RAMONES
Radioactivity Monitoring in Ocean Ecosystems

Deliverable

Document ID: D6.6 Data management plan

31/10/2021

Document Info

Project Information			
Acronym	RAMONES		
Name	Radioactivity Monitoring in Ocean Ecosystems		
Start Date	1 Jan 2021	End Date	31 Dec 2024
Program	H2020-EU.1.2.2. - FET Proactive		
Call ID	H2020-FETPROACT-2020-2	Topic	FETPROACT-EIC-08-2020 - Environmental Intelligence
Grant No	101017808	Instrument	RIA
Document Information			
Document Id	D6.6		
Document Title	Data management plan		
Due Date	31-May-2021	Delivery Date	31-November-2022
Lead RAMONES partner	NKUA (1)		
RAMONES partners (part.)	NKUA(1), IST-ID(2), EVOL(3), ENS(4), NTUA(5), PLOA(6), UDUR(7), UCA(8)		
Editor(s)	Theodoros Mertzimekis (NKUA)		
Authors (s)	Paraskevi Nomikou, Eleni Petra, Effie Zafeirakopoulou, Stavroula Kazana, Konstantina Bejelou, Theodoros Mertzimekis, (NKUA), Antonio Pascoal, Luis Sebastiao, David Cabecinhas (IST-ID), Konstantin Kebkal (EVOL), Javier Escartin (ENS), Konstantinos Karantzalos, Valsamis Ntouskos (NTUA), Angelos Mallios (PLOA), Lydia Maigne (UCA)		
Contributor (s)			
Reviewer(s)	Angelos Mallios (PLOA)		
Workpackages	WP6 - Project Coordination and Innovation Management		
Version	V1.0	Stage	Submitted
Distribution	Public		
Keywords	Data management, Innovation Management, Intellectual Property Rights	Type	ORDP

Document Change Record

Version	Date	Description	Editor
1.0	08/11/2022	Structure/text update	NKUA
1.1	14/11/2022	Deliverable resubmitted	NKUA

Disclaimer

RAMONES is a European Innovation Council (EIC) FET Proactive project in the Environmental Intelligence Scope B, related to radically novel approaches to resilient, reliable and environmentally responsible in-situ monitoring, funded by European Union under Horizon 2020 FET proactive programme, via grant agreement No. 101017808.

RAMONES project's main objective is to close the current marine radioactivity under-sampling gap and foster new interdisciplinary research in ocean ecosystems. RAMONES will invest a significant effort to provide tools to enable long-term data acquisition missions, rapid deployments, low cost per information byte, and propose new AI and Robotics-driven and supported methodologies, being ambitious to eventually offer scaled-up solutions to researchers, policy makers and communities. All of these may be achieved by combining state-of-the-art (SoA) methodologies and equipment from various disciplines in a well-balanced synergy, and designing new and effective methodologies targeting the marine environment, which will provide efficient response to existing natural and man-made hazards, and shape future policies for the global population. RAMONES will additionally contribute to shaping a blueprint on Environmental Intelligence in the EU and worldwide.

This document has been produced through funding from the European Commission. The content of this document is a product of the RAMONES project Consortium and it does not necessarily reflect the opinion of the European Commission. The editor, author, contributors and reviewers of this document have taken any available measure in order for its content to be accurate and lawful. However, neither the project consortium as a whole, nor the individual partners that have implicitly or explicitly participated in the creation and publication of this document may be held responsible for any damage, financial or other loss or any other issue that may arise as a result of using the content of this document or any of the project outputs that this document may refer to.

The European Union (EU) was established in accordance with the Treaty on the European Union (Maastricht). There are currently 27 member states of the European Union. It is based on the European Communities and the member states' cooperation in the fields of Common Foreign and Security Policy and Justice and Home Affairs. The five main institutions of the European Union are the European Parliament, the Council of Ministers, the European Commission, the Court of Justice, and the Court of Auditors (http://europa.eu.int/).

List of acronyms and definitions

Acronym	Description
DMP	Data Management Plan
EC	European Commission
EOSC	European Open Science Cloud
EU	European Union
GA	General Assembly
KPI	Key Performance Indicator
TB	Technical Board
URL	Uniform Resource Locator
WP	Work package
WG	Working Group
Publication	Means any conference paper, article for a journal, portion of a book, poster, presentation, broadcast or other means of disclosure related to the Project or which contains Confidential Information, Results or Background but excludes any thesis of a student enrolled with one of the parties.
Results	Mean any tangible or intangible output of the Project, such as data, knowledge and information whatever their form or nature, whether or not they can be protected, which are generated in the performance of the Project as well as any attached rights, including intellectual property rights.
Software	means sequences of instructions to carry out a process in, or convertible into, a form executable by a computer and fixed in any tangible medium of expression.

Table of Contents

Abstract

This report entitled "*Data management plan*" (Deliverable D6.6) summarizes RAMONES activities and actions towards comprehensive Data Management aligned to EC guidelines and FAIR Data principles[1]. In particular, D6.6 describes RAMONES data management plan and discusses the data management lifecycle to answer questions regarding both Ethical and Research data management. Moreover, specific questions are addressed regarding the overall Research Data Management plan (Section 3), including:
- how can data be gathered, ingested and exploited by RAMONES project
- how long will datasets and products be kept
- 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).

