

Security ECONomics service platform for smart security investments and
cyber insurance pricing in the beyond 2020 netwOrking era



WP1 – Project Management and Coordination
Deliverable D1.5 “Data Management Plan”








Editor(s):	Christos Xenakis, ?
Author(s):	Farnaz Mohammadi (UPRC)
Dissemination Level:	Public
Nature:	Report
Version:	1



Project Profile

Contract Number	823997
Acronym	SECONDO
Title	Security ECONomics service platform for smart security investments and cyber insurance pricing in the beyond 2020 netwOrking era
Start Date	Jan 1 st , 2019
Duration	48 Months

Partners

Logo	Partner		Country
	Full Name	Short Name	
 University of Piraeus	University of Piraeus Research Centre	UPRC	Greece
	UNIVERSITY OF SURREY	SURREY	United Kingdom
	Cyprus University of Technology	CUT	Cyprus
	UBITECH LIMITED	UBI	Cyprus
	LSTech Espana	LST	Spain
	Cromar Insurance Brokers LTD	CRO	Greece
	Fogus Innovations & Services P.C.	FOG	Greece

Document History

VERSIONS

Version	Date	Author	Remarks
0.1	24/05/2019	Farnaz Mohammadi	Initial draft
0.2	20/06/2019		
1.0	30/06/2019		Final Draft

Executive Summary

This deliverable presents **the first version of Data Management Plan (DMP)** for the SECONDO project. The deliverable outlines how the research Data collected/generated will be handled during and after the SECONDO project, describes which standards and methodology for Data collection/generation will be followed, and whether and how Data will be shared. DMP is a mandatory report for all projects participating under the Open Research Data Pilot (ORDP) in Horizon 2020.

SECONDO is a part of Marie Skłodowska-Curie Actions [1](part of Horizon 2020 project) participating in the ORDP, which aims to reuse, improve, and maximise access to research Data collected/generated by Horizon 2020 projects. ORDP also takes into account the need to balance openness and protection of scientific information, commercialization and Intellectual Property Rights (IPR), privacy concerns, security as well as Data management and preservation questions.

The document is based on the template and guidelines provided by the European Commission [2] [3], will be updated regularly with the purpose of supporting the Data management life cycle for all Data that will be collected, processed or generated by the project (**living document**).

It should be noted that SECONDO will not collect or process Personal Data to conduct its research.

Table of Contents

Executive Summary.....	3
Table of Contents	4
List of Figures.....	5
List of Tables	5
Table of Abbreviations	6
1. Introduction	7
2. SECONDO FAIR Data Principles	8
2.1 DATA Summary	8
2.1.1 Purpose of the Data Collection/Generation and its relation to the objectives of the project.....	9
2.2 Allocation of Resources	10
2.3 Data Sharing.....	10
2.3.1 Methods for Data sharing	11
2.4 Data Security	11
2.4.1 Data Protection	13
3. SECONDO Data Sourcing and Data Sharing Architecture	15
3.1 Overview of ROs’ scenarios	16
RO1. Design and develop an extended risk analysis metamodel.	16
RO2. Design and develop a scenario-based risk management module that facilitates in both cost-effective risk management and optimised security investments.	16
RO3. Design and develop a cyber insurance module that estimates cyber insurance exposure and derives coverage and premiums.	17
RO4. Use smart contracts and a blockchain to empower cyber insurance claim.	17
3.2 SECONDO Data Sources	18
4. Data Archiving and Preservation (including storage and backup).....	18
5. Ethical Aspects and Privacy and Security Requirements.....	19
5.1 General Data Protection Regulation (GDPR).....	21
5.2 Security and Authentication Legislation	22
6. The SECONDO FAIR Dataset Template Questionnaire	22
7. References.....	31
8. Appendix	33

List of Figures

No table of figures entries found.

List of Tables

Table 6-1: SECONDO FAIR Dataset Template Questionnaire.....	22
Table 8-1: Protection of Personal Data.....	33
Table 8-2: Human Participation	34
Table 8-3: Other Ethics Issues	34

Table of Abbreviations

Abbreviation	Explanation
DMP	Data Management Plan
ORDP	Open Research Data Pilot
IPR	Intellectual Property Rights
ERC	European Research Council
FAIR	Findable, Accessible, Interoperable, Reusable
RO	Research Objective
GRAM	Quantitative Risk Analysis Metamodel
TO	Technical Objective
BO	Business Objective
RAOHM	Risk Analysis Ontology and Harmonisation Module
SEAM	Social Engineering Assessment Module
BDCPM	Big Data Collection and Processing Module
ESaaS	Economics-of-Security-as-a-Service
CSIM	Cyber Security Investment Module
GTM	Game Theoretic Module
CRMM	Continuous Risk Monitoring Module
CICPM	Cyber Insurance Coverage and Premiums Module
HDPa	Hellenic Data Protection Authority
SMW	Social Media Website
DOA	Description Of Action
EUI	European University Institute

1. Introduction

SECONDO aims at achieving the following features simultaneously: efficiency, security, user privacy, and flexibility of contract expressiveness. This deliverable addresses the main elements of the Data Management policy that will be used by the project participants regarding all Datasets. It also establishes some procedural mechanisms for participants with the responsibilities of Data Controllers and Processors. Through the SECONDO project, **Data Controllers** and **Processors** are defined as below [4] :

- **Data Controller** means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law;
- **Data Processor** means a natural or legal person, public authority, agency or other body which processes Personal Data on behalf of the controller;

The DMP establishes a set of guidelines meeting each of the fundamental topics to be considered. These guidelines cover aspects such as applicable policies, roles, standards, infrastructures, sharing strategies, Data processing, storage, retention and structure, legal compliance and compliance with market standards and best ethical and privacy practices, identification, accessibility, intelligibility, legitimate use for other purposes. These guidelines will be adopted at the early stages of the project.

For each Dataset in H2020 the following aspects should be considered:

- Making Data Findable,
- Making Data openly Accessible,
- Making Data Interoperable,
- Increase Data Re-use,
- Allocation of recourses and Data security

The Data collected/generated during the project will be owned by the partners which have contributed to producing that Data (Data controller). The extent up to which this Data will be made available and which restrictions will be imposed on its re-use will be decided on a case-by-case basis by the Data controller. Moreover, Data controller determines the purposes and means of Personal Data processing and will decide the purpose for which Personal Data is required and what Personal Data is necessary to fulfil that purpose.

The partners will comply with the Findable, Accessible, Interoperable, Reusable (FAIR) guidelines of the H2020 programme, which state that Data will be made as available as possible, so long that does not negatively affect the commercial advantage of the partners. The Horizon 2020 FAIR DMP template [5] has been designed to be applicable to any Horizon 2020 project that produces, collects or processes research Data.

The Data will be shared among partners using internal repositories or through direct communication,

and with the public through the project’s website or public repositories. The Data will be preserved up to **three (3) years** after the end of the project at the partners’ repositories and cloud infrastructures, according to each partner’s internal policy.

