regular post in double envelopes (internal sealed envelope and external opaque envelope). Please note that the names of the individuals in the above table have been omitted for the needs of this deliverable for confidentiality reasons. These team members must read the guidelines and follow the procedures indicated in next chapters by the PSO. Each contributor should guarantee that they have the necessary resources and facilities to deal with EU-classified information according to each National Security Authority (NSA) of its member state.

3.5.1.3 *Security concepts in the project*

The project is establishing security measures for protecting the information exchange via electronic means during the elaboration of the deliverable and possible dependences in other deliverables as well.

These procedures are basic principles and minimum standards to protect European Union Classified Information (EUCI) under tag RESTREINT UE/EU RESTRICTED only and they are not enough for other classified cases such as CONFIDENTIEL UE/EU CONFIDENTIAL or TRES SECRET UE/EU TOP SECRET which are more strict and require additional mechanisms (certification of authorised people, used equipment and separate work zones under controlled access).

The procedures consider basic concepts such as:

- The Need-to-Know

  The term need-to-know refers to the need of an individual to have access to EUCI in order to be able to perform a professional function or task. Persons may only be granted access after their need-to-know has been determined. The need-to-know is central to protecting EU and Member State interests and is applicable to H2020 projects.

- Awareness of the Security Rules

  Individuals may only be granted access to EU RESTRICTED information after they have been briefed on the security rules for protecting EUCI (and have acknowledged their responsibility with regard to protecting such information). The holder of any EU RESTRICTED information is responsible for protecting it and for handling it in accordance with the rules and procedures set in the project Grant Agreement and the Consortium Agreement. These rules are in line with the data management plan of the project.

- Originator consent

  The principle of originator consent applies to all EU RESTRICTED information originated for H2020 research. Responsibility for classifying, declassifying or downgrading such information rests solely with the originator. The European Commission/REA are the originators of all information created during H2020 research project.

- Continued protection

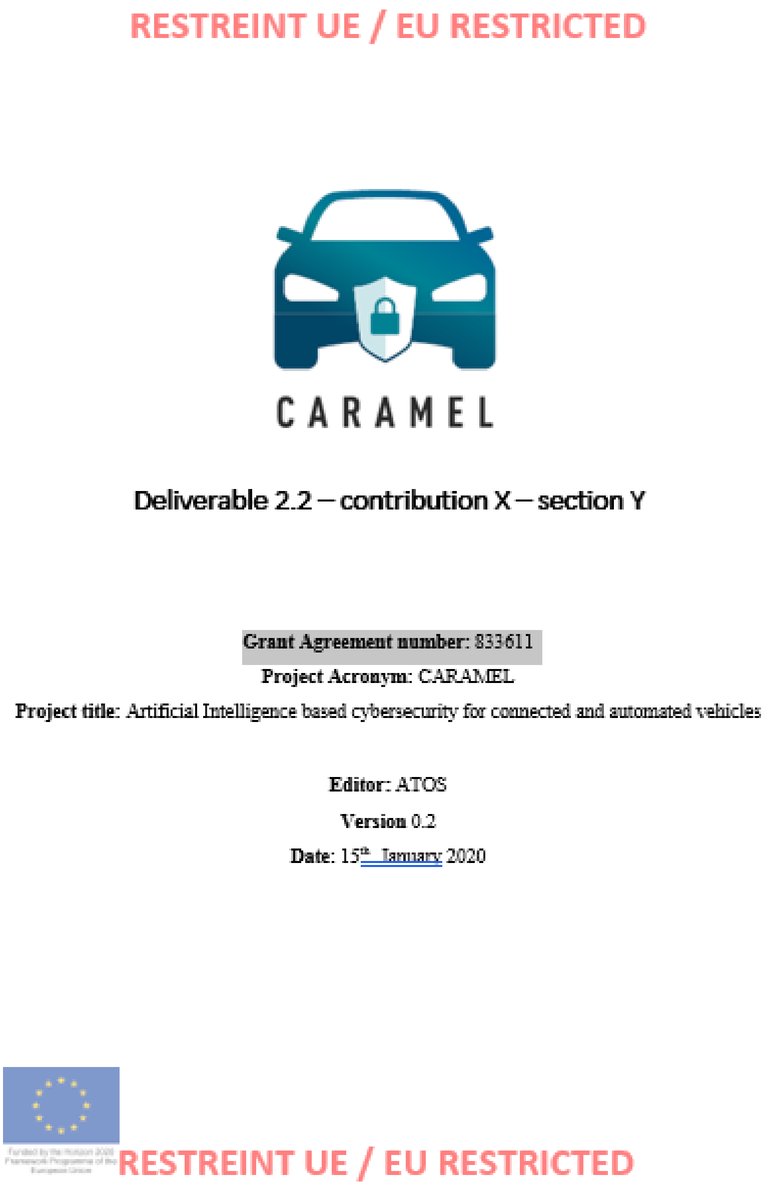
  All EUCI provided under the grant agreement must continue to be protected in the event of the termination of the grant agreement. Compromise of EUCI occurs when, as a result of a breach of security, it has wholly or in part been disclosed to unauthorised persons, in such case the PSO and the coordinator must be alerted immediately to be in communication with the EC officer and national authority.