---

[1] Wilkinson, M. D. et al., (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3,** 160018 (2016). doi.org/10.1038/sdata.2016.18

1. Introduction

1.1. Context

The report is the outcome of the work performed in "T6.3 Data and Innovation Management" of the WP6 "*Project Coordination and Innovation Management*" of the H2020 RAMONES project. This report is the deliverable D6.6 "*Data management plan*" of the project. Overall, as RAMONES embraces future emerging technologies numerous challenges arise with respect to the management of the outcomes and innovation. As such the task T6.3 is tackling a number of topics related to effective innovation management and avoidance of any relevant conflict. The task will deal with challenges related to (a) software (b) scientific knowledge (c) exploitation aspects and (d) data. For data, RAMONES will adopt the Data Management Platform offered by OpenAIRE /EUDAT and FAIR data principles through specific Data Management Plan. RAMONES will collect and utilize numerous datasets from a variety of sensors and detectors. The focus of this report is on data management and related procedures. Moreover, software licence management includes all activities related to adopting Free and Open Source Licences in order to implement RAMONES solutions on a high level. The RAMONES DMP, generated through the ARGOS DMP service[2], provided by OpenAIRE, is detailed in the appendix of this report.

1.2. Structure of the document

The rest of this document is structured as follows:

- Chapters 2-4 describe the data management principles within RAMONES including a reference to the collected data, the strategy as well as ethics and policies.
- Chapter 5 concludes this report.
- The Appendix provides the RAMONES DMP

---

[2] https://argos.openaire.eu/about

2. RAMONES Data Types and Lifecycle

2.1. Data Description and Data Types

In the framework of RAMONES a number of instruments for sensing radioactivity in the marine environment will be designed, developed and validated. In particular, the instruments are the following ones:

GASPAR - Gamma Spectrometer for Marine Radioactivity Studies. A new multi-component γ-ray spectrometer for extended operation in extreme underwater environments.

SUGI - Submarine Gamma Imager. A novel real-time underwater γ-radiation imager.

αSPECT - Underwater Alpha Spectrometer. A novel spectrometer will be constructed for monitoring alpha radiation in the marine environment.

CHERI - Cherenkov Imager. A novel imager for underwater Cherenkov radiation detection.

γ-Sniffers – portable gamma radiation spectometers. Lightweight Cadmium-Zinc-Telluride (CZT) / Silicon Photomultiplier (SiPM) spectrometers.

RAMONES instruments, therefore, will acquire and process

- ✔ 1D signals (spectra) in the form of single-column datasets (histograms) in various format e.g., ASCII files.

- ✔ 2D signals (images) imaging per pixel radioactivity levels in the FOV of the camera

Additionally, other kind of data that are collected, processed and exchanged among RAMONES nodes (Gliders/Slocums, Benthic Lab, USV, Control room) include:

- ✔ Positioning/Localisation data

- ✔ Communication data

- ✔ Water quality data (e.g. CTD)

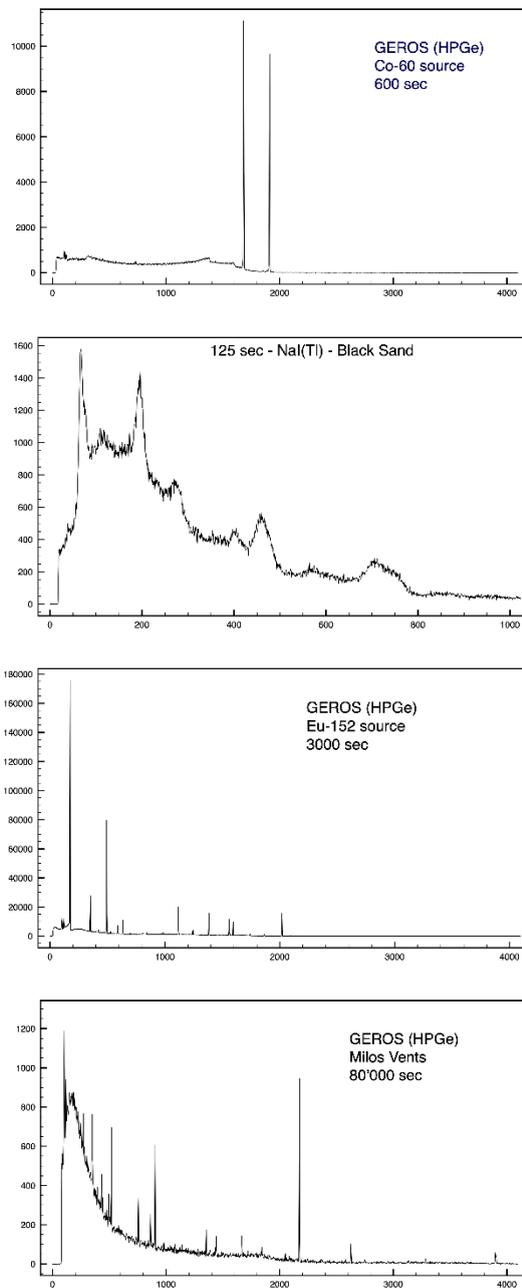- ✔ Navigation and control data etc
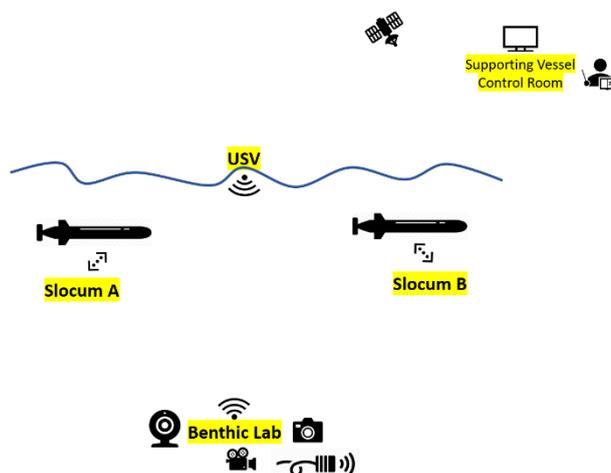
Data management plan