SECONDO DMP should be updated as a minimum in time with the periodic evaluation/assessment of the project. Furthermore, the consortium can define a timetable for review in the DMP itself. Regarding [5], the SECONDO DMP needs to be updated over the course of the project whenever significant changes arise, such as (but not limited to):

- Using new Dataset
- Changes in consortium policies (e.g. innovation potential, decision to file for a patent)
- Changes in consortium composition and external factors (e.g. new consortium members joining or old members leaving).

Regarding [Participant Portal H2020 Online Manual \[6\]](#), as part of making research Data findable, accessible, interoperable and re-usable (FAIR), SECONDO FAIR DMP should include information on:

- The handling of research Data during and after the end of the project
- What Data will be collected, processed and/or generated?
- Which methodology and standards will be applied?
- Whether Data will be shared/made open access?
- How Data will be curated and preserved (including after the end of the project).

The SECONDO FAIR Dataset Template Questionnaire (Table 6-1) includes a set of questions that all Data controllers are required to fill in for each Dataset [3], [7], [8]. The questionnaire template has been reviewed by the Project Coordinator (UPRC) and the Ethics Board for completeness and compliance with the FAIR DMP directives.

Zenodo [9] will be used as the project Data and publication repository and will be linked to the SECONDO project-site at OpenAIRE. Zenodo is a simple and innovative service that enables researchers, scientists, EU projects and institutions to share and showcase multidisciplinary research results (Data and publications) that are not part of existing institutional or subject-based repositories.

2. SECONDO FAIR Data Principles

2.1 DATA Summary

As part of making research data Findable, Accessible, Interoperable and Re-usable (FAIR), a DMP should [6]:

To be Findable:

- Data/Metadata are assigned a globally unique and eternally persistent identifier.
- Data are described with rich Metadata.
- Data/Metadata are registered or indexed in a searchable resource.
- Metadata specify the Data identifier.

To be Accessible:

- Data are retrievable by their identifier using a standardized communications protocol.
- the protocol is open, free, and universally implementable.
- the protocol allows for an authentication and authorization procedure, where necessary.
- Metadata are accessible, even when the Data are no longer available.

*Particularly, SECONDO Database will be accessible through the SECONDO project for **three (3)** years following the end of the project. During this period, unless otherwise decided by the consortium members, the Database functionality will remain the same as during the project duration.*

To be Interoperable:

- Data/Metadata use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- Data/Metadata use vocabularies that follow FAIR principles.
- Data/Metadata include qualified references to other Data/Metadata.

There is not a standard for allowing Data exchange between researchers, institutions, organizations, countries, etc. (e.g. adhering to standards for Data annotation, Data exchange, compliant with available software applications, and allowing re-combinations with different Datasets from different origins). Thus, it always needs of a human interpretation of the Data structure to manually create a Data map. However, the utilization of standards for Data capturing and the documented annotation will ease the Data exchange.

To be Re-usable:

- Data/Metadata have a plurality of accurate and relevant attributes.
- Data/Metadata are released with a clear and accessible Data usage license.
- Data/Metadata are associated with their provenance.
- Data/Metadata meet domain-relevant community standards.

SECONDO Data will be licensed under Creative Commons license, to the extent that it may be subject to such licensing (likely the “CC BY”). Applicable Data will become available at the end of the project. The Data can be re-used by other scientists and interested parties. Parts of the Data may become available prior to this as a result of journal publications. There will be no embargo period.

2.1.1 Purpose of the Data Collection/Generation and its relation to the objectives of the project

SECONDO will propose a unique, scalable, highly interoperable **Economics-of-Security-as-a-Service (ESaaS) platform** that encompasses a comprehensive cost-driven methodology for estimating cyber risks and determining the residual risks. The SECONDO platform will establish a new paradigm in risk management for enterprises of various sizes, with respect to the GDPR framework, while it will enable formal and verifiable methodologies for insurers that require estimating premiums. **SECONDO will not collect or process Personal Data to conduct its research. The collection, processing and use of**

Personal Data is only admissible if expressly permitted by any legal provision or if the Data subject has expressly consented in advance.

2.2 Allocation of Resources

The costs for Data preparation to be FAIR are unknown at this stage but will be estimated in the future. Expenses may consist of additional publication and documentation costs of the repositories where applicable. Data preparation and management costs during the project will be covered by the project.

UPRC, as the Project Coordinator for **SECONDO**, will be responsible for DMP updates, and Data archiving and publication within repositories. No additional funding is provided for Data management activities for those deciding to participate in the pilot. Costs relating to open access to research Data will be eligible as part of the grant, independent from the participation in the pilot, provided the general eligibility conditions specified in the Grant Agreement are followed.

2.3 Data Sharing

The Data controller will determine the details of how Data will be shared, including access procedures, embargo periods (if any), outlines of technical mechanisms for dissemination and necessary software and other tools for enabling re-use, and definition of whether access will be widely open or restricted to specific groups. Similarly, the Data controller will identify the repository where Data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.).

During the project, any potential user that wants the get access would be guided to:

- Submit a "Request" to the Dataset controller from the **SECONDO** consortium.
This request will contain:
 - Full name
 - Organization and department
 - Email address
 - Description of intended use
- After reviewing the request, if the Data controller approves it, the user will receive an email with a special link to verify the email address.
- Then the user is asked to agree to and sign the following terms of access:
[RESEARCHER_FULLNAME] (the "Researcher") has requested permission to use the Dataset. In exchange for such permission, Researcher hereby agrees to the following terms and conditions:
 - Researcher shall use the Database only for non-commercial research and educational purposes.
 - Data Controller makes no representations or warranties regarding the Dataset, including but not limited to warranties of non-infringement or fitness for a particular purpose.
 - Researcher accepts full responsibility for his or her use of the Dataset and shall defend and indemnify Data Controller, including their employees, trustees, officers and agents, against any and all claims arising from Researcher's use of the Dataset,

including but not limited to Researcher's use of any copies of copyrighted Dataset that he or she may create from the Dataset.

- Researcher may provide research associates and colleagues with access to the Dataset provided that they first agree to be bound by these terms and conditions.
- Data Controller reserves the right to terminate Researcher's access to the Database at any time and without justification.
- If Researcher is employed by a for-profit, commercial entity, Researcher's employer shall also be bound by these terms and conditions, and Researcher hereby represents that he or she is fully authorized to enter into this agreement on behalf of such employer.
- The law and jurisdiction of the Data Controller's country shall apply to all disputes under this agreement.