3.5.1.4 *Working on classified deliverable and information exchange*

D2.2 deliverable must be treated according to RESTREINT UE / EU RESTRICTED handling rules, also when they are works-in-progress.

Marking and storage of paper copies and equipment

D2.2 contributions (pieces of document, diagrams, text, etc) are classified as EU RESTRICTED and any (intermediate) files have to be clearly marked top and bottom, as shown in next figure. When not in use by authorized personnel, hardcopies of documents or portions classified as EU RESTRICTED must be locked down in a filing cabinet that satisfies national rules for storing documents of equivalent classification of EU RESTRICTED.



RESTREINT UE / EU RESTRICTED

CARAMEL

Deliverable 2.2 – contribution X – section Y

Grant Agreement number: 833611

Project Acronym: CARAMEL

Project title: Artificial Intelligence based cybersecurity for connected and automated vehicles

Editor: ATOS

Version 0.2

Date: 15th January 2020

Funded by the Horizon 2020 Framework Programme of the European Union

RESTREINT UE / EU RESTRICTED

Figure 1: CARAMEL D2.2 – EU Restricted Cover

All equipment (i.e. laptops) used to create, edit or view information classified as EU RESTRICTED must be also stored in the same manner in a cabinet with a key as documents classified as EU RESTRICTED.

NOTE: The Quality Assurance process as outlined in the Quality Assurance Plan of the project will also apply to this D2.2 classified deliverable, but only partners with a specified Need to Know will be involved in the QA process. Partners involved with the QA process for classified deliverables will need to have access to a computer approved for EU RESTRICTED.

Requirements for standalone computer equipment

Each partner involved in the production of classified deliverables has to procure (at least) one computer approved for RESTREINT UE / EU RESTRICTED, as indicated by NSA for securing stand-alone computer systems.

Minimum requirements for stand-alone computer equipment include at least the following (some NSA may have additional requirements):

- Not connected to network
- No active wireless connections
- No active built-in camera or microphone
- Full disk encryption
- No network printers
- 64-bit processor
- BIOS password
- The equipment must be labelled "RESTREINT UE / EU RESTRICTED"
- Should run Windows 10 Professional or macOS version 10.13 or later with latest security patches applied.
- No private laptops are allowed to be used for handling EU RESTRICTED information.

Sending EU RESTRICTED contributions among partners - Encryption software

During the elaboration of deliverable, it will be necessary to exchange text, diagrams, figures, etc. The electronic EU RESTRICTED documents must be stored encrypted. The document or pieces of it cannot be uploaded to a collaborative environment as the Confluence repository of CARAMEL.

If you need to send EU RESTRICTED classified data over the internet, you must encrypt all information.

In CARAMEL, we have chosen the off-line encryption tool Zed! (https://www.zedencrypt.com/) for use within the project. It is relatively inexpensive, has EAL3+ evaluation, and has the option of downloading a free read-only version, implying that only partners that need to send EU RESTRICTED have to purchase the Zed! Pro version that costs 39,90€. These costs are eligible.