Figure 1. *Indicative examples of the expected 1D signals (spectra) gather by RAMONES spectrometers.*
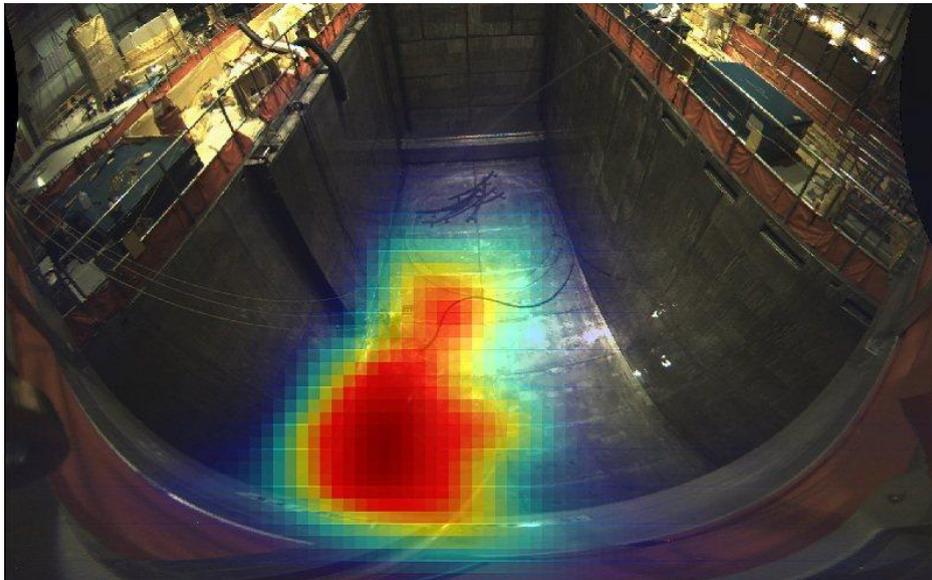
Figure 2. *Indicative example (laboratory conditions) of the expected 2D signals (images) gather by RAMONES spectrometers. Courtesy of H3D Inc.*

All these data are the internal datasets collected by RAMONES while their processing within RAMONES will deliver additional datasets and products (e.g., 2D, 3D maps).

The details of the datasets which will be handled in RAMONES are also presented in the DMP of Deliverable D6.6 "Data Management Plan".

2.2. Source Code

Software has a special role in the Open Science landscape in EU H2020 program. Software is the carrier of intelligence and research-based knowledge in actionable form. Being one of the stops where data get processed, transformed or presented, it has a very important role in the development of new knowledge and as such access to it is essential for carrying out quite a few of the objectives of Open Science (repeatability, reproducibility, transparency).

Although Open Science is best served by open source code access, as it imposes the least restrictions on how to exploit the code and the research behind it, several paradigms and cases may be presented that do not substantially compromise its overall objectives.

RAMONES follows an open, flexible approach with respect to software. Source code in RAMONES is managed according to partner policies that dictate licensing, control, authoring, versioning and management. RAMONES establishes the policies that assure quality of the delivered software. Although it is not a mandate of the project, the majority of code in RAMONES is licensed under an OSS (Open Source Software) scheme. As such, it may support a depositing policy that will allow researchers to inspect, replicate, reproduce and reuse both the code and the research behind it.

In general – assuming that respective license allows such a distribution - the source code of RAMONES software will be deposited in Zenodo (https://zenodo.org/) as a minimal assembly of information which consists of:

- A README file that contains at least the name of the software package, a short description, a reference to the project, software authors, other contributors, contact point(s), links to any active showcasing or servicing deployments and, if present, links to the code versioning system used for the code and binary releases.
- A LICENSE file that contains both the copyright and the licensing information of the source code.
- The source code itself.
- If the source code contains 3rd party works that cannot be separated, then those should be contained along with their relevant licensing and copyright information.
- The code should be packaged in a compressed form, excluding binary elements (i.e. executables), deployment specific/configuration files, sensitive information, passwords, etc.