2.3.1 Methods for Data sharing

The methods used to share Data will be dependent on a number of factors such as the type, size, complexity and sensitivity of Data. Data can be shared by any of the following methods [10]:

- **Under the auspices of the Principal Investigator**
Investigators sharing under their own auspices may securely send Data to a requestor or upload the Data to their institutional website. Investigators should consider using a Data-sharing agreement to impose appropriate limitations on the secondary use of the Data.
- **Through a third party**
Investigators can share their Data by transferring it to a Data archive facility to distribute more widely to the scientific community, to maintain documentation and meet reporting requirements. Data archives are particularly attractive for investigators concerned about managing a large volume of requests for Data, vetting frivolous or inappropriate requests, or providing technical assistance for users seeking to help with analyses.
- **Using a Data enclave**
Datasets that cannot be distributed to the general public due to confidentiality concerns, or third-party licensing or use agreements that prohibit redistribution, can be accessed through a Data enclave. A Data enclave provides a controlled secure environment in which eligible researchers can perform analyses using restricted Data resources.
- **Through a combination of methods**
Investigators may wish to share their Data by a combination of the above methods or in different versions, in order to control the level of access permitted.

Note: During the SECONDO project, Data controllers could use **a combination of methods** for Data sharing.

2.4 Data Security

Regarding [Guide on Good Data protection practice](#) [11], to process Data in a secure manner, each Data controller must:

- Take technical and organisational measures to prevent any unauthorised access
- Establish clear access rules

- Organise the processing in a way that gives you the best possible control, for example by allowing for tracking of access (logbook)
- If someone processes the Data on your behalf, make sure that this processor ensures for appropriate security safeguards.

In practical terms, these measures could result in:

- User authentication: The way to verify the identity of a user
- Access control: Mechanism to allow or deny access to certain Data
- Storage security: Storing Data in a way that prevents unauthorised access, for example by:
 - Operating system controls (authentication & access control)
 - Use of passwords to access electronic files (e.g. use the text editor function to save a document password-protected)
 - Local encrypted storage (enable the full disk encryption, enable the file system, enable the text editor encryption)
 - Database encryption: turning Data into a form that makes them unintelligible (for anyone not having access to the key)
- Communication security: Safe electronic communication for transferring the Data can take the following forms:
 - Encrypted communication (SSL/TLS); (e.g., use web services whose URL starts with ‘https: //’ and not only http ://)
 - Firewall systems and access control lists (e.g. make sure the firewall service is enabled on your PC)
 - Anti-virus & anti-malware systems
 - Protect Data and Data carriers when they are physically transferred (paper notes, laptop etc.).
- Other IT technical controls such as installing security updates, anti-virus protection, local back-ups, blocking of certain software installation, etc.

Regarding the guidelines on implementation of open access to scientific publications and research Data, participants of the ORDP need to take the following three steps [12]:

- Deposit research Data needed to validate the results presented in scientific publications, including associated Metadata, in the repository as soon as possible. Also, other Data (for instance Data not directly attributable to a publication, or raw Data), including associated Metadata, should be deposited – that is, according to the individual judgement by each project, specified in the Data management plan.
- Take measures to enable third parties to access, mine, exploit, reproduce and disseminate (free of charge for any user) this research Data, for instance by attaching [a Creative Commons Attribution Licence](#) (CC BY) to the Data deposited, or by waiving all interests associated to copyright and Database protection.
- Provide information via the chosen repository about the tools available in order for the beneficiaries to validate the results, e.g., specialised software or software code, algorithms and analysis protocols. Where possible, these tools or instruments should be provided.

All SECONDO software/toolkit modules will encapsulate state-of-the art security, authentication and authorization mechanisms. The robustness of such modules is ensured by years of developments in the field (the basic building-blocks stem from previously funded EU projects or from already functioning commercial solutions) and will be tested through dedicated penetration / hacking tests and challenges. In addition, Data protection methods will be made available through a set of secure APIs and Smart Contracts. Moreover, privacy-preserving smart contracts will be leveraged to hide sensitive client information and meanwhile, secure encryption technique will be considered in Data storage.

Privacy-preserving techniques will be used in Data storage and smart contract to protect clients’ privacy. Privacy-preserving smart contracts will be leveraged to hide sensitive client information and meanwhile, secure encryption technique will be considered in Data storage.

The conceptual security and privacy taxonomy will be applied. It contains three main Big Data security and privacy principles:

- Data confidentiality topic: safeguarding the confidentiality of Personal Data.
- Data provenance topic: safeguarding the integrity and validation of Personal Data.
- Public policy, social, and cross-organizational topics: safeguarding the specific Big Data and privacy and Data protection requirements.

In SECONDO, a Byzantine-fault-tolerance-like algorithm will be used to randomly select a group of clients as validators. To achieve security, access control will be used to guarantee that only registered clients can read information from the ledger.

2.4.1 Data Protection

As mentioned in SECONDO DOA, no real Data will be used in the context of the project. However, with SECONDO being a GDPR-compliant platform by design, we describe the procedures and technical measures that would be applied if real Data are being processed.

A key issue in considering observational research using social media is whether the proposed project meets the criteria as human subjects’ research, and if so, what type of review is needed. A human subject is defined by federal regulations as a living individual about whom an investigator obtains Data through interaction with the individual or identifiable private information [13].

An important area of concern with **Social Media Website (SMW)** research is the protection of confidentiality. Similar to other types of research involving survey or interview Data, protection of participant identities is critical. Website research may initially be perceived as lower risk, because participant information can be collected in absence of some protected information such as address or phone number. Online Data can present increased risks; studies that publish direct text quotes from an SMW may directly identify participants. Entering a direct quote from an SMW into a Google search engine can lead to a specific Web link, such as a link to that person’s LinkedIn profile, and thus identify the participant.

Personal Data refers to any information relating to an identified or identifiable natural person, meaning by identifiable person the one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity. Data is considered personal when someone is able to connect information to a specific person, even when the person or entity that is holding the Personal Data cannot make the connection directly (e.g. name, address, e-mail), but has or may have access to information allowing such identification (e.g. through telephone numbers, credit card numbers, license plate numbers, etc.).

The fundamental right to the protection of Personal Data is explicitly recognised in Article 8 of the Charter of Fundamental Rights of the European Union, and in Article 16 of the Treaty on the functioning of the European Union, according to which everybody has the right to the protection of Personal Data concerning them. 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.

If Data controllers intend to process sensitive data in the project, or if there is a possibility that sensitive data (See section: 5: Ethical Aspects and Privacy and Security Requirements) may be processed (unintended processing of sensitive data), more solid justification to the Ethics Committee have to be provided by Data Controllers

SECONDO Data processing must be lawful, fair and transparent. It should involve only Data that are necessary and proportionate to achieve the specific task or purpose for which they were collected. Therefore, **SECONDO will only collect the Data that is needed for the research objectives**, since collecting unnecessary/unrelated Data for the research project may be deemed unethical and unlawful. The Data are to be processed only for scientific purposes comprising processing operations that are performed for purposes of study and systematic research to develop scientific knowledge for the specific sector addressed by SECONDO.

SECONDO will not collect or process Personal Data to conduct its research. Any real users that will take part in the assessment of the implemented software do not have to provide Personal Data and in case some non-sensitive Data are needed, the users will be informed and sign the appropriate consent and agreements.

