



## Develop SEA-EU Open Research Data Policies and Toolkit

Deliverable 5.2

(due to M36, Dec 2023)

# Contents

1. INTRODUCTION
2. DEVELOPMENT OF WORK
  - 2.1 SEA-EU Research Data Policy Framework
  - 2.2 Toolkit
3. DETAILED DESCRIPTION OF RESULTS
  - 3.1 SEA-EU Research Data Management Policy Framework
  - 3.2 SEA-EU Research Data Management Toolkit
4. RECOMMENDATIONS AND FOLLOW-UP

APPENDIX I: Opinion Survey on Research Data Management

APPENDIX II: reSEARCH-EU Opinion Survey on Research Data Management: Survey Results and Analysis

APPENDIX III: SEA-EU Research Data Management Policy Framework

## Deliverable identification

Deliverable No. and Title	D5.2. Develop SEA-EU Open Research Data Policies and Toolkit
Leader	Mr. Kevin J. Ellul
Related task(s)	Tasks 5.1 and 5.3
Authors	reSEArch-EU Open Research Data Officers
Dissemination level	Public
Due submission date	30/11/2023
Submission	15/11/2023
Project number	101017454
Start date of project	01/01/2021
Duration	36 months
<p>Task 5.2 aimed to prepare an Open Research Data Policy for the SEA-EU Alliance. The policy was developed and agreed to, with terms of reference that encompass the national, regional and university policies on data management and data sharing. The open research data policy entails a shared understanding of data sharing practices, either in place or planned, at each of the six participating SEA-EU universities, taking on board the roles of the individual university Libraries, their Knowledge Transfer Offices, Committees for Public Engagement with Research, Legal Offices, and other relevant bodies that champion, monitor and review such policies.</p> <p>A tool-kit was also developed, to provide guidelines for the implementation of the open research data policy.</p>	

## Versions and contributions history

1. Version	Date	Modified by	Reason
First draft of the deliverable report	19 <sup>th</sup> October 2023	Open Science Data Officers and Technical Managers	Reviewing by individual partners
Second draft of the deliverable report	12 <sup>th</sup> November 2023	Godfrey Baldacchino	Internal review
Final draft	28 <sup>th</sup> October 2023	Prepared by Kevin Ellul and Josianne Vella	Final version
Final version	12 <sup>th</sup> November 2023	Godfrey Baldacchino	Internal review
Final version	17 <sup>th</sup> November 2023	Kevin Ellul, Josianne Vella, Godfrey Baldacchino and Aldo Drago	Completion and submission

### 1. INTRODUCTION

The implementation of Research Data Management (RDM) Policies is a pertinent tool to facilitate and ensure that research data generated by SEA-EU institutions is organised and managed in a harmonised fashion which supports archiving, dissemination and sharing. To this effect, a pertinent deliverable of WP5 was to develop and implement a SEA-EU Open Research Data Management Policy Framework. Additionally, an RDM Toolkit was developed to facilitate and guide researchers with the sharing of their research data in line with the FAIR Data Principles.

### 2. DEVELOPMENT OF THE WORK

#### 2.1 SEA-EU Research Data Policy Framework

The purpose of a SEA-EU Research Data Management Policy Framework is to serve as a guideline that is expected to be adopted by the SEA-EU Alliance for the implementation of research data management policies within the respective SEA-EU universities. Subsequently, this Policy Framework aims to support SEA-EU institutions

with the adoption and implementation of Research Data Management Policies that facilitate the

appropriate curation and management of data, to secure its longevity and its potential to be shared.

To this effect, the UM assumed the responsibility to compile a draft of this Framework Policy. This was completed by March 2022. This draft was communicated with the SEA-EU Open Research Data Officers (ORDOs), feedback was gathered and the necessary amendments were affected. Moreover, as agreed during the Open Research Data Staff Week held in Malta between 24th and 26th January 2022, the UM was also responsible to design an opinion survey ([Appendix I](#)) intended to share the main elements of the

Policy Framework with the research communities of the SEA-EU universities so as to assess the acceptance of the various elements of the policy. This survey also served to involve researchers as co-constructors in the policy-making process and to raise general awareness. The opinion survey was distributed amongst academics of the six participating institutions with a time window of five weeks, from the third week of August till the end of September 2022. Three hundred and thirty-six (336) responses have been received. An analysis report ([Appendix II](#)) was compiled by the UM and discussed with the SEA-EU ORDOs. Where applicable, feedback gathered from the opinion survey was reflected within the draft Policy Framework. The final version of the Policy Framework ([Appendix III](#)) was then communicated with the various stakeholders and approval was sought from the Rectorates of the respective SEA-EU universities. The final version of the Policy Framework was presented by UM during the reSEArch-EU Open Science Staff Week in Cadiz held on 21<sup>st</sup> and 22<sup>nd</sup> March 2023.

Subsequently, the SEA-EU Research Data Management Policy Framework was formally presented to and approved by the SEA-EU Governing Board on 29<sup>th</sup> June 2023.

## 2.2 Toolkit

The scope of developing a SEA-EU Research Data Management (RDM) Toolkit is to support researchers through the entire lifecycle of Research Data Management. It provides information to assist researchers to effectively manage research data, as well as to develop suitable data management plans.

To this effect, UBO, UM and UNIST carried out extensive research to identify the main elements which constitute an RDM Toolkit. This was done by reviewing the literature, as well as consulting various RDM toolkits developed by various reputable research agencies and universities around the world. This was followed by an ORDOs' brainstorming session whereby the structure and the main elements of the SEA-EU RDM Toolkit were identified and agreed upon. It was also determined that the Toolkit should cater for the exigencies of the SEA-EU Alliance and to have an added value with respect to various other toolkits that already exist. The added benefits include the consolidation of Open Science policies, resources and tools available within the SEA-EU Alliance at institutional, national and European level; as well as, providing a thematic

approach pertaining to data repositories and data journals for the archiving and publishing of research data. Subsequently, the structure for the RDM Toolkit was developed and this

was presented by UNIST during the reSEARCH-EU Open Science Staff Week in Cadiz held on 21<sup>st</sup> and 22<sup>nd</sup> March 2023.