As it is the common practice in software code management, RAMONES partners are already using solutions for code versioning and distribution. Those are not affected by the policy for Zenodo deposition of source code, which is selected for two main reasons:

a. to empower the long-term preservation policy of source code independently of the sustainability and durability of the various source code management systems.
b. to align the form of source code management planning independently of technologies and policies used by each partner, i.e. formatting, access rules, protocols, etc.

Regarding the granularity of software deposits those shall be defined by each partner at a level high enough to enclose self-contained entities.

Regarding timeline of deposits, the suggestion is that those should take place at time intervals that the software is considered mature to conduct research via its use. At least

one deposit at the end of the project is suggested for Open Source services. Coordinated deposits at milestones is the best option.

Summing up the Source Code Management Planning, it is expected that at key points in the service lifetime, copyright owners will (a) extract an instance of the source code from their respective code version control systems (b) compress the code in a common format (e.g. tar, tar.gz, zip), (c) author and package a few files of metadata (README, LICENSE) (d) upload the information package to Zenodo, along with code metadata and (e) obtain a DOI for the delivery of the package.

2.3. Reports and Publications

RAMONES produces several types of presentation of its work. Those can be roughly summarized in the following:

- **Project report deliverables**, as described in its Description of Action. Those reports are of several confidentiality levels, yet most fall under the "Public" access class. As such they shall be published in open access repository and the EC project management platform, as supported by the relevant software. Confidential reports shall be only uploaded in EC project management platform and protected appropriately by it.
- **Scientific Publications**, that may be produced by any scientific and/or technological research sector of the project. Complying with project funding mandate, all scientific publications will be published under Open Access principles, Green or Gold compliance depending on the case, yet targeting the project established KPIs in place.
- **Project dissemination artefacts**, which may consist of web and/or press articles, presentations, webcasts, spectral datasets, etc. This will be delivered under an Open Access license, most preferably CC-BY-4.0[3]. Reasoned deviations may exist.
- **Training material**, which may consist of documents, presentations, webinars etc. Those will be delivered under an Open Access license. Owners are suggested to adopt one of Creative Commons 4.0 license, CC-BY-SA-4.0[4] most preferable, yet reasoned deviations may exist.
- **Technical documentation material**, which may consist of documents, wiki pages, diagrammatic descriptions etc. Those will be delivered under an Open Access license. Owners are suggested to adopt one of Creative Commons 4.0 license, CC-BY-SA-4.0 most preferable, yet reasoned deviations may exist.

---

[3] https://creativecommons.org/licenses/by/4.0/
[4] https://creativecommons.org/licenses/by-sa/4.0/deed.ast

3. Data Management Strategy in RAMONES

This section describes the Data Management Strategy applied to make RAMONES data Findable, Accessible, Interoperable and Reusable (FAIR) through detailing the internal policy and adopting Open Data guidelines and platforms (e.g., Argos https://argos.openaire.eu).

3.1. RAMONES Internal Consortium Policy

As the project embraces future emerging technologies, business value, open access, and numerous challenges arise with respect to the management of the information. As such the task aims to tackle a number of topics of key importance for the success of the work and avoidance of any conflict.

RAMONES is an early adopter of the first publicly available platforms to support actionable Data Management Planning, like Argos - OpenAIRE (https://argos.openaire.eu).

RAMONES will not only comply with the Open Access publication policy and the provisioning of a Data Management Plan, but also adopts a robust path to support the FAIR research data principle.

RAMONES Data Management plan follows the guidelines of EC to promote data FAIRness in describing the life cycle of the data being gathered, processed and/or generated by the project by collecting information on:

- the handling of (research) data during and after the end of the project
- the particular data sets to be collected, processed and/or generated and their nature
- the methodology, technologies and standards that will be applied for data packaging, handling, description, sharing etc.
- the licensing / restrictions on data managed by the project
- the policies with respect to data curation and quality
- the policy with respect to data preservation (esp. beyond end of the project).

Moreover, for the employed DMP, Argos is an open extensible service that simplifies the management, validation, monitoring and maintenance and of Data Management Plans. Argos allows actors (researchers, managers, supervisors, etc.) to create actionable DMPs that may be freely exchanged among infrastructures for carrying out specific aspects of the Data management process in accordance with the intentions and commitment of Data owners. In order to facilitate the adoption of the platform, RAMONES has engaged

Data management plan

OpenAIRE experts since the beginning of the project to provide support on the use of the platform both in live presentations and as web training. In particular, Argos is based on OpenDMP which is by itself a free and open source software used for actionable machine Data Management Planning, that closely follows RDA DMP WG developments and specifications.