To secure the confidentiality, accuracy, and security of Data management, the following measures will be taken:

- All Personal Data obtained in SECONDO studies will be transmitted to partners within the consortium only after anonymization. Keys to identification numbers will be held confidentially within the respective research units. In situations where re-identification of study participants becomes necessary, for example the collection of additional Data, this will only be possible through the research unit and in cases where informed consent for such cases has been given.
- Personal Data are entered to secure websites. Data are processed only for the purposes outlined in the patient information and informed consent forms of the respective case studies.

Use for other purposes will require explicit patient approval. Also, Data are not transferred to any places outside the consortium without patient consent.

- None of the Personal Data will be used for commercial purposes, but the knowledge derived from the research using the Personal Data may be brought forward to such use as appropriate, and this process will be regulated by the Grant Agreement and the Consortium Agreement, in accordance with any generally valid legislation and regulations.
- No vulnerable or high-risk groups are used (e.g. children, adults unable to consent, people in dependency relationship, vulnerable persons) will be addressed during the development and progress of the SECONDO project;
- Persons are approached in their professional capacity;
- The purpose of collecting contact Data of potential stakeholders is to ask them about their willingness to be involved in SECONDO network and for obtaining professional opinions and consultation only;
- Information about the objectives of the project, structuring of the Stakeholder Network and details about Data processing will be provided in advance (as a governance document) to all external stakeholders;
- Minimum and limited amount of Personal Data will be collected;
- Personal contact Data will be kept internally within the SECONDO partners and will not be accessible to external organizations or individuals.
- Personal Data shall always be collected, stored, and exchanged in a secure manner, through secure channels during the project.

Regarding Data confidentiality, SECONDO partners must keep any Data, documents or other material confidential during the implementation for the project and for three years after end of the project.

Note: Appendix (Table 8-1, Table 8-2, Table 8-3) has to be filled by the SECONDO Data controllers.

3. SECONDO Data Sourcing and Data Sharing

In general, Data may be grouped into four main types based on methods for collection:

- **Observational Data:** captured in real time, typically cannot be reproduced exactly.
- **Experimental Data:** from labs and equipment, can often be reproduced.
- **Simulation Data:** from models, can typically be reproduced if the input Data is known.
- **Derived or Compiled Data:** after Data mining or statistical analysis has been done, can be reproduced if analysis is documented.

The categories of Data processed in SECONDO are:

- **Experimental**
Dataset captured from a real infrastructure, such as external sources e.g. social media and other internet-based sources, including Darknet to establish research activities with a

static dataset.

- Simulation
Dataset captured in real time from a testbed or lab infrastructure to monitor it and test optimization strategies from internal organisation sources, e.g. network infrastructure.
- Derived or Compiled Data
*The intelligent **Big Data Collection and Processing Module (BDCPM)** uses specialised crawlers to acquire risk-related Data.*

3.1 Overview of Research Objectives (ROs') scenarios

The interactions mapped in the Research Objectives (ROs) scenarios have determined the Data sources, as well as the connections that will take place in SECONDO. A short description of each ROs' scenarios can be found below:

RO1. Design and develop an extended risk analysis metamodel.

One of the key contributions of the SECONDO programme in the area will be the design, analysis and implementation of a Quantitative Risk Analysis Metamodel (GRAM) that will utilise advanced security metrics to quantitatively estimate the exposed cyber risks. To implement the desired functionalities the following SECONDO modules will be implemented:

- **Risk Analysis Ontology and Harmonisation Module (RAOHM)**
RAOHM receives the outcomes of the existing risk analysis tools and harmonises them using a common vocabulary with straightforward definition in order to be used by GRAM (Leader: UPRC)
- **Social Engineering Assessment Module (SEAM)**
SEAM interacts with users to devise their behaviour using penetration testing approaches and it provides specific numeric results on risky actions, (i.e. percentage of users that open suspect files or execute Trojans, etc.) (Leader: UPRC).
- **Intelligent Big Data Collection and Processing Module (BDCPM)**
BDCPM uses specialised crawlers to acquire risk-related Data either from internal organisation sources, e.g. network infrastructure or external sources such as social media and other internet-based sources, including Darknet. (Leader: LST)

RO2. Design and develop a scenario-based risk management module that facilitates in both cost-effective risk management and optimised security investments.

Cyber Security Investment Module (CSIM) will be designed and implemented. CSIM will provide decision support for organisations that seek an optimal equilibrium point (i.e. balance) between spending on cyber security investment and cyber insurance fees.

CSIM will use the following results/procedures/modules outcome as an input:

- Costs for attacking and defending will be investigated and they will be given as an input to CSIM. (Leader: CUT)
- the outcome of the provided extended and QRAM
- the results of BDCPM that provides analytics on Internet sources regarding state-of-the-art security solutions as well as their cost. (Leader: LST)
- The outcome of the Game Theoretic Module (GTM) that models all possible attacking scenarios and defensive strategies, (i.e. available security controls), by employing attack graphs (Leader: FOGUS)
- The outcome of the Econometrics Module (ECM) that provides estimates of all kinds of costs of potential attacks and it takes into account costs, (i.e. purchase, installation, execution, etc.), of each possible security control using a set of existing econometric models; (Leader: CUT)
- The outcome of the Continuous Risk Monitoring Module (CRMM) that assesses on a continuous basis the performance of the implemented risk-reducing cyber security controls allowing the adaptation of the cyber insurance contract to the changing IT environment and the evolving cyber threat landscape (Leader: UBI)

RO3. Design and develop a cyber insurance module that estimates cyber insurance exposure and derives coverage and premiums.

the Cyber Insurance Coverage and Premiums Module (CICPM) will compute premium curves and coverages as a function of the organisation’s security level (can be used by clients). CICPM will communicate with CRMM for monitoring the conditions that violate cyber insurance contract agreements toward resolving conflicts.

CICPM will use the following results/outcomes/policies as an input to propose the insurance calculation tool:

- The outcome of the proposed QRAM.
- The defending policies selected to be applied in order to provide optimal protection strategies as well as the results of the related econometric parameters that justify the cost effectiveness of the considered security investments (Leader: UPRC).
- The results of analytics on cyber insurance environment and market (Leader: CRO).
- The underwriter’s strategy (Leader: SURREY).

RO4. Use smart contracts and a blockchain to empower cyber insurance claim.

SECONDO will deploy a blockchain, which is a distributed decentralised Database that maintains continuously growing blocks of Data records, in which all blocks are tightly chained together against information tampering. SECONDO will use a private ledger, which provides secure access control on Data records, to hold an inventory of assets and information regarding security and privacy risk measurable indicators of an organisation (cyber insurance client). The ledger will be updated based on information received from CRMM. By using smart contracts, the traditional physical-based paper

process and endorsement will be turned to digital formats that brings convenience on Data management (Leader: SURREY).

3.2 SECONDO Data Sources

In the context of the project, SECONDO will not collect /process Personal Data to conduct its research. The major Data sources, as these have been identified in SECONDO Description Of Action (DOA), are described below.