The hosting service for the Toolkit was also discussed between ORDOs. SubjectPlus, being an open source tool, was selected and recommended as a platform for hosting the content pertaining to the RDM Toolkit. UCA installed SubjectPlus and UBO, UM and UNIST were responsible for the uploading of the Toolkit content. (<https://subjectsplus.github.io/>)

### 3. DETAILED DESCRIPTION OF RESULTS

#### 3.1 SEA-EU Research Data Management Policy Framework

The SEA-EU Research Data Management Policy Framework ([Appendix III](#)) is made up of eight sections as follows:

1. **Preamble** which provides the context to the policy.
2. **Introduction** which specifies the purpose and scope of the document.
3. **Definitions** which describe the terminology used within the policy.
4. **Policy Guiding Principles** which constitute the main clauses of the policy.
5. **Roles and responsibilities** which have to be undertaken by the SEA-EU Universities and Researchers.
6. **Data Management Planning** which defines the importance of compiling a Data Management Plan.
7. **Research Data Management Policies in relation to existing Institutional Policies** which ensure that this overarching policy framework aligns with the universities' regulatory framework and existing guidelines pertaining to research.
8. **Support for the Implementation of Research Data Management Policies** which highlights the importance of collaboration between various entities within respective SEA-EU Universities.

The Policy's fundamental Guiding Principles comprise seven clauses which stipulate that Research Data should be:

1. accurate, complete, authentic and reliable
2. published on a discipline specific or an institutional data repository to ensure access and reuse. The notion of "*as open as possible and as closed as necessary*" should be applied

3. in line with the FAIR Data Principles so as to support findability, accessibility, interoperability and reusability
4. compliant with statutory, ethical and contractual requirements
5. documented by compiling a Data Management Plan
6. managed by the Principal Investigator; and
7. made available for consultation and reuse as quickly as possible, when feasible.

### 3.2 SEA-EU Research Data Management Toolkit

The SEA-EU Research Data Management Toolkit is made up of four main components:

1. **Introduction** which highlights the scope of the Toolkit and includes a definition of Research Data Management and FAIR Data Principles
2. **Policies** which consist of links to policies and infrastructures together with concise descriptions of the resources available within the SEA-EU Alliance at institutional and national level
3. **Research Data Lifecycle** which presents the process of conducting research and identifying the main phases and related data activities taking place during the data lifecycle.
4. **Data Management Plan** which consists of DMP tools that are freely available for researchers via online applications for the creation, management and sharing of DMPs.

This Toolkit is based on a five-phased data lifecycle model, this being:

1. **plan and document** - identification of data collected or used to answer the research question or hypothesis by creating a Data Management Plan. Researchers should also take into consideration ethics and research integrity during this phase.
2. **collect** - refers to the method of gathering, measuring and analysing data from various relevant sources. Important aspects of data collection include standardisation, structure and organisation of data, data quality, documentation and metadata.
3. **analyse and store** - research data can be categorised as quantitative or qualitative, depending on the methodology used during the data collection process. Data analysis refers to the process of manipulating raw data to determine useful insights and draw conclusions. Appropriate data storage is crucial for ensuring that data is securely stored, accessible and shared in a way that maximises its value.
4. **archive and publish** - archiving data implies that a copy of the data is kept in a secure location, predominantly either in general or thematic data repositories. The choice of the repository must comply with the FAIR Data Principles. Publishing of data refers to the process of making data publicly available whether

in restricted or open access. Also, decisions on what other researchers can do with data need to be determined by making use of licences.

5. **access and reuse** - access refers to the process of identifying how data can be retrieved. Digital Object Identifiers (DOIs) make data easier to find while data citation makes it easier for others to reference content. Data reuse implies using data for other purposes than originally collected for.

## 4. RECOMMENDATIONS AND FOLLOW-UP

To maximise the benefits of the SEA-EU Research Data Management (RDM) Policy Framework, it is being recommended that:

1. the Policy Framework is periodically reviewed to reflect developments pertaining to emergent Research Data Management practices
2. appropriate visibility is provided to enhance uptake and adherence
3. it serves as a basis to align and harmonise institutional RDM policies
4. adequate training and support is provided to all stakeholders involved in the research lifecycle
5. appropriate infrastructure is provided and sustained to enable researchers to adhere with the FAIR Data Principles

The scope of the SEA-EU Research Data Management Toolkit is to translate theory into practice. Subsequently, it is being recommended that the Toolkit is:

1. promoted through various channels both at institutional and SEA-EU level
2. regularly reviewed, maintained and monitored so as to ensure that its content remains relevant and that it has no broken links
3. provided with the necessary resources to ensure that hosting services are supported beyond December 2025 (when UCA's responsibility for supporting the Toolkit ends).

## APPENDIX I

### Opinion Survey on Research Data Management

#### Preamble

Open Science is a concept that is gaining momentum as scientific endeavours and timely technological solutions are needed to address growing challenges for sustainable development, climate related targets, and higher performance in applied research and innovation. Scientific knowledge and scholarly application need to be more transparent, following a culture of sharing, exchange and access such that results and methods can be replicated and validated, while codes and analysis can be passed on for additional application and development without unnecessary duplications and delays in further advancement. Open Science thus goes much beyond the peer-review and open publishing of scientific literature.



It entails a complete and full delivery of the underlying research data, scientific methods, analysis and computing programs on top of the theoretical aspects.

**To this effect, the six partner universities composing the European University of the Sea (SEA-EU) partner universities are preparing to implement Research Data Management Policies as a fundamental step to ensure that research data is organised in a harmonised fashion throughout the entire research lifecycle which supports archiving and sharing, where appropriate.** A baseline Policy Framework is necessary to provide common principles and guidelines to policymakers responsible for research management within the SEA-EU partner

Universities. This facilitates the process of how research data should be managed, preserved and disseminated in order to maximise the potential of the research output in support of core values and missions. **In practical terms, the Policy Framework aims to support SEA-EU institutions with the adoption and implementation of Research Data Management Policies that facilitate the appropriate curation and management of data, to secure its longevity and its potential to be shared**

This questionnaire is intended to pave the way towards the creation and successful implementation of a Research Data Management Policy at the *<insert the name of the University>*. **Aside from affirming the structure and the key elements of the planned policy, it also serves to provide a clear picture about the expectations and experiences within the academic community, and shed some light on the current practices employed in the area of managing and publishing of research data.**

#### **Instructions on how to compile this questionnaire**

This questionnaire is intended to solicit and obtain feedback from University academics on their awareness about open science principles, their readiness to share their own research data, and the extent to which they value accessing the data generated by others.

This questionnaire is kept simple so that it can be answered in less than 10 minutes. Mandatory questions are indicated by a red asterisk. The answers are intentionally graded over five levels of agreement/disagreement so that a statistical analysis can be achieved, but additional (non-mandatory) comments are highly encouraged by those who wish to stress important aspects.

#### **Consent form**

##### **About the survey:**

This questionnaire is intended to get feedback from University academics on their awareness about open science principles, their readiness to share their own research data, and the extent to which they value accessing the data generated by others.