Data Management Responsible

RAMONES does not hold a separate role for Data Management responsibilities. The responsibilities are shared by the entire consortium and few key project individuals will be responsible for providing RAMONES data management plan including:

✔ WP1 - Design and Develop Novel Underwater Gamma Radiation Instruments, Dr. Mertzimekis (NKUA)

✔ WP2 - Autonomous Marine Robotics for Collaborative Radioactivity Mapping, Dr. Pascoal (IST-ID)

✔ WP3 - Onsite Integration, Testing and Pilot Demonstrators, Dr. Mallios (PLOA)

✔ WP4 – Multi-modal Data Analytics and Environmental Modelling, Dr. Karantzalos (NTUA)

✔ WP5 – Citizen Awareness, Dissemination and Communication Activities, Dr. Nikolopoulos (UDUR)

As a result of the Data Management plan, the following will be created: i) RAMONES Inventory of Results: a database where all research results will be collected, ii) Outreach, Engagement and Liaison Activities Plans. In particular,

✔ **Creating an inventory of research results:** Collecting and storing research outcomes through RAMONES Inventory where all research results are listed.

✔ **Identifying and analysing innovation opportunities:** The innovation management through the Outreach Task Force will preliminary analyse each solution based on the information collected and mapped in the RAMONES Inventory. As part of analysing each solution, possible individual innovation management and IPR strategy could be investigated. This strategy includes the main steps in order to prepare for the exploitation of the solution, and also include the potential business models and elaborate the main steps for IPR.

✔ **Data/Publication Access Policies and Licensing:** WP leaders have to decide and list the core competences of each technology. After analysing the core competences, the innovation management and RAMONES decision bodies decide which of them should be:

- o Public: can be shared with the external stakeholders, and the users. Public information about the technologies does not hinder the added value of the product and does not jeopardize competitive advantage of the technology.
- o Protected by Non-disclosure Agreement (NDA): part of the core competence should be protected by NDA. This information is essential to share with external stakeholders, in order to foster collaboration and added value of the technology, however, it is an essential part of the competitive advantage, which should not be given away.
- o Protected by IPR: This should be negotiated with the involved project partners, the Technical and Project Manager.

Long term preservation and Archiving

RAMONES long-term preservation policy addresses both research artefacts and their metadata handled by its processes.

RAMONES strategy with respect to long-term preservation and archiving builds on the following cornerstones:

- Data and their metadata shall be preserved in forms usable by any offered solution at least for the contractual duration of the project. A five (5) years extended preservation is targeted as a best practice.
- Data and metadata repositories shall handle long term preservation and archiving on a best-effort manner, in which case they are required to offer transparency on the measures taken in the direction.
- Data and metadata managers (WP leaders) shall use open schemes (protocols/formats, etc.) that minimize the risk for long-term preservation braking changes to occur.
- Security policy applied by each system in place must be sufficient to support at least the most restrictive form of data access required to secure the data and metadata deposited.
- Data managers are responsible for clearance of rights and application of those in the underlying preservation system, as well as for ensuring that the protection features cover the minimum requirements posed by data copyright owners.
- In case additional/redundant depositing/cataloguing services exist that can enhance the options of data/metadata to survive long time frames, their use is

encouraged by the project, as long as those comply with the data protection rules settled by the data owner/manager.

Data security

Acting in the area of Open Science and data FAIRness, it is often misunderstood that data security is of lesser importance than in other cases, however this is far from being true. Not only data security is an essential aspect of Data Management Planning, but it also poses additional challenges due to the breadth and mix of policies that may be present in a discipline. In the general case, this creates a vast landscape of options that needs to be nailed down to the specific requirements in order to be strategically and technologically addressed in the context of RAMONES, that operates within a focused plan and a limited timescale.

Thus, starting from (a) requirements posed by the specific instrumentation, (b) need and relevance of data security in general and (c) technology relevance in the project, in the following we present the cornerstones of data security policy in RAMONES.

Data confidentiality

Data confidentiality addresses both data and metadata, yet policies differ substantially among those.

RAMONES strongly suggests **open publication of metadata** that may be reused in any foreseeable case inside and outside the project, independently of the access type granted to data. Yet temporal or internal reference data may well deviate from the rule. Those metadata depending on policies, may be openly accessible to 3rd parties.

In addition to delivery of metadata, it is essential that a declarative, comprehensible license is provided. If targeting open access under restrictions, the choice of one of CC 4.0 flavours is suggested, however data owners are free to select any existing or provide a new licensing scheme that covers their needs. It is evident though that along with the license a clear statement of ownership of the data is required. Actionable licensing is most preferable, however there are little chances that this is enforceable by several systems in place.

Regarding data, although RAMONES suggests an **open access data** delivery as the baseline policy, however as most data handled are results of prior or current work or

derivatives with restrictive licenses this will be applied at data owner's discretion. Thus, a data owner may opt for private, restricted or open access (with or without limitations) of data. In case data are kept private, those may not be accessed by other users inside or outside the RAMONES ecosystem. Restricted use may have different meanings depending on the developed solution. Open access on the other hand, can be provided also under restrictions, as for instance a requirement for user registration or limited use of data, etc.

Another aspect of data protection option is the establishment of an embargo period on fresh data, in order to protect the interests of the owner, yet not substantially impact their reuse by other scientist. Common practices of 6- or 12-months embargo (even 24 in several cases) are usually employed in the Open Access world.