UPRC and LIST are nodes of QRAM that RAOHM (Leader: UPRC) as a main part of SECONDO Risk analysis module, receives the outcomes of the existing risk analysis tools and harmonises them using a common vocabulary with straightforward definition. And internal organisation sources, e.g. network infrastructure or external sources such as social media and other internet-based sources, including Darknet will be used by BDCPM to acquire risk-related Data. In the context of the QRAM, SEAM (Leader: UPRC) interacts with users to devise their behaviour using penetration testing approaches and it provides specific numeric results on risky actions, (i.e. percentage of users that open suspect files or execute Trojans, etc.)

For the CSIM phase, costs for attacking and defending will be investigated and they will be given as an input to CSIM (Leader: CUT), and the results of BDCPM that provides analytics on Internet sources regarding state-of-the-art security solutions as well as their cost. (Leader: LST). The outcome of the Game Theoretic Module (GTM) that models all possible attacking scenarios and defensive strategies, (i.e. available security controls), by employing attack graphs (Leader: FOGUS).

ECM provides estimates of all kinds of costs of potential attacks and it takes into account costs, (i.e. purchase, installation, execution, etc.), of each possible security control using a set of existing econometric models; (Leader: CUT). CRMM assesses on a continuous basis the performance of the implemented risk-reducing cyber security controls allowing the adaptation of the cyber insurance contract to the changing IT environment and the evolving cyber threat landscape (Leader: UBI).

CICPM will compute premium curves and coverages as a function of the organisation’s security level (can be used by clients). CICPM will communicate with CRMM for monitoring the conditions that violate cyber insurance contract agreements toward resolving conflicts. CICPM will use the QRAM’s outcome, Cyber insurance ontology (Lead: UPRC), results of analytics on cyber insurance environment and market (Leader: CRO), the underwriter’s strategy (Leader: SURREY).

As mentioned before, SECONDO will use a private ledger to hold an inventory of assets and information regarding security and privacy risk measurable indicators of an organisation (cyber insurance client). By using smart contracts, the traditional physical-based paper process and endorsement will be turned to digital formats that brings convenience on Data management (Leader: SURREY).

4. Data Archiving and Preservation (including storage and backup)

The collected Data will be stored in secure servers, only accessible to the consortium members. If any identifiable Data are required for the research purposes, access to and distribution of this Data will be

granted only after explicit permission and after the agreement of the user participants. Authentication will be required to access stored Data on the research site. Authorized consortium members will have access to the Data after authentication with a centralized server and on a need to know basis. Consortium members will have access rights to add Data to the identity Database. No editing or reading rights will be granted to them to prevent alteration/disclosure of private Data, if a specific permission is not granted by the respective user participant.

All technical partners participating in SECONDO have previous experience in storing and processing user Data. This implies that all of them have the appropriate competence and infrastructure to address the processing of SECONDO user Data. This will assure secure storage, delivery and access of Personal Data, as well as managing the rights of the users. In this way, there is complete guarantee that the accessed, delivered, stored and transmitted content will be managed by the right persons, with well-defined rights, at the right time.

Depending on each Dataset the Data archiving and preservation procedures that will be put in place for long-term preservation of the Data will be responsibility of the corresponding Data Controller. This includes the indication of how long the Data should be preserved, what is its approximated end volume, what the associated costs are and how these are planned to be covered. As mentioned before, privacy-preserving smart contracts will be leveraged to hide sensitive client information and meanwhile, secure encryption technique will be considered in Data storage.

5. Ethical Aspects and Privacy and Security Requirements

Privacy and Data protection are fundamental rights, which needs to be protected. Privacy can mean different things in different contexts and cultures. It is therefore important to detail the purpose of the research according to the different understandings of privacy. In the context of research, privacy issues arise whenever data relating to persons are collected and stored, in digital form or otherwise. The main challenge for research is to use and share the data, and at the same time protect personal privacy. Moreover, Data protection aims at guaranteeing the individual's right to privacy. It refers to the technical and legal framework designed to ensure that Personal Data are safe from unforeseen, unintended or malevolent use. Data protection therefore includes e.g., measures concerning collection, access to Data, communication and conservation of Data. In addition, a Data protection strategy can also include measures to assure the accuracy of the Data. In the context of research, privacy issues arise whenever Data relating to persons are collected and stored, in digital form or otherwise. The main challenge for research is to use and share the Data, and at the same time protect personal privacy [7]. In order to ensure respect for Data protection and privacy, the European University Institute (EUI) has adopted a Data Protection Policy [14] that must be respected by all EUI members and which is inspired by the EU Data protection rules. If the research is exclusively carried out at the EUI's premises, the applicable Data protection framework is the EUI's Data Protection Policy, complemented when necessary by local privacy and Data protection laws.

In legal terms, ‘processing of Personal Data’ means: ‘any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organisation, storage, adaptation or alteration, retrieval, consultation, use, disclosure, transmission,

dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction’ [11]. Additionally, if a study will use Personal Data on an individual who can be identified, this may fall under the remit of the Data Protection Act 2018. It is the Host Institution’s responsibility to ensure that the provisions of the Act are met [15].

Article 2 of the EUI’s Data Protection Policy indicates some categories of data that are more sensitive than other personal data and therefore require special treatment (‘Sensitive Data’). Sensitive Data are those revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, genetic Data, biometric Data, Data concerning health and Data relating to sexual orientation or activity. As a rule, the processing of sensitive data is prohibited. However, Article 8 of the EUI’s Data Protection Policy provides for specific circumstances, which allow for the processing of sensitive data. The most common in research is upon the **Data subject’s explicit consent**.

As mentioned before, an important area of concern with Social Media Website (SMW) research is the protection of confidentiality. Similar to other types of research involving survey or interview Data, protection of participant identities is critical. Website research may initially be perceived as lower risk, because participant information can be collected in absence of some protected information such as address or phone number. Online Data can present increased risks; studies that publish direct text quotes from an SMW may directly identify participants. Entering a direct quote from an SMW into a Google search engine can lead to a specific Web link, such as a link to that person’s LinkedIn profile, and thus identify the participant.

SECONDO will not collect or process personal data to conduct its research. Therefore, a data protection impact assessment shall not be conducted.

Nevertheless, the SECONDO consortium and the advisory board will monitor closely the activities of the project and in case there is a requirement for collecting/processing personal data a risk evaluation will be conducted.

The Ethics Board is formed by the following persons, who are closely involved in ethical procedures within the project and to whom any issue arising during the project, especially involving end-users would be reported.

Partner	Name
UPRC	Christos Xenakis
SURREY	Emmanouil (Manos) Panaousis
CUT	Michael Sirivianos
UBI	Dimirtios Alexandrou
LST	Evangelos Kotsifakos
CRO	Nikos Georgopoulos
FOGUS	Dimitrios Tsolkas

The Ethics Board will define a proper procedure for informing the Data subjects about any ethical related issue (privacy, GDPR compliance etc), its possible consequences and how their fundamental rights will be safeguarded. The Ethics Board will make sure that the Data subjects have understood

this information by asking for their consent. These procedures will be kept in a dedicated git repository that will only be accessible by the Ethics Board and the SECONDO Platform Administrators. This repository was defined in deliverable D1.1 - Quality Assurance Plan, while the procedures will be reported in deliverable D6.2 – Platform Assessment.