##### **Participation:**

You are being asked to participate in this survey because you are an academic/researcher. Your participation in this survey is entirely voluntary.

##### **Benefits & Risks:**

You will not gain any material benefit in your participation in this survey. However, your responses will help us pave the way towards the creation and successful implementation of a Research Data Management Policy. There are no known foreseeable risks involved in participating in this survey.

### **Confidentiality:**

All information obtained in this survey is strictly confidential unless disclosure is required by the law. The survey is strictly anonymous; your name and contact information will not be revealed at any point.

By signing this consent form you are agreeing that you read and you fully understand the contents of this document and are willing to take part in this survey.

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

## Definitions

### Research Data Management

Research Data Management (RDM) is a term that describes the organisation, storage, documentation, preservation, and sharing of data collected and used in a research undertaking. It involves the everyday management of research data during the lifetime of a research undertaking (e.g., using consistent file-naming conventions which describe the type of data within the file, the initials of the Principal Investigator and date). It also involves a strategy for the collection, backup and storage of data, data documentation, and ethical and legal requirements related to data, data sharing, data archiving and data destruction.

### Data Management Plan

A Data Management Plan (DMP) is a plan that outlines how data is managed from the point of collection at the start of a research undertaking all the way through to its analysis and elaboration of results and how it will be used beyond the original research undertaking. Typically, a DMP will cover such areas as data types, formats and volumes of data collected, metadata, quality control, scientific integrity, specifics concerning access and information concerning publications (as may be applicable).

### Data Repository

A Data Repository (DR) is an online platform which collects, preserves and disseminates research data.

### FAIR Data Principles

The FAIR Data Principles are a set of guiding principles that intend to optimise the reusability of research data by improving their Findability, Accessibility, Interoperability, and Reuse.

### Respondent's profile

Your scientific discipline (please select the appropriate):

- ✓ Natural Sciences
- ✓ Social Sciences
- ✓ Applied Sciences
- ✓ Medical and Health Sciences
- ✓ Arts & Humanities

Your age group (please select the appropriate):

- ✓ <30
- ✓ 30 -40
- ✓ 41-50
- ✓ 51-60
- ✓ >60

Your position (please select the appropriate):

- ✓ Early Career Researcher
- ✓ Recognised Researcher
- ✓ Senior Researcher
- ✓ Leading Researcher

## SECTION A: General principles

\*Q1. Research data should be shared with others, as with publications, keeping full recognition of source and authorship.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

\*Q2. Research data should be organised and published on a designated discipline-specific or institutional Data Repository with an open data licence for consultation and reuse, according to agreed common protocols.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

\*Q3. Published research data should adhere to the FAIR Data Principles (i.e., data should be findable, accessible, interoperable and reusable).

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

\*Q4. Research funding programmes and initiatives should impose criteria in favour of open research data (unless the data is protected by a law).

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

\*Q5. There are certain data sets that should absolutely never be openly shared.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Give examples of research data that should have restricted access: \_\_\_\_\_

\*Q6. What would you define as research data? (Please select one or more options)

- Databases
- Notebooks or Lab books
- Photos
- Videos
- Algorithms
- Texts
- Surveys
- Maps
- Translations
- Samples
- Procedures
- Laws
- All of the above
- Other: \_\_\_\_\_

\*Q7. What types of data do you use in your research? (Click all that apply)

- Geographical
- Statistical
- Environmental
- Substance emissions
- Personal data (GDPR )
- Sensitive data (military/trade secrets...)
- National security
- Other: \_\_\_\_\_

\*Q8. Do you use published data in your research?

- Yes
- No

\*Q8.1. If Yes, what types and from which repositories?

\_\_\_\_\_

\*Q9. Could you name any specific laws/principles concerning your research data?

\_\_\_\_\_

\*Q10. Where do you store your data in the short term? (Click all that apply)

- Laptop hard drive
- External hard drive
- Laboratory server
- University server
- National or regional service
- International service (please give the name): \_\_\_\_\_
- Other: \_\_\_\_\_

\*Q11. Where do you store your data in the long term? (Click all that apply)

- Laptop hard drive
- External hard drive
- Laboratory server
- University server
- National or regional service
- International service (please give the name): \_\_\_\_\_
- Other: \_\_\_\_\_

\*Q12. Where do you back-up your data? (Click all that apply)

- Laptop hard drive
- External hard drive
- Laboratory server
- University server
- National or regional service
- International service (please give the name): \_\_\_\_\_
- Other: \_\_\_\_\_
- I don't back-up data
- I am not permitted to back-up the data for security reasons

\*Q13. How do you share data with project partners? (Click all that apply)

- Google Drive
- ResearchGate
- Institutional service
- Other: \_\_\_\_\_

## SECTION B: Roles and responsibilities of the stakeholders

### B1. The University

\*Q14. The University should provide advice and support to researchers on data management practices as well as associated issues, such as data protection, research integrity, research ethics and Intellectual Property Rights, and with the compilation of Data Management Plans.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

\*Q15. The University should provide an Institutional Data Repository (or equivalent) that collects, preserves and provides access (when possible) to research data.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

\*Q16. The submission of research data generated within the University to an Institutional Data Repository should be mandatory if such a service exists.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

\*Q17. The University should offer support and training to affiliated researchers on how to deposit and access research data uploaded on Data Repositories.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

AQ1. What additional roles and responsibilities do you feel the University should have? (Click all that apply)

- ✓ The University should help you find an appropriate subject repository in case it does not have its own platform.
- ✓ In addition to support, advice, and training on research data management and associated issues, as delineated in Q6 and Q9, the University should also offer courses on Open Science, Open Data, and others.
- ✓ The University should provide easily accessible resources (e. g. a dedicated webpage) that provide guidance on good data management practices.
- ✓ The University should actively promote good data management practices.
- ✓ The University should monitor compliance with the RDM policy and supporting processes.
- ✓ Other: \_\_\_\_\_

## B2. Principal Investigators

**The principal investigator is a researcher responsible for a research undertaking, of any size, conducted for, on behalf of, or in association with the University; on the University premises; or using University facilities.**

\*Q18. The principal investigators should be responsible for the proper handling and publication of the research data collected by them or their team.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

\*Q19. The principal investigators should ensure that a Data Management Plan is created before research undertakings commence.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

\*Q20. The principal investigators should determine if and when a Data Management Plan needs to be updated.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

AQ2. What additional roles and responsibilities do you feel the principal investigators should have, if any? (Click all that apply)