Sending EU RESTRICTED deliverable to EC

The project coordinator and the lead editor will decide in agreement to the project officer (European Commission) the suitable way to deliver D2.2 report. EU RESTRICTED documents can be sent either by post or by appropriately encrypted electronic means. If an electronic means is chosen, the project may provide the PO with the necessary authentication key to read the document.

If regular post is chosen, the hardcopies of classified deliverable must be placed in an addressed envelope sealed with security tape, with classification markings top and bottom as illustrated in next figure. In turn, this envelope should be placed in an opaque and unmarked envelope, and sent to the addressee using registered mail of the PO.

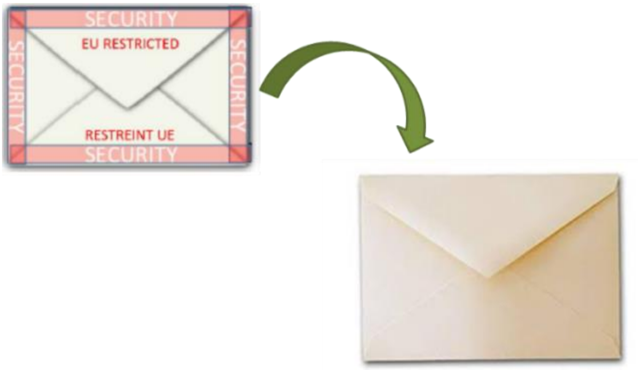


Figure 2: CARAMEL D2.2 – Sending Report Process

Responsibility

The adherence to these rules is the responsibility of

- The work package 2 leader
- Any partner responsible for contributing to EU-Classified deliverable D2.2
- The PSO and SAB members if reading the classified deliverable

3.6 *Making data interoperable*

The interoperability of the project's data is directly related to the impact of the project and the possibilities of re-use, migration on different platforms and extrapolation of results in different applications. Thus, interoperability is in fact a design principle for the project partners. As far as this is possible, open file formats, open vocabularies and other relevant standard will be used in order to maximize interoperability.

3.6.1 License of the data: how and when

Data generated and used within the project will be made publicly available, in the cases that this is possible, following the FAIR data directive [6]. When this is not possible, licensing options will be examined on a case by case basis, taking into account all applicable factors.

3.6.2 Data quality assurance process

All the CARAMEL's deliverables and outputs will be peer-reviewed by the project partners. Through this approach we expect to ensure high data quality within the project, promoting project data reuse and sharing.

3.7 *Data security*

In order to prevent unauthorized access, modification, replication or destruction of the project's data, the following measures are to be ensured:

- Identification security: Data is stored in online repositories which are password protected and/or grant access only upon correct identification. Different layers of security are implemented in order to protect data of higher sensitivity (users' personal data, etc.)
- Location security: Access to the premises of the partners, where the project is being developed, is restricted.
- Workstation security: People working on the project are strongly encouraged to remain protected against a possible data breach by password protecting all computers and through the use of an up to date antivirus software. Additionally, the sharing of confidential information via email is highly discouraged.

- All data use in the project will be regularly backed up and, in most cases, will be residing on cloud storage facilities, preventing this way the possibility of loss of data due to hardware failure.

More details about this can be found on ANNEX III.

3.7.1 Digital Data Storage

Securing stored data involves preventing unauthorized people from accessing it as well as preventing accidental or intentional destruction, infection or corruption of information. Data encryption is a popular mechanism that is utilised by CARAMEL but nevertheless, it is just one of many techniques and technologies that can be used to implement a tiered data-security strategy. Steps to secure data involve understanding applicable threats, aligning appropriate layers of defence and continual monitoring of activity logs taking action as needed. The proper method of storage along with levels of access for privileged users are important considerations for comprehensive protection. Improperly stored information along with overly permissive accounts are a centralized theme in many high-profile breaches.

Data-in-storage must be protected from unauthorized access, modification and loss and as such the measures are implemented by all CARAMEL partners:

- Data availability must be guaranteed.
- Confidential data must be stored using access protection.
- Strictly confidential information must only be stored in an encrypted mode
- Confidential data must not be stored in online services that are not approved by the CARAMEL Consortium.
- Any exception from this measure must explicitly be approved
- Modifications to data with high integrity requirements must be documented and approved by the partners.

More details about this can be found on ANNEX III.