As mentioned in D8.1_GEN-requirement_no2, Professor **Konstantinos Lambrinoudakis**, as a member of the Hellenic Data Protection Authority (HDPa) he is participating in privacy and GDPR related events, conferences and talks.

5.1 General Data Protection Regulation (GDPR)

If Data controllers intend to use Personal Data that were collected from a previous research project, they must provide details regarding the initial Data collection, methodology and informed consent procedure, to the extent that consent is the appropriate legal basis. They must also confirm that they comply with the Data protection principles and that they for example have permission from the Data controller to use the Data in the SECONDO project.

Where the planned use of Data is predicated on the ‘legitimate interests’ of the Data controller, the nature and purpose of the Dataset must be set out in detail, together with the safeguards (e.g. anonymisation or pseudonymisation techniques) that warrant its use in SECONDO project (GDPR , Article 89).

If Data controllers intended Data processing is based on national legislation or international regulations authorising the research, or a demonstrable overriding public interest (e.g. public health, social protection) allows to use a particular Dataset, they must make reference to the relevant Member State or Union law or policy.

Regarding [16], one of the best ways to mitigate the ethical concerns arising from the use of Personal Data is to anonymize them so that they no longer relate to identifiable persons. Data that no longer relate to identifiable persons, such as aggregate and statistical Data, or Data that have otherwise been rendered anonymous so that the Data subject cannot be re-identified, are not Personal Data and are therefore outside the scope of Data protection law. However, even if the plan is to use only anonymized Datasets, significant ethics issues may still be raised, and the Database would become rather unusable. These ethics issues could relate to the origins of the Data or the manner in which they were obtained. Therefore, the source of the Datasets intended for use must be specified and any ethics issues that arise must be addressed. The potential for misuse of the research methodology or findings must also be considered, as well as the risk of harm to the group or community that the Data concern.

Where it is necessary to retain a link between the research subjects and their Personal Data, Data controllers should, wherever possible, pseudonymize the Data in order to protect the Data subject’s privacy and minimize the risk to their fundamental rights in the event of unauthorized access. However, in SECONDO, because of using only simulated and/or synthetic Data for the purposes of validation during the project, no pseudonymisation will be used. Data will be protected by other means of Data security.

When Personal Data moves across borders outside the Union it may put at increased risk the ability of natural persons to exercise Data protection rights in particular to protect themselves from the unlawful use or disclosure of that information.

National authorities in the Member States are being called upon by Union law to cooperate and exchange Personal Data so as to be able to perform their duties or carry out tasks on behalf of an authority in another Member State. cross-border cooperation and agreements to deliver effective Data protection are essential, particularly if the EU is to maintain its values and uphold its principles.

To achieve this, the European Data Protection Supervisor (EDPS) regularly interacts with EU and international Data Protection Authorities (DPAs) and Regulators to influence and develop cross-border enforcement.

5.2 Security and Authentication Legislation

- **The Directive (EU) 2016/1148 on Network and Information Security (NIS Directive)**, provides legal measures to boost the overall level of cybersecurity in the EU and is the first piece of EU-wide cybersecurity legislation. The goal of the NIS Directive is to establish a minimum level of (cyber) security for network and information systems across the EU, particularly for those operating essential services. The Directive addresses specifically operators of essential services and digital service providers. However, it is up to the Member States to assess which entities meet the criteria of the definition of an operator of an essential service. Member States must identify the operators of essential services.
- **The Regulation on ENISA, the "EU Cybersecurity Agency", and repealing Regulation (EU) 526/2013, and on Information and Communication Technology cybersecurity certification (Cybersecurity Act) [17]** is adopted by the European Parliament on the 12th of March 2019. This Act aims to strengthen Europe's cybersecurity, by replacing existing national cybersecurity certification schemes in European schemes which will define security objectives. For one thing, SECONDO will comply with the Cybersecurity Act's principles of security by design and by default.

6. The SECONDO FAIR Dataset Template Questionnaire

This section gathers all FAIR forms completed with information from Data Controllers. The following questionnaires have been addressed by the responsible partners with a level of detail appropriate to the project's progress. The SECONDO FAIR Dataset Template Questionnaire (Table 7-1) includes a set of questions that all Data Controllers are required to fill in for each Dataset [3], [7], [8]. The questionnaire template has been reviewed by the Project Manager, Ethics Board for completeness and compliance with the FAIR DMP directives.

As mentioned before, the DMP is intended to be a living document in which information can be made available gradually through successive updates as the implementation of the project progresses. The Data Controllers will be responsible to update their respective tables every time significant changes occur.

Table 6-1: SECONDO FAIR Dataset Template Questionnaire

Project Acronym		Project Number
SECONDO		823997
	Description	
Title		Name of the Dataset <i>Please provide a meaningful name so that we can refer to it unambiguously in the future</i>
Task		SECONDO task/subtask where Dataset was generated <i>Describe the overall setting of the use case in a scenario style, clarify how things will really happen during pilots, who will be involved, who will benefit, etc.</i>
Data owner/controller		Names and addresses of the organizations or people who own/control the Data
Time period covered by the Dataset		Start and end date of the period covered by the Dataset
Subject		Keywords or phrases describing the subjects or content of the Data
Language		All languages used in the Dataset
Variable list and codebook		All variables in the Data files, with description of the variable name, length, type, values
Data quality		Description of Data quality standards and procedures to assure Data quality
File inventory		All files associated with the project, including extensions
File formats		Format of the file
File structure		Organization of the Data file(s) and layout of the variables, where applicable
Necessary software		Names of any special-purpose software packages required to create, view, analyse, or otherwise use the Data
Details on the procedures for obtaining informed consent		Please give details on the procedures for obtaining informed consent from the Data subjects (e.g. providing an information sheet together with the consent form). In case of children/minors and/or adults unable to give informed consent, indicate the tailored methods used to obtain consent. According to the H2020 Guidelines, if the Data subjects are unable to give consent in writing, for example because of illiteracy, the non-written consent must be formally documented and independently witnessed. Please explain how you intend to document oral consent. In the very exceptional case that it can't be recorded please give reasons. If you will use deception for another type of Data subjects, you must obtain retrospective informed and free consent as well as debrief the participants.