- ✓ The principal investigators should deposit the Data Management Plan into an Institutional Data Repository (or equivalent).
- ✓ The principal investigators should be aware of all suitable platforms for depositing research data and the Data Management Plan, in case the University does not offer its own service.
- ✓ The principal investigators should attend training and courses on research data management and associated issues and work with the University on refining these courses.
- ✓ Where a research project is conducted in collaboration with external research partners, the principal investigators should ensure that suitable agreements for the ownership, use, and preservation of research data are established and agreed in writing by the parties concerned before the project starts.
- ✓ Other: \_\_\_\_\_

### B3. Researchers



**A researcher is a member of staff of the University who has an appointment of employment and who performs research as defined in this document. It also includes postgraduate students registered at the University who are also undertaking research as part of their studies.**

\*Q21. The researcher should ensure that research data is accurate, complete, authentic and reliable.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

\*Q22. The researcher should ensure that research data is managed and stored with appropriate security, including protecting any confidential, personal and sensitive information contained within.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

\*Q23. The researcher should publish research data in an appropriate digital format, along with sufficient descriptive metadata, in order to ensure adherence to the FAIR Data Principles.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

AQ3. What additional roles and responsibilities regarding research data practice do you feel that researchers should have? (Click all that apply)

- ✓ The researcher should provide a statement describing how and on what terms any supporting research data may be accessed either in an associated research article or in the descriptive metadata.
- ✓ When depositing research data in an external repository (in the case the University does not have its own suitable platform), the researcher should opt for such a repository that equips the records with persistent identifiers (e.g. DOIs).
- ✓ The researcher should create a back-up of their research data upon completion of every stage of the research project
- ✓ The format the researcher uses to publish research data should be open-source and non-proprietary
- ✓ Other: \_\_\_\_\_

## SECTION C: Data Management Planning

\*Q24. A Data Management Plan should be created and maintained for every research undertaking.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

\*Q25. A Data Management Plan should address the creation, management, documentation, storage and sharing of research data, and the production of descriptive metadata in order to ensure adherence to the FAIR Data Principles.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

\*Q26. A Data Management Plan should specify where the research data is deposited after the research undertaking has been concluded.

- Strongly agree
- Agree
- Undecided
- Disagree
- Strongly disagree

Additional comments: \_\_\_\_\_

AQ4. Is there anything else that, in your opinion, a DMP should explicitly address? If yes, kindly specify.

\_\_\_\_\_

Thank you for your answers!

## APPENDIX II

# reSEArch-EU

## Opinion Survey on Research Data Management

### Survey Results and Analysis

Contents

Introduction

4

## Respondents' profile

Section A: General principles	6
Section B: Roles and responsibilities of stakeholders	16
<b>1. The University</b>	<b>16</b>
<b>2. Principal Investigators</b>	<b>18</b>
<b>3. Researchers</b>	<b>20</b>
Section C: Data Management Planning	22
Concluding Remarks	23

## List of Tables

Table 1: Q1	6
Table 2: Q2	7
Table 3: Q3	7
Table 4: Q4	8
Table 5: Q5	8
Table 6: Q6	9
Table 7: Q7	11
Table 8: Data types mentioned by researchers	12
Table 9: Q10	14
Table 10: Q11	15
Table 11: Q11	15
Table 12: Q13	16
Table 13: Q14	16
Table 14: Q15	17
Table 15: Q16	17
Table 16 AQ1	18
Table 17: Q17	19
Table 18: Q18	19
Table 19: Q19	19
Table 20: AQ2	20
Table 21: Q20	20
Table 22: Q21	21
Table 23: Q22	21
Table 24: AQ3	22
Table 25: Q23	22
Table 26: Q24	22
Table 27: Q25	23

## List of Figures

Figure 1: Respondents' profile	5
Figure 2: Q6 graphical representation	10
Figure 3: Q7 graphical representation	11

## Introduction

The aim of this opinion survey is to pave the way towards the successful implementation of the Research Data Management Policy Framework for the SEA-EU Alliance within the six partner institutions. This survey also serves to gather an insight about the expectations and experiences within our academic communities in the area of managing and publishing research data.

The survey was distributed amongst academics of the six participating institutions with a time window of five weeks, from the 3rd week of August till the end of September 2022.

The population studied for the survey analysis has been treated as representative of all 6 member universities within the SEA-EU Alliance. The breakdown of the number of responses received from each university is provided below:

- University of Malta – 41 responses
- University of Cádiz (Spain) – 27 responses
- Université de Bretagne Occidentale in Brest (France) – 10 responses
- University of Kiel – 89 responses
- University of Gdańsk – 128 responses
- University of Split – 42 responses
- Total number of valid responses - 337

Although the number of respondents varied from one University to the other, no university-specific variances were noted. To this effect, homogeneous responses were noted across all Universities. For this reason, and since the scope of this opinion survey was to gather an insight about the expectations and needs of our researchers, it was decided that the final analysis should be presented collectively. Thus, any conclusions and suggestions are applicable to all the alliance, rather than university-specific.

## Respondents' profile

The number of valid responses for the questionnaire was 337; this will be considered the total population,  $n$ . Based on the respondents' ages, the highest number of respondents fall within the age group of 41-50, with a total of 106 respondents, equivalent to a percentage of 31.5%, falling within this bracket. The lowest number of respondents are over 60 years of age, with a total of 32 respondents and a percentage of 9.5%, falling within this age range.

Based on discipline, the highest number of respondents come from the Natural Sciences with a total of 160 participants, equivalent to 47.48% of the total population. The lowest number of respondents based on discipline come from the Medical and Health Sciences, where 31 respondents participated, equivalent to 9.2%.

On grounds of appointment within the University, the highest number of respondents were Early Career Researchers, with a total of 126 respondents, equal to 37.4% of the total population. The lowest number of respondents based on position were Leading Researchers, amounting to 42 participants, which is equivalent to 12.5%.