Data dependability

Another aspect that data security policy in RAMONES refers to data trustworthiness. This is addressed with policies in several directions.

- Data integrity needs to be safeguarded both for data and metadata by storage service and catalogues and, if possible, further augmented by author seals.
- Authorship shall be well registered and tacked so that repudiation or any other relevant false claims are prevented.
- Versioning of data / metadata is ensured so that consumers may precisely address the artefact in question.
- Metadata description is adequate to support introspection of data and accurate consumption of the latter.
- Provenance is tracked adequately to support comprehension of data origins, generation procedures and ownership respect.
- Strong referencing of data is provided so that any further claims may be routed appropriately.
- Long term preservation is enabled allowing data to be dependable in the future.

3.2. Making RAMONES Data FAIR

According to the Guidelines on FAIR Data Management in Horizon 2020[5] this section of the Data Management Plan is expected to give answers to the following questions:

---

[5]

https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-dissemination_en.htm

✔ Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?
✔ Will be provided possibilities for re-use?
✔ Do you provide clear version numbers?
✔ What metadata will be created?

✔ In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Based on the above guidelines and on the directions adopted by RAMONES, the FAIR Data Management policy will be accomplished by the following principles and Implementation Plan:

Findable data. The RAMONES portal will offer browsing and keyword search functionality to enable users to discover and find relevant content and data in an intuitive manner. Additionally, Zenodo will be employed. Zenodo's metadata is compliant with DataCite's Metadata Schema minimum and recommended terms, with a few additional enrichments. Metadata of each record is indexed and searchable directly in Zenodo's search engine. Metadata of each record is sent to DataCite servers during DOI registration and indexed there. The data catalogues will support versioning. A globally unique and persistent ID (PID) will be assigned to every resource via Zenodo.

Accessible data. The RAMONES portal will provide continuous access to the developed solutions. Additionally, data and metadata in Zenodo will be retained for the lifetime of the repository.

Interoperable data. Zenodo uses JSON Schema as internal representation of metadata and offers export to other popular formats such as Dublin Core or MARCXML. Interoperability will be supported by the provision of dedicated APIs, if necessary, that allow the import and export of the content in standard formats (XML, JSON) and standard data models.

Re-usable data. The public content within the RAMONES will be available for download and re-use with no restrictions or embargo. The content will be available under permissive licenses, (CC-BY 4.0, CC-0 or comparable) but certain conditions (e.g. Noncommercial use=NC) and/or exceptions may also apply. In Zenodo, the code is open source, and built on the foundation of the Invenio digital library which is also open source. The work-in-progress, open issues, and roadmap are shared openly in GitHub. All

Data management plan

meta data is openly available under CC0 license, and all open content is openly accessible through open APIs when necessary. License is one of the mandatory terms in Zenodo's metadata and is referring to an Open Definition license [26]. Data downloaded by the users is subject to the license specified in the metadata by the uploader.

Open Access Policy. Scientific publications, that will be in focus by almost all project activities, set another field of regulation in the project. RAMONES will target 100% Open Access publications, being compliant with the mandate of H2020. More than 30% (current target) of those publications will be filed under Gold Open Access approach and the rest will be filed under Green Open Access (self-archived with the support of OpenAIRE repositories). The project will reserve financial resources for achieving open access in cases where this is not the de facto situation.

4. Data Management - Ethical Aspects and Privacy Policy

4.1. Privacy Policy and EU General Data Protection Regulation

According to the Guidelines on FAIR Data Management in Horizon 2020, this section of a Data Management Plan is expected to give answers to the following questions:

- ✔ Are there any ethical or legal issues that can have an impact on data sharing?
- ✔ Is informed consent for data sharing and long-term preservation included in any form of questionnaires dealing with personal data?

Privacy and Data Protection become relevant in the context of RAMONES where the data processed as part of the project analysis constitutes personal data as defined by the EU General Data Protection Regulation (GDPR) i.e. any information relating to an identified or identifiable natural person (data subject), whereby "an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person."

Research in RAMONES but not only is clearly not the preserve of academia. The interface, including the exchange of data, between research organisations and the wider research ecosystem are highly complex. Scientific publishers, developers, entrepreneurs, commercial, governmental and non-profit funding sources in the commercial, governmental and non-profit sectors, all potentially have a stake. In RAMONES, we share the same opinion and act accordingly with the European Data Protection Supervisor[6] and argue that the processing of personal data for the purposes of 'academic expression' implies: (i) processing directly linked to the freedom of academics to disseminate information, (ii) their freedom to distribute knowledge and truth without restriction, such as with publications, dissemination of research results, and (iii) the sharing of data and methodologies with peers and exchanges of views and opinions.

---

[6] European Data Protection Supervisor, 2020. Preliminary Opinion on data protection and scientific research

Concerning the ethical or legal issues that can have an impact on data sharing in RAMONES no ethical or legal issues are foreseen for any of the generated dataset.

With regards to the second basic question, if informed consent for data sharing and long-term preservation is included in questionnaires dealing with personal data, RAMONES is requesting compliance for collection and reuse of data. A dedicated work has taken place already within the 6 first months of the project concerning a full privacy policy for the protection of natural persons concerning the processing of personal data and on the free movement of such data (GDPR). It describes the policies and procedures set in place by RAMONES to protect the privacy of users, how the confidentiality of such information is ensured, laws, rights of data subjects and a communication path for further clarifications.