		Deception requires strong justification and appropriate assessment of the impact and the risk incurred by both researchers and participants.
Measures taken to prevent the risk of enhancing vulnerability/stigmatization of individuals/groups		<i>Please indicate any such protective measures (e.g. use of anonymization techniques, use of pseudonyms, non-disclosure of audio-visual materials, voice records, etc.)</i>
Description of the processing operations (i.e. what you do with Personal Data and how)		<p>Processing of ‘Personal Data’ means any operation or set of operations which is performed upon Personal Data, whether or not by automatic means, such as:</p> <ul style="list-style-type: none"> •Collection (digital audio recording, digital video caption, etc.) •Recording •Organization and storage (cloud, LAN or WAN servers) •Adaptation or alteration (merging sets, amplification, etc.) •Retrieval and consultation •Use •Disclosure, transmission, dissemination or otherwise making available (share, exchange, transfer, access to the Data by a third party) •Alignment or combination •Blocking, deleting or destruction, etc. <p><i>Please describe in detail the processing operations that you will perform for conducting your research and give detailed feedback on participants. Indicate also if a copy of notification/authorization for tracking or observation is required.</i></p> <p><i>Any type of research activity may involve processing of Personal Data (ICT research, genetic sample collection, research activities involving personal records (financial, criminal, education, etc.), lifestyle and health information, family histories, physical characteristics, gender and ethnic background, location tracking and domicile information, etc.)] any method used for tracking or observing.</i></p>

1. Data Summary

1.1 Purpose

1.2 Types and formats of Data
1.3 Re-use of existing Data
1.4 Origin
1.5 Expected size
1.6 Data utility
2. FAIR Data
2.1 Making Data findable (Dataset description: Metadata, persistent and unique identifiers e.g.)
2.1.1 Are the Data produced and/or used in the project discoverable with Metadata, identifiable and locatable by means of a standard identification mechanism?
2.1.2 What naming conventions do you follow?

2.1.3 Will search keywords be provided that optimize possibilities for re-use?
2.1.4 Do you provide clear version numbers?
2.1.5 What Metadata will be created?
2.2 Making Data openly Accessible
<i>which Data will be made openly available and if some Datasets remain closed, the reasons for not giving access; where the Data and associated Metadata, documentation and code are deposited (repository?); how the Data can be accessed (are relevant software tools/methods provided)?</i>
2.2.1 Which Data produced and/or used in the project will be made openly available as the default?
2.2.2 How will the Data be made accessible?
2.2.3 What methods or software tools are needed to access the Data?
2.2.4 Is documentation about the software needed to access the Data included?
2.2.5 Is it possible to include the relevant software?

2.2.6 Where will the Data and associated Metadata, documentation and code be deposited?
2.2.7 Have you explored appropriate arrangements with the identified repository?
2.2.8 If there are restrictions on use, how will access be provided?
2.2.9 Is there a need for a Data access committee?
2.2.10 Are there well described conditions for access?
2.2.11 How will the identity of the person accessing the Data be ascertained?
2.3 Making Data Interoperable <i>(which standard or field-specific Data and Metadata vocabularies and methods will be used)</i>
2.3.1 Are the Data produced in the project interoperable?
2.3.2 What Data and Metadata vocabularies, standards or methodologies will you follow to make your Data interoperable?

2.3.3 Will you be using standard vocabularies for all Data types present in your Data set, to allow interdisciplinary interoperability?
2.3.4 In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?
2.4 Increase Data re-use <i>(which Data will remain re-usable and for how long, is embargo foreseen; how the Data is licensed; Data quality assurance procedures)</i>
2.4.1 How will the Data be licensed to permit the widest re-use possible?
2.4.2 When will the Data be made available for re-use?
2.4.3 Are the Data produced and/or used in the project useable by third parties, in particular after the end of the project?
2.4.4 How long is it intended that the Data remains re-usable?
2.4.5 Are Data quality assurance processes described?
3 Allocation of resources
3.1 What are the costs for making Data FAIR in your project?

3.2 How will these be covered?
3.3 Who will be responsible for Data management in your project?
3.4 Are the resources for long term preservation discussed?
4 Data security
4.1 What provisions are in place for Data security? Please indicate any methods considered for secure Data storage and transfer of sensitive Data.
<i>Please indicate any methods considered for Data storage.</i>
4.2 Is the Data safely stored in certified repositories for long term preservation and curation?
5 Ethical aspects / Protection of Personal Data notification of processing operations
5.1 Are there any ethical or legal issues that can have an impact on Data sharing?
5.2 Is informed consent for Data sharing and long-term preservation included in questionnaires dealing with Personal Data?

5.3 Name of the Processor(s) <i>Please indicate the names of any other natural or legal person that may process the Data. If processors can be categorised into groups please refer to them by groups and not necessarily by name, otherwise indicate their names.</i>
5.4 Lawfulness of Processing <i>Data Controllers must process only those Personal Data that are necessary during the project and for a specific purpose. Processing Personal Data that are not essential to the research may even expose Data Controllers to allegations of 'hidden objectives', i.e. processing information with the Data subjects' permission for one purpose and then use that information for another purpose, without specific permission.</i>
5.5 Categories of Data Subjects <i>Please indicate the categories of Data subjects involved in the processing operations of the project.</i>
5.6 Categories of Personal Data <i>Please list concretely the categories of Personal Data that you will process:</i> <ul style="list-style-type: none"> • Normal Personal Data: name, home address, e-mail address, location Data etc. • Sensitive Data: religious beliefs, political opinions, medical Data, sexual identity, etc.
5.7 Rights of Data subjects <i>Regarding Article 16 of the EUI's Data Protection Policy, Data subjects enjoy the following rights concerning their Personal Data:</i> <ul style="list-style-type: none"> • To be informed whether, how, by whom and for which purpose they are processed • To ask for their rectification, in case they are inaccurate or incomplete • To demand their erasure in case the processing is unlawful or no longer lawful ('right to be forgotten') • To block their further processing whilst the conditions under letters b) and c) of this Article are verified. <i>Note: Please indicate how you will ensure the Data subjects' rights. E.g. participants will be free to withdraw at any time without justification. The Data collected prior to the withdrawal will be deleted. In such a case, you may need to ensure the erasure of the collected Data while maintaining anonymity. In order to do so, you may use a pseudonym for each participant ensuring that the key to the pseudonyms is password-protected and available only to the Data Controller.</i>
5.8 Safeguards taken to protect the Data subjects' identity.

Regarding Article 2 of the EU's DP Policy, Identifiable persons can be identified directly or indirectly, in particular by reference to an identification number or to one or more factors specific to their physical, physiological, genetic, mental, economic, cultural or social identity. Please provide details on the measures taken to avoid direct or indirect identification of the Data subjects, e.g. by using anonymisation techniques or pseudonyms. E.g. names of the Data subjects will not be disclosed, at any time, in audio recording and published material Pseudonyms (a reversible system of coding in order to be able to re-contact participants if needed) will be used in all documentation, and any additional information that may reveal the identity of participants will be concealed when publishing.

Destroy any residual information that could lead to the identification of participants at the end of the project. You must explain this procedure clearly to participants during the 'recruitment' process.