A tabular representation of the respondents' profile is provided in [Figure 1](#) below:

	Applied Sciences	Arts & Humanities	Medical and Health Sciences	Natural Sciences	Social Sciences	Grand Total
<b>&lt;30</b>	2	3	6	23	10	44
Early Career Researcher	2	3	5	23	9	42
Recognised Researcher					1	1
Senior Researcher			1			1
<b>&gt;60</b>	3	8	3	10	8	32
Leading Researcher	1	1	2	5	1	10
Recognised Researcher		2	1	2	2	7
Senior Researcher	2	5		3	5	15
<b>30-40</b>	9	7	12	54	16	98
Early Career Researcher	6	6	6	38	11	67
Leading Researcher			1			1
Recognised Researcher	1	1	4	15	5	26
Senior Researcher	2		1	1		4
<b>41-50</b>	11	7	8	50	30	106
Early Career Researcher	1	3	2	3	5	14
Leading Researcher	2		2	8	2	14
Recognised Researcher	3	1	2	19	12	37
Senior Researcher	5	3	2	20	11	41
<b>51-60</b>	13	8	2	23	11	57
Early Career Researcher	1	1			1	3
Leading Researcher	6	3	1	6	1	17
Recognised Researcher	2			5	3	10
Senior Researcher	4	4	1	12	6	27
<b>Grand Total</b>	<b>38</b>	<b>33</b>	<b>31</b>	<b>160</b>	<b>75</b>	<b>337</b>

**Figure 1: Respondents' profile**

## Section A: General principles

Respondents were asked about their views on sharing research data in the same way as one would share research publications, keeping full recognition of source and authorship. Respondents were presented with a five-point Likert scale, 1 being “Strongly agree” and 5 being “Strongly disagree”. The Likert scale was used in the majority of the questions in this survey and the results will all be presented in the same way to ensure consistency.

Based on the statistics gathered, one can conclude that the majority of the population agrees with research data sharing, with over 80% choosing Options 1 and 2. Only 13% of the population disagrees with data sharing, whilst a minority of respondents is indifferent.

Scale	1	2	3	4	5
No. of respondents	202	68	23	14	30
Percentage	59.9	20.2	6.8	4.2	8.9

Table 1: Q1

A manual content analysis was carried out to categorise the comments proposed by 32 of the respondents with regards to the aforementioned question. Seven (7) out of the 32 respondents who brought forward their comments mentioned confidentiality and ethical issues, stating that whilst agreeing with data sharing, this should not be at the cost of exposing sensitive or confidential information, such as clinical data. Four (4) participants claimed that data should only be published when the Principle Investigator (PI) no longer needs the data to generate further research output. One (1) of these four (4) participants added that the extent to which data is shared is to be decided by the PI. Three (3) participants mentioned that data should be shared depending on the funding; specifically, if the research was financed by public funds, then it should be available to the public. Conversely, 2 researchers stated that research incurs financial resources, and because of this, they are not willing to share for free. Nine researchers emphasised that whether or not they agree that data should be shared, it highly depends on the type of research data, discipline and content. Two (2) respondents mentioned that data should be shared, cohering FAIR principles.

A few reasons from respondents who are not in favours of sharing data were also brought forward, with respondents mentioning the following concerns:

- Publishing data for failed experiments will lead to a large amount of “wrong” data which could also be detrimental to the researcher’s reputation
- Intellectual property issues vis-a-vis patentable research data
- Misuse
- Freeriding



Participants were asked whether they agree or disagree that research data should be organised and published on a designated discipline-specific or institutional Data Repository with an open data licence for consultation, according to agreed common protocols. The data gathered is represented in Table 2, below:

Scale	1	2	3	4	5
No. of respondents	157	81	50	24	25
Percentage	46.6	24.0	14.8	7.1	7.4

Table 2: Q2

Once again, the comments section for Q2 provided room for researchers to express their concerns, amongst which the following were pointed out:

- Organisation of data is time-consuming.
- Sharing of data could lead to researchers losing their appointment by virtue of plagiarism in a very competitive environment.
- Repositories should cater for various discipline-specific data, unless more than one repository is in place; out of the participants who had an opinion on this, 3 stated that they prefer a subject-specific repository, whilst 2 participants believe that an inclusive repository should be available.
- Four (4) participants seemed to lack knowledge about the subject, clearly stating that they are not sure whether this will affect their eligibility to get a patent; they are not sure what to answer as they believe that the data will be too exposed; they are not sure by what the survey means by “research data”.

A total of 217 participants, equivalent to 64%, chose the same value on the Likert scales for Q1 and Q2. This shows that it is likely that researchers who agree that research data should be shared with others, also agree that such data should be organised in a way that it can be reused according to agreed protocols. Some comments and concerns raised in Q2 differ from those raised in Q1.

With a total of 263 (78 %) respondents indicated their agreement that research data should adhere to the FAIR Data Principles, 37 (11 %) respondents being indifferent about this, and another 37 (11 %) respondents disagreed. One can conclude that the majority of the survey participants believe that research data should be findable, accessible, interoperable and reusable. The detailed responses received for this question are outlined in Table 3.

Scale	1	2	3	4	5
No. of respondents	199	64	37	13	24
Percentage	59.1	19.0	11.0	3.9	7.1

Table 3: Q3

The concerns put forward in the comments section of Q3 included the following:

- Time and human resources; some researchers pointed out that there should be a person dedicated to help with the data preparation.

- Depending on the nature of the data and the discipline.
- Reusability should still respect and acknowledge original authors.

An appreciable number of researchers replied favourably to the question of whether research funding programmes and initiatives should impose criteria in favour of open research data, unless this is protected by law (see Table 4).

Scale	1	2	3	4	5
No. of respondents	148	84	47	29	29
Percentage	32.9	17.5	31.2	10.1	8.3

Table 4: Q4

Comments on this question included the following:

- Policy should allow for an embargo to be placed on the data.
- This should be a university policy rather than imposed by funding bodies.
- Infrastructure needs to be made available.
- Policy should exclude data related to inventions.

Respondents were asked whether they agree that certain data sets should in no circumstance be openly shared. The majority of researchers agreed that some data sets should not be openly shared with a fair share of respondents being indifferent or uncertain about this. The responses are provided in Table 5.

Scale	1	2	3	4	5
No. of respondents	111	59	105	34	28
Percentage	32.9	17.5	31.	10.1	8.3

Table 5: Q5

Examples of research data that should have restricted access, as suggested by the survey respondents included the following:

- Private data including personal information, sensitive data, financial data, private messages, medical data and health records, political and religious opinions
- Security related data e.g. data related to country's safety e.g. military information
- Intellectually protected data
- Data which is subject to patent application
- Data prior publishing of research output or conclusion of investigations
- Any data, as decided by the author
- Biological weapon design
- Data with information/locations on critically endangered/vulnerable species; data that could put any person or animal at risk
- Data shared between institutions, that is not intended to be shared further
- Notebooks and lab books; some suggested these should be available on request and perhaps with supervised access
- Personal notes, certain photos and videos
- Maps
-

- Experimental protocols which could be used by inexperienced people and possibly cause harm
- Data that could be used against particular human groups
- Research experimental therapies at an early stage
- Copies of archival materials
- Data of critical infrastructure
- Data before anonymisation
- Data that can be a source of bias, data that can be misunderstood when given without context
- Nuclear physics, certain chemical procedures, information about viruses
- Datasets obtained from commercial entities
- Provisional data
- Dangerous data / wrong data e.g. wrongfully attained data
- Harmful content e.g. child pornography
- Data that have potential commercial value
- Personal data collected from minors without explicit consent from parent and child
- All data
- Data that could potentially be abused
- Data about sites that should be protected e.g. heritage sites
- People's personal narratives
- Data related to rare diseases, genomic and genetic data
- Data not specifically consented for sharing
- Qualitative interviews with vulnerable persons
- Curated data (raw data could be open)

Survey respondents were asked to tick the items that they would define as research data. Each item, as outlined in Table 6, could be ticked more than once. A graphical representation of the replies is provided in Figure 2.