4 ANNEX I - INFORMATION SHEET

INFORMATION SHEET

Project: CARAMEL

Aim of the study

Description of the study and incentive

Privacy and anonymity

- All data gathered in our projects will be processed anonymously and only be used within this project. All participants personal info will be coded (for example using pseudonyms) in the analysis and reporting of the data. This means that your name will not be linked to the gathered information.

5 ANNEX II - INFORMED CONSENT FORM

INFORMED CONSENT FORM

Project: CARAMEL

Aim of the study

Description of the procedure

Permission

I, ____, agree with the content of this document and agree to participate in the CARAMEL project. I agree/ not agree with the usage of pictures/movies with the CARAMEL consortium.

Date:

Signature:

6 ANNEX III – Dedicated Data Storage and Processing Facility provided by T-Sys

T-Systems has setup a lab environment for exclusive use by project partners in the CARAMEL project in the Telekom Security server room in Berlin.

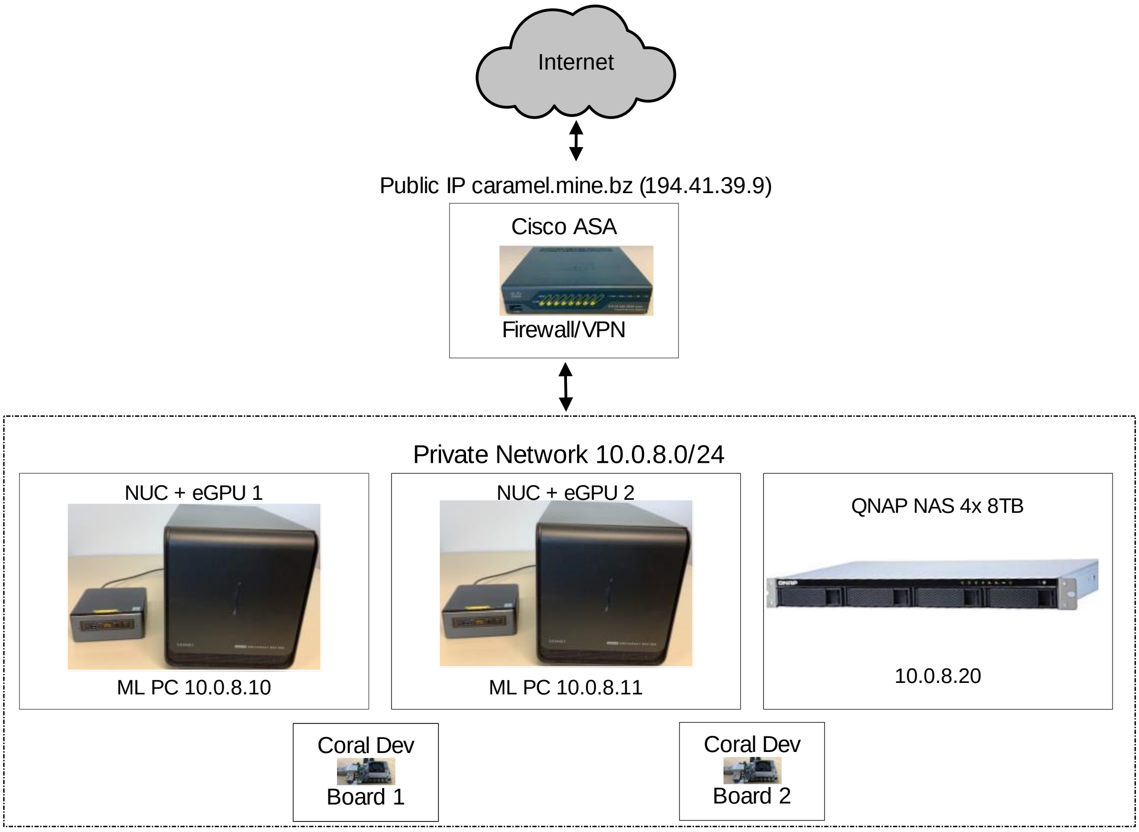


Figure 3: CARAMEL ML Lab Setup in Berlin (Telekom Security Lab)

Figure 3 shows the different components of the system:

- Cisco ASA firewall and VPN endpoint appliance
- QNAP NAS devices
- 2 Linux-based PCs with GPU for machine learning purposes
- 2 Coral Dev Boards (target hardware for the anti-hacking device developed by T-Systems)

6.1 *Access control*

Full access to the NAS device and the PCs is only possible by establishing a VPN connection using the Cisco AnyConnect client software (or compatible software). This VPN access is protected by UserID and password. The UserIDs, passwords, real names, and email addresses are known to one T-Systems employee tasked with maintaining the lab infrastructure. Project members have been advised not to use these passwords in other contexts.