6. Other issues

6.1 Do you make use of other national/funder/sectorial/departmental procedures for Data management? If yes, which ones?

7. References

- [1] “Marie Skłodowska-Curie Actions | Horizon2020,” [Online]. Available: <https://ec.europa.eu/programmes/horizon2020/en/h2020-section/marie-sklodowska-curie-actions>.
- [2] “Open research data and data management plans information for erc grantees,” February 2018. [Online]. Available: https://erc.europa.eu/sites/default/files/document/file/ERC_info_document-Open_Research_Data_and_Data_Management_Plans.pdf.
- [3] “Guidelines on FAIR Data Management in Horizon 2020,” European Commission/Directorate-General for Research & Innovation, 2016.
- [4] “GDPR EU,” [Online]. Available: <https://www.gdpreu.org/the-regulation/key-concepts/data-controllers-and-processors/>.
- [5] ProgrammeEuropeanCommissionH2020, “Guidelines on FAIR Data Management, Directorate- General for Research & Innovation,” [Online]. Available: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.

- [6] “Data management, European Commission, Research and Innovation, Participant Portal H2020 Oline Manual,” [Online]. Available: https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm.
- [7] “Guide on Good Data protection practice, This Guide should be consulted in parallel with the EUI’s Data Protection Policy as well as with the EUI Code of Ethics in Academic Research.,” [Online]. Available: <https://www.eui.eu/Documents/ServicesAdmin/DeanOfStudies/ResearchEthics/Guide-Data-Protection-Research.pdf>.
- [8] “ERC Open Research Data Management Plan Template,” April 2017. [Online]. Available: <https://erc.europa.eu/content/erc-data-management-plan-template>.
- [9] “Zenodo,” an interdisciplinary open data repository service maintained by CERN, Geneva. Datasets can be located via the Zenodo search engine., [Online]. Available: <https://www.zenodo.org/>.
- [10] “Data sharing guidelines | Cancer Research UK,” [Online]. Available: <https://www.cancerresearchuk.org/funding-for-researchers/applying-for-funding/policies-that-affect-your-grant/submission-of-a-data-sharing-and-preservation-strategy/data-sharing-guidelines>.
- [11] “Guide on Good Data protection practice, This Guide should be consulted in parallel with the EUI’s Data Protection Policy as well as with the EUI Code of Ethics in Academic Research.,” [Online]. Available: <https://www.eui.eu/Documents/ServicesAdmin/DeanOfStudies/ResearchEthics/Guide-Data-Protection-Research.pdf>.
- [12] [Online]. Available: https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/oa-pilot/h2020-hi-erc-oa-guide_en.pdf.
- [13] Megan A. Moreno, Natalie Goniou, Peter S. Moreno, Douglas Diekema, “Ethics of Social Media Research: Common Concerns and Practical Considerations,” *CYBERPSYCHOLOGY, BEHAVIOR, AND SOCIAL NETWORKING*, vol. 16, no. 9, 2013.
- [14] “PRESIDENT DECISION No. 10/2019 of 18 February 2019,Regarding Data Protection at the European University Institute (EUI),” 209. [Online]. Available: <https://www.eui.eu/Documents/AboutEUI/Organization/DataProtection/PresDecision10-2019-DataProtection.pdf>.

- [15] “Marie Curie Research Grants Terms and Conditions,” [Online]. Available: <https://www.mariecurie.org.uk/globalassets/media/documents/research/funding-for-research/marie-curie-research-programme/marie-curie-research-grants-terms-and-conditions.pdf>. [Accessed 2019].
- [16] EuropeanCommision(EC), “Ethics and data protection by,” 2018. [Online]. Available: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-data-protection_en.pdf.
- [17] “REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL,” European Commission, 2018.

8. Appendix

Ethics Issues tables that have to be filled in by the SECONDO Data Controllers.

Table 8-1: Protection of Personal Data

Protection of Personal Data	Information to be provided	Documents to be provided
Does your research involve Personal Data collection and/or processing?	<p>Description of the technical and organizational measures that will be implemented to safeguard the rights and freedoms of the Data subjects/research participants – including procedures for Data collection, storage, protection, retention, transfer, destruction or re-use</p> <p>Description of the security measures that will be implemented to prevent unauthorized access to Personal Data or the equipment used for processing, methods of storage and exchange (LAN, cloud, etc.)</p> <p>Description of the anonymization/ pseudonymization techniques that will be implemented or explanation on why the research Data will not be anonymized/ pseudonymized</p> <p>Detailed information on the informed consent procedures regarding Data processing</p>	<p>Data Management Plan, if required</p> <p>Data Protection Impact Assessment, if required</p> <p>Informed Consent Forms, Information Sheets/Specific Privacy Statements, other consent documents (opt-in processes, etc.) (if relevant).</p> <p>Copy of authorization for Data transfer from non-EU country (if required) or any other legal basis under Chapter V of the General Data Protection Regulation 2016/679.</p>
If YES:	Does it involve the collection or processing of special categories of Data (Data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic Data, biometric Data for the purpose of uniquely identifying a natural person, Data concerning health or Data concerning a natural person's sex life or sexual orientation)	Check if special derogations pertaining to the rights of Data subjects or the processing of genetic, biometric and/or health Data have been established under the national legislation of the country where the research takes place and submit a declaration of compliance with respective national legal framework(s). Justification for the processing of special categories of Data must be included in the grant agreement
	Does it involve tracking or observation or profiling of participants? (profiling)	In case the research involves profiling, the beneficiary must provide explanation how the Data subjects will be informed of the existence of the profiling, its possible consequences and how their fundamental rights will be safeguarded
Does your research involve further processing of previously collected Personal Data ('secondary use')? (including use of pre-existing Data sets or sources, merging existing Data sets)?	<p>Confirmation that the Data used in the project is publicly available and can be freely used for the purposes of the project</p> <p>Confirmation that the beneficiary has lawful basis for the Data processing and that the appropriate technical and organizational measures are in place to safeguard the rights of the Data subjects</p>	<p>Evidence of public access and terms of use (e.g. print screen from website).</p> <p>Informed consent forms, Information sheets, other consent documents.</p> <p>Copies of permissions (if required).</p>

Table 8-2: Human Participation

Humans		Information to be provided	Documents to be provided
Does your research involve human participants?		Please provide information in one of the subcategories below:	
If YES:	- Are they volunteers for social or human sciences research?	Details on recruitment and informed consent procedures.	Copies of relevant Ethics Approvals. Informed Consent Forms. Information Sheets.
	- Are they persons unable to give informed consent?	Details on the procedures used to ensure that there is no coercion on participants.	
	- Are they vulnerable individuals or groups?	Details on the type of vulnerability. Details on recruitment and informed consent procedures.	
	- Are they children/minors?	Details on the age range. Details on children/minors' assent procedures. Describe the procedures to ensure welfare of child/minor. Justification for involving minors.	
	- Are they patients?	Details on the nature of disease/condition/disability. Details on recruitment and informed consent procedures.	
	- Are they healthy volunteers for medical studies?	Details on incidental findings policy.	Copies of relevant Ethics Approvals.

Table 8-3: Other Ethics Issues

Other Ethics Issues	Information to be provided	Documents to be provided
Are there any other ethics issues that should be taken into consideration? <i>Please specify</i>	Any relevant information.	Any relevant document.