4.2. Lawful bases for data processing, including cases that RAMONES users are children

RAMONES solutions used by any internal or external user, should make it possible to obtain their consent or justify the processing of any dataset by means of the framework/ contract into which they have entered in order to use the relevant solutions.

Where the users are children, one will have to take into consideration Article 8 of the GDPR. Whereas RAMONES results/products will be offered through the cloud (i.e. online interactive service) directly to a child, typically a student, the processing of the personal data of the child shall be lawful where the child is at least 16 years old. Where the child is below the age of 16 years, such processing shall be lawful only if and to the extent that consent is given or authorised by the holder of parental responsibility over the child. The controller must make reasonable efforts to verify in such cases that consent is given or authorised by the holder of parental responsibility over the child, taking into consideration available technology. We note that Member States may provide by law for a lower age for those purposes.

4.3. Ethical and compliance implications of Machine Learning algorithms

In RAMONES we have considered ways to tackle the risks associated with the development and solution provision of AI-based algorithms both during development and operation phases. In particular, RAMONES will follow the set of EC non-binding Ethics guidelines for trustworthy AI[7] *i.e.,* implement 'human-centric' approach to AI that is respectful of European values and principles: *"The human-centric approach to AI strives to ensure that human values are central to the way in which AI systems are developed, deployed, used and monitored, by ensuring respect for fundamental rights, including those set out in the Treaties of the European Union and Charter of Fundamental Rights of the European Union, all of which are united by reference to a common foundation rooted in respect for human dignity, in which the human being enjoys a unique and inalienable moral status. This also entails consideration of the natural environment and of other living beings that are part of the human ecosystem, as well as a sustainable approach enabling the flourishing of future generations to come."*

RAMONES solutions (e.g. WP4) will consider all critical factors[8] of human agency and oversight, robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental well-being, accountability. Therefore, transparency in RAMONES is of paramount importance in order to ensure that AI is not biased and AI systems been deployed are explainable.

4.4. Data controllers and processors

RAMONES consortium members are expected to be in the role of the data controllers with regard to the data gathered from the solutions and instrumentation. However, subject to future commercial arrangements for the backend technologies exploitation, it may be possible for consortium members to act as data processors for their clients. In such cases, the responsibilities towards data subjects would primarily fall on the clients. However, under Article 28 of the GDPR [27], RAMONES consortium members would still need to assist the clients in fulfilling their obligations under GDPR towards data subjects whose data would be processed by the system.

---

[7] https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

[8] Tambiama Madiega, 2019. EU guidelines on ethics in artificial intelligence: Context and implementation, EPRS | European Parliamentary Research Service

5. Conclusions

This report D6.6 "*Data management plan*" (Deliverable D6.6) summarizes RAMONES activities and actions towards comprehensive Data Management processes within the RAMONES project. All aspects have been described in detail, aiming to set clear and transparent procedures minimizing any management risk.

If needed, this deliverable will be updated, when the GA body will decide, that the processes within RAMONES require changes in order to:

- ensure efficient data management processes within and outside RAMONES
- utilize the research results in a more efficient manner
- generate research results of higher quality
- help the project partners collaborate more efficiently
- exploitation processes require more elaborated data about the results

The update of this plan will be required especially when the progress of the project will reach a level, when it will be clear, what should be licenced from external parties.

Data management plan

References

-   ARGOS service [https://argos.openaire.eu/]

-   Creative Commons Licenses [https://creativecommons.org/licenses/]

-   DataCite's Metadata Schema [https://schema.datacite.org/]

-   Dublin Core Ontology [https://www.dublincore.org/resources/glossary/ontology/]

-   EC Digital Future Policies: Open Data. [https://ec.europa.eu/digital-single-market/en/open-data]

-   eInfraCentral [https://github.com/eInfraCentral]

-   EOSC ENHANCE project [https://www.eosc-portal.eu/enhance]

-   European Commission (2019). European Open Science Cloud (EOSC) strategic implementation plan. [https://op.europa.eu/s/n5Q1]

-   General Data Protection Regulation [https://gdpr-info.eu/]

-   Gezelter, D. (2009). What, exactly, is Open Science? [http://www.openscience.org/blog/?p=269]

-   Invenio digital library [https://invenio-software.org/]

-   Open Archives Initiative Protocol for Metadata Harvesting [https://www.openarchives.org/pmh/]

-   Open Definition license [https://opendefinition.org/]

-   OpenAIRE [https://www.openaire.eu/]

-   OpenDMP [https://gitlab.eudat.eu/dmp/OpenAIRE-EUDAT-DMP-service-pilot/]

-   Pontika, N. et al, (2015). Fostering open science to research using a taxonomy and an eLearning portal. In *Proc. Int'l Conf. on Knowledge Technologies and Data-Driven Business* (pp. 1-8). [DOI: 10.1145/2809563.2809571]