Research data	D at a b a s e s	N o t o o k s	P h o t o s	V i d e o s	A l g o r i t h m s	T e x t s	S u r v e y s	M a p s	T r a n s l a t i o n s	S a m p l e s	P r o c e s s e s	L a w s	A l l o f t h e a b o v e
No. of respondents	257	89	190	160	167	176	180	159	86	173	177	60	82
Percentage	76.26	26.41	56.38	47.48	49.55	52.23	53.41	47.18	25.52	51.34	52.52	17.80	24.33

Table 6: Q6

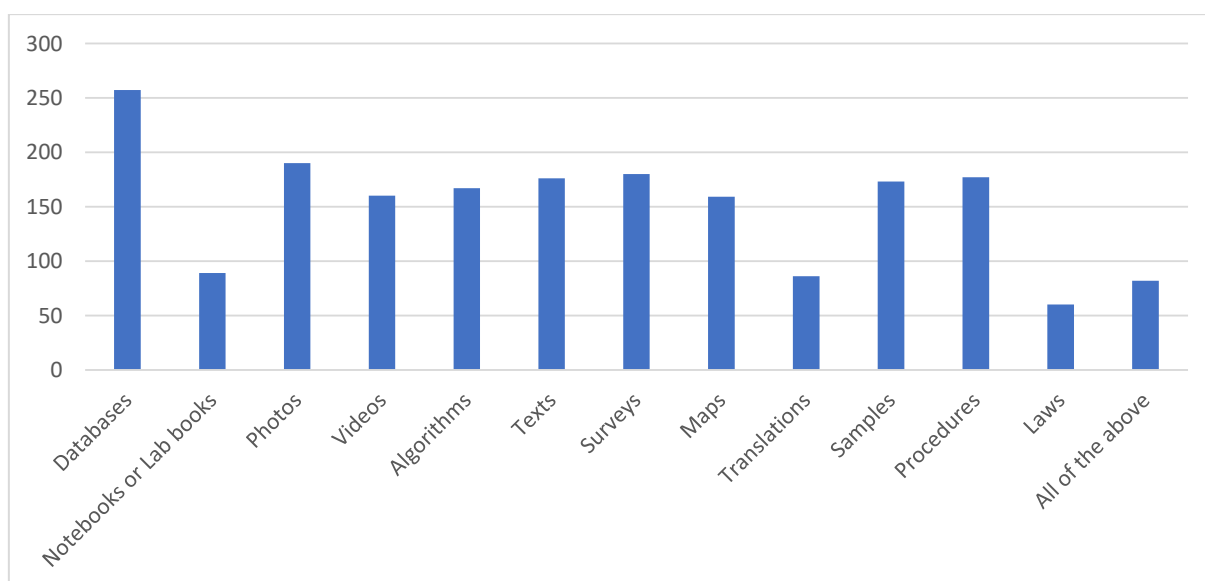


Figure 2: Q6 graphical representation

In addition to the above, a minority of participants added other items to the list that they would consider research data. These are listed below:

- Historical artefacts
- Personal narratives
- Data from lab experiments on which a publication is based, possibly also "failed" experiments
- Metadata
- Medical Histories, files and Medical Images or results
- People's narratives
- Books
- Calculating scripts
- Software /scripts
- Depends on the project. Anything generated during research which is not the object in question and not the result
- Experimental data
- Reports

Survey participants were asked to tick the types of research data that they use in their research. The data types listed below were included by the survey creators, nevertheless academics were free to include any other data types used in their research.

- Geographical
- Statistical
- Environmental
- Substance emissions
- Personal data (GDPR)
- Sensitive data (military/trade secrets...)
- National security

Although most of the respondents ticked one or more of the provided options, a number of participants added other data types to the list. Table 7 and Figure 3 display the number of participants who selected the given options, whilst Table 8 is a categorised representation of additional data types mentioned by the researchers.

Research data	Geographica l	Statistica l	Environmenta l	Substanc e emissions	Persona l data (GDPR)	Sensitive data (military/trad e secrets...)	National security
No. of respondents	103	246	146	27	84	19	8
Percentage of total no. of respondents	30.56	73.00	43.32	8.01	24.93	5.64	2.37