The same UserID and password combinations are setup for SSH access to the PCs as well as for individual SMB shares on the NAS device. Users don't have root access on the PCs. Individual machine learning or other specific environments can be started as Docker containers.

The NAS device is also accessible without going through the VPN connection via HTTPS. The TLS server certificate is obtained from the letsencrypt project and updated automatically. The same

UserID/password combinations as above are used for access to the services offered by the NAS device via HTTPS. The following services are available:

- Individual download and upload folders
- WP-specific specific folders to separate data access concerns (WP membership managed by NAS-based groups)
- CARAMEL project-wide download and upload folders
- Docker Registry for project-specific Docker images (based on Harbor software)
- Individual file sharing services for project members (based on Plik)

Only T-Systems has management access to the NAS.

Access to the dev boards will be granted on a case-by basis. Credentials will be defined by T-Systems and communicated to project partners that need to work with the devices.

6.2 *Security of Data-at-rest*

The NAS has only encrypted volumes. The encryption key must be entered by the T-Systems employee tasked with maintaining the NAS upon every boot of the NAS device. This way, there will be no data loss even when the whole NAS system is stolen.

The disks in the PCs are not encrypted. Sensitive data sets should therefore be stored on the NAS device.

6.3 *Logging*

NAS device and PCs provide the usual Linux logging facilities. The logging system on the NAS will be configured such that user logins and upload/download actions are logged appropriately and to the extent supported by the NAS device software (QNAP QTS operating system).

Log files might contain personally identifiable information that are sensitive and should not be stored longer than necessary to detect misuse of the platform or the data sets stored on the platform. Therefore, T-Systems will ensure that log data is deleted regularly during the course of the project and most certainly at the end of the project when the hardware in the lab is repurposed for other activities.

6.4 *System Security*

The NAS device and the Cisco ASA firewall will be updated as soon as patches become available by the respective vendors QNAP and Cisco.

The PCs offer no services to the Internet without going through the VPN. Therefore, system updates for the PCs will be handled more conservatively in order not to disrupt ongoing activities by project partners with incompatible updates.

All hardware is located in a secure environment (the Telekom Security lab server room in Berlin, Germany, Holzhauser Str. 4-8) that is protected by smartcard access (with additional PIN protection). Only a limited number of employees has access to these premises.

References

[1] Research Ethics. European Commission [Online] Available: https://ec.europa.eu/research/swafs/index.cfm?pg=policy&lib=ethics

[2] Charter Of Fundamental Rights Of The European Union. Official Journal of the European Communities. 2000. Online] Available: https://www.europarl.europa.eu/charter/pdf/text_en.pdf

[3] The European Code of Conduct for Research Integrity: Revised Edition. All European Academies, Berlin 2017. [Online] Available: https://allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf

[4] Horizon 2020 Rules for Participation. European Commission. [Online] Available: https://ec.europa.eu/research/participants/data/ref/h2020/legal_basis/rules_participation/h2020 -rules-participation_en.pdf

[5] The Horizon 2020 Regulation of Establishment. European Commission. [Online] Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32013R1291

[6] Guidelines on FAIR Data Management in Horizon 2020. European Commission. [Online] Available: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

[7] OpenAIRE. European Open Science Infrastructure, for open scholarly and scientific communication. [Online] Available: https://www.openaire.eu/

[8] Zenodo. Open-access repository. CERN. OpenAIRE. [Online] Available: https://zenodo.org/

[9] Security rules for protecting EU classified information. European Commission [Online] Available: https://op.europa.eu/en/publication-detail/-/publication/41a6eeeb-cc70-11e4-ab4d-01aa75ed71a1/language-en