-   RDA DMP Common Standards Working Group [https://www.rd-alliance.org/groups/dmp-common-standards-wg]

-   SKOS Schema [https://www.w3.org/TR/2008/WD-skos-reference-20080829/skos.html]

-   Software Deposit: What to deposit, M. Jackson - Jisc [DOI: 10.5281/zenodo.1327325] [https://doi.org/10.5281/zenodo.1327325]

- Wilkinson, M. D. et al, (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3,** 160018 (2016). [DOI: https://doi.org/10.1038/sdata.2016.18]

- ZENODO [https://zenodo.org/]

APPENDIX - Data Management Plan Information

Data Management Plan Information

RAMONES project DMP

The data management plan of RAMONES project https://ramones-project.eu/

Funder

European Commission||EC

Grant

Radioactivity Monitoring in Ocean Ecosystems

Organisations

Ecole Normale Supériere de Paris, GiPStech Srl, Université Clermont Auvergne, PLOATECH, Durham University, National Technical University of Athens, Organization of the Instituto Superior Técnico for Research and Development, EVOLOGICS

Researchers

All project partners All project partners

Datasets

Title: RAMONES spectra

Template: Horizon 2020

Indicative data from RAMONES project

Dataset Description

1.1 Data Summary

1.1.1 What is the purpose of the data collection/generation and its relation to the objectives of the project?

Data management plan

["To obtain information"]

Comment: To collect data and observe levels of radioactivity

1.1.2 What types of data will the project generate/collect?

["sensor data", "observational (e.g., sensor data, data from surveys)"]

observations from sensors

1.1.3 What formats of data will the project generate/collect?

["Text files", "Numerical", "Instrument specific formats"]

in most cases the format will be ASCII

1.1.4 What is the origin of the data?

["Primary data"]

1.1.5 What is the expected size of the data?

GB (gigabyte)

Comment: 50

1.1.6 To whom might it be useful ('data utility')?

["Researchers", "Research communities", "Education"]

Both for the research community and academia

2.1 Reused Data

2.1.1 Will you re-use any existing data and how?

No

3.1.1 Making data findable, including provisions for metadata

3.1.1.1 Will you use metadata to describe the data?

No

3.1.1.3 Will your metadata use standardised vocabularies?

No

3.1.1.7 Will you use naming conventions for your data?

Yes

3.1.1.8 Please provide more details and examples on used naming conversions

date of data collections will be indicated

date of data collections will be indicated

3.1.1.9 Will you provide clear version numbers for your data?

Yes

3.1.1.10 Will you provide persistent identifiers for your data?

Yes

3.1.1.11 Persistent identifiers

DOI

3.1.1.12 Will you provide searchable metadata for your data?

No

3.1.1.15 Will you use standardised formats for your data?

Yes

7Zip format

3.1.1.18 Are the file formats you will use open?

Yes

3.1.1.20 Do supported open-source tools exist for accessing the data?

Yes

okeanos IAAS11

3.1.1.22 Will you provide metadata describing the quality of the data?

No

3.1.2 Making data openly accessible

3.1.2.1 Are there ethical or legal issues that can impact sharing the data?

No

3.1.2.2 Will your data be openly accessible?

some

3.1.2.3 Please provide the URL of the data which can be made available

RAMONES data

https://ramones-project.eu/

3.1.2.4 How will the data be made available?

["Project website", "Repository of Archive"]

Zenodo

3.1.2.6 Is the storage sufficiently secure for the data and does the storage provide backup and recovery procedures?

secure with backup and recovery

3.1.2.7 Are there any methods or tools required to access the data?

No

Data management plan

3.1.2.10 Will you also make auxiliary data that may be of interest to researchers available?

at end of project

3.1.3 Making data interoperable

3.1.3.1 Will you use a controlled vocabulary for your data?

No

3.1.3.2 Will you provide a mapping to more commonly used ontologies?

No

3.1.4 Increase data reuse

3.1.4.4 What internationally recognised licence will you use for your data?

Creative Commons: CC-BY 4.0, CC-0

3.1.4.5 Do you have documented procedures for quality assurance of your data?

No

3.1.4.8 Will you provide any support for data reuse?

Yes

3.1.4.9 How long do you intend to support data reuse?

More than 10 years

4.1 Allocation of resources

4.1.1 How will the cost of making your data findable, accessible, interoperable and reusable be covered?

["Use of national infrastructure", "Use of institution infrastructure"]

4.1.2 Will you identify a data manager to manage your data, if not who will be responsible for the management of your data?

Yes

4.1.3 Identify the people or roles that will be responsible for the management of the project data

Konstantinos Karantzalos

4.1.4 How do you intend to ensure data reuse after your project finishes?

["Institutional archive", "National archive"]

5.1 Data Security

5.1.1 What do you plan to do with research data of limited use?

Kept on secure, managed storage for limited time

6.1 Ethical aspects

6.1.1 Are there any ethical or legal issues that can have an impact on data sharing?

No

6.1.2 What are the methods used for processing sensitive/personal data?

["National laws"]

7.1 Other

7.1.1 Do you make use of other procedures for data management?

No