Table 7: Q7

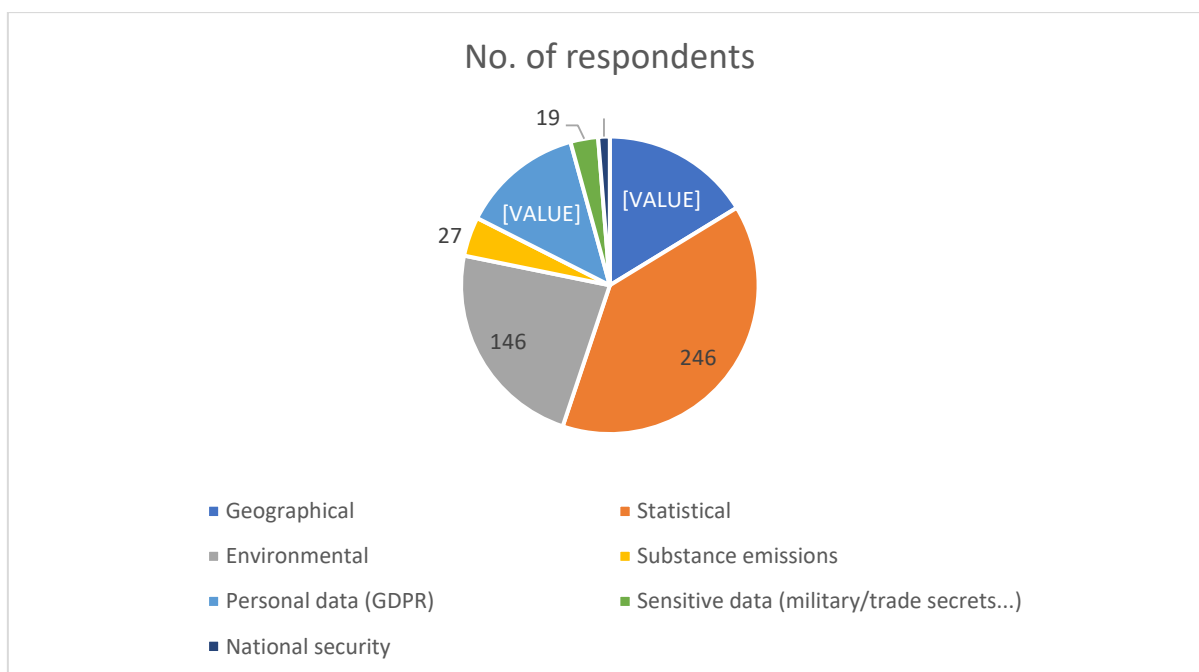


Figure 3: Q7 graphical representation

Data type:	No. of respondents
Historical / archival	5
Secondary	5
Astronomical observations	1
Scientific (laboratory, biological, chemical)	13
Experimental	9
Physiological	1
Voice recordings, photos, videos	7
Acoustic recordings	1
Archaeological	1
Personal	1
Medical	7

Simulated	2
Qualitative	5
Legal	2
Samples	2
Theoretical	1
Electoral	1
Algorithms / source codes	3
Technical	2
Descriptive	1

Table 8: Data types mentioned by researchers

When asked whether or not they use published data in their research, the majority of respondents answered affirmatively. In fact, 227 (67%) claimed they use published data in their research, whilst 110 (33%) respondents asserted that they do not use published data. Survey participants who currently use published data in their research were asked to specify the types of data they use as well as the repositories from which such data is retrieved. It transpires that, predominantly, researchers used subject-specific databases for their research. Such databases include:

- CIFAR: collections of images used in Artificial Intelligence
- GitHub and GitLab: software codes
- Copernicus, ESA and PANGEA: Earth observations and spatial data
- NCBI, bioRxiv, BIOTIC, OMIM, SRA: genetic and biological data
- BORIS: event logging software for video/audio coding and live observations
- Neotoma, ArkeoGIS: fossil, paleoecological and archeological data
- ORBIS: private company data
- Refinitiv LSEG Group, World Bank Data Portal: financial and economic data
- EMODnet: marine observations
- OECD, Eurostat and POIROT: statistical data

In addition to the above, a number of researchers claimed to be using government databases such as NSO, clinicaltrials.gov and the General Elections Register. Some respondents also indicated the use of publishing companies who supply datasets to support research outputs, such as 'Web of Science' by Clarivate, 'Scopus' by Springer and 'Data in Brief' by Elsevier. A number of participants use general purpose repositories such as Zenodo and Data Mendeley, whilst others claimed using National and University repositories.

Various laws and regulations were mentioned when researchers were asked to identify laws and principles concerning the research data in their fields. The General Data Protection Regulation (GDPR) was frequently mentioned, together with other National Laws and Regulations such as Copyright Laws, Heritage Laws and Animal Protection Laws. Other principles such as ethical standards, good clinical practice and several protocols were mentioned.

The following is an exhaustive list of the unique laws and principles identified through the survey.

- Copyright Law
- GDPR

## Heritage Laws

- Open data commons
- Bioethics
- Intellectual property laws
- Regulations on the protection of endangered species
- Patent laws & regulations
- Animal protection law
- Geological Data Act
- Genetic safety regulations
- DFG Code of Conduct
- GenTG (Genetic Engineering Act)
- laws concerning gene technology
- laws concerning biological safety
- Data Use Act
- Hospital Act
- Employment and Industrial Relations Act
- Product Safety Act
- Environment Protection Act
- Law for hydrographic measurements
- Law for the protection of sanitary data
- Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information
- Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes
- GNU General Public License
- Statistical confidentiality
- PRISMA 2020 principles
- CONSORT statement
- INSPIRE directive
- Code of relations
- Code of ethics
- Declaration of Geneva
- 3Rs framework
- MIT licence for code
- Creative Commons License
- Principles of portability, reproducibility, readability
- Good scientific practice
- Confidentiality agreements
- University collective agreement
- FAIR principles
- Principles of Good Laboratory Practice (GLP)
- Good clinical practice
- Publisher policies
- WHO recommendations
- International Commission on Zoological Nomenclature

## Declaration of Helsinki

- Nagasaki protocol
- Cartagena protocol
- Biosafety protocol
- Nagoya Protocol

When asked where researchers store data in the short term, it transpired that the majority use a laptop hard drive. This being said, most researchers use more than one storage medium. Results obtained for this question are recorded in Table 9.

Storage medium	Laptop hard drive	External hard drive	Laboratory server	University server	National or regional service	International service
No. of respondents	280	198	69	144	10	52
Percentage	83.09	58.75	20.47	42.73	2.97	15.43

Table 9: Q10

Some participants claimed using other storage media that were not included in the list of selectable options. This includes workstation hard-drive, pen drive and internal/private cloud. Another participant mentioned email as a storage medium, whilst 2 participants store their data in paper-based format.

The international data storage systems used by the survey participants include cloud-based services such as GDrive, Dropbox, OneDrive and Amazon AWS. The following international storage systems were also identified:

- Zenodo
- GitHub
- GitLab
- Allas csc
- arXiv
- PRIDE
- ProteomeXChange
- GenBank
- Copernicus
- Protein Data Bank
- Biological Magnetic Resonance Data Bank
- HAL archives
- Global Telecommunication System (GTS)
- Zotero
- NCBI
- Earthquake Source Model Database
- IPCC database
- AWS Storage Solutions by Amazon
- RedCap platform



Participants were asked about which media they use for backing-up data. The majority of respondents use external hard drives, followed by laptop hard drives and University

servers, respectively. Out of the participants who back their data on their laptop hard drives, there are 21 researchers who claimed that this is their only back up medium. Table 10 shows the breakdown of the replies for this question.

Storage medium	Laptop hard drive	External hard drive	Laboratory server	University server	National or regional service	International service	Do not back-up data	Not permitted to back-up data (security reasons)
No. of respondents	157	229	52	129	11	44	12	1
Percentage	46.6	67.9	15.4	38.3	3.3	13.1	3.6	0.3

Table 10: Q11

In addition to the above, participants also mentioned the following media that they use for data back-up:

- cloud services
- pendrive
- ELAB
- GitHub
- workstation hard drive
- private cloud
- email
- paper

When asked about international services for back-ups, the aforementioned services were brought up. The only distinctive answer indicated the use of Backblaze, an American cloud storage and data backup company based in San Mateo, California.

Survey results show that more than half the respondents use Google drive to share data with project partners, whilst over 40% use an institutional service.

Data sharing medium	Google Drive	ResearchGate	Institutional service
No. of respondents	174	41	140
Percentage	51.6	12.2	41.5

Table 11: Q11

Whilst researchers do use Google Drive, ResearchGate and Institutional service to share data with their peers, a number of alternative cloud-based options were indicated, namely:

- Next cloud
- GitHub
- GitLab
- Dropbox
- Box
- National clouds
- BonaRes
- REDCap

Aside from the above, one of the participants mentioned Overleaf, which is available as a cloud-based or on-premises solution. Flash drives, including pendrives and external hard drives are also being used by researchers to share data. Some researchers claimed using emails, Microsoft teams, Private websites and Academia.eu, whilst others favour FTP clients, and large file transfers solutions such as WeTransfer and Gros fichier. In addition to all the mentioned solutions, two researchers opt for sharing by handing the physical copies.

## Section B: Roles and responsibilities of stakeholders

### 1. The University

Survey results show that, predominantly, academic members of staff agree that the University should provide advice and support to researchers on data management practices and associated issues, such as data protection, research integrity, research ethics, Intellectual Property Rights, and with the compilation of Data Management Plans (see Table 12).

Scale	1	2	3	4	5
No. of respondents	227	60	21	11	18
Percentage	67	18	6	3	5

Table 12: Q13

The fact that there needs to be human resources available to do the organisation, management and uploading of research data was repeatedly highlighted in the comments. Rather than general information sessions, Universities need to have human resources that are experienced in the research field and then train these personnel in how to manage data in that field.

As can be observed in Table 13, most researchers agree that the University should provide an Institutional Data Repository (or equivalent.) The aim of this repository would be to collect, preserve and provide access (when possible) to research data.

Scale	1	2	3	4	5
No. of respondents	<b>185</b>	<b>72</b>	<b>39</b>	<b>20</b>	<b>21</b>
Percentage	55	21	12	6	6

Table 13: Q14

Time constraints were mentioned in the comments associated with Q14. A number of researchers seem to insist on having subject-specific repositories and others maintain that publishing data in a repository would comprise the possibility of publishing with reputable Journals. It was also suggested that a National repository would be better than a University repository. Some researchers do not see the use of having an institutional repository when there are other international repositories available.

Survey participants were asked about their stand on the compulsory submission of research data generated within the University on an Institutional Data Repository if the service exists. The results obtained in this question are shown in Table 14.

Scale	1	2	3	4	5
No. of respondents	<b>78</b>	<b>80</b>	<b>86</b>	<b>44</b>	<b>49</b>
Percentage	23	24	26	13	15

Table 14: Q15

Researchers' opinion on mandating open research data is quite evenly distributed, even though the majority do agree, one can notice that the predominance is not so clearly identified.

When commenting on the above, 9 researchers mentioned that they agree with mandating open research data, but with certain criteria and conditions, namely the mandate should consider the type of data being deposited (not all types of data should be deposited), and perhaps certain data should be in restricted access. Four (4) respondents insisted that sharing of data should be at the researchers' discretion. Three (3) researchers mentioned the possibility of having duplicate data, therefore it is suggested that data is to be deposited on the University's repository only if it is not submitted elsewhere. Another 3 researchers commented that such mandate should only be in place for research funded by the University. Two (2) survey participants mentioned that the University should consider allowing an embargo period on the research data.

The greater part of the survey respondents agree that the University should offer support and training to affiliated researchers on how to deposit and access research data uploaded on Data Repositories, remarking that high quality training should be offered but not necessarily to academics. Some academics persistently ask that there

should be employees who do the work that they consider administrative. Others are against training, stating that they do not have the time for it. There are a few who assert

that they do not need such training and that the repository should be user-friendly, thus not requiring any training.

Scale	1	2	3	4	5
No. of respondents	<b>200</b>	<b>78</b>	<b>25</b>	<b>17</b>	<b>17</b>
Percentage	59	23	7	5	5

Table 15: Q16

Table 16 outlines additional roles and responsibilities as identified by the survey creators, together with the number of respondents that feel that each particular function is to be fulfilled by the University.

Roles and Responsibilities of the University	No. of respondents	Percentage
The University should help you find an appropriate subject repository in case it does not have its own platform.	227	67
In addition to support, advice, and training on research data management and associated issues, the University should also offer workshops on Open Science, OpenData, and others.	223	66
The University should provide easily accessible resources (e.g. a dedicated webpage) that provide guidance on good data management practices.	257	76
The University should actively promote good data management practices.	255	75
The University should monitor compliance with the RDM policy and supporting processes.	147	43

None.

9

2

Table 16 AQ1

For the most part, all the options were marked, barring the proposal for the University to monitor compliance. Comments put forward regarding this question include that the University should strongly support researchers to create the Data Management Plan (DMP), accentuating once again that HR should be supplied. It was also suggested that the University should have a consultation service that can assist researchers with queries on data protection and provide legal advice.

## 2. Principal Investigators

Although a total of 69% are favourable, even if on different levels, that the Principal Investigators (PIs) should be responsible for proper handling and publication of the research data collected by them or their team, there is a significant number of researchers who are against this notion, as can be observed in Table 17 below.

Scale	1	2	3	4	5
No. of respondents	<b>128</b>	<b>103</b>	<b>63</b>	<b>24</b>	<b>19</b>
Percentage	38	31	19	7	6

Table 17: Q17

A total of 13 respondents asserted that the responsibility should not solely fall on the PI, but on all parties involved in the research process. Once again, the idea of the University having a dedicated department or dedicated staff members to assume responsibility for data handling and publication, has been raised by various survey participants.

205 out of the 337 survey participants agree to different extents that the PI should ensure that a DMP is created before research undertakings commence.

Scale	1	2	3	4	5
No. of respondents	<b>106</b>	<b>99</b>	<b>81</b>	<b>34</b>	<b>17</b>
Percentage	31	29	24	10	5

Table 18: Q18

Some researchers argued that the PI does not always know what data will be gathered, as this might change over time. Others claimed that DMP could be updated or open enough to cater for “accidentally” discovered data. Some participants have put forward the idea that the University provides a template for the DMP. Avoiding bureaucracy and increasing efficiency were emphasised.

The general feedback obtained through the survey supports the idea of the PI as the person responsible to determine if and when a DMP needs to be updated. A good number of researchers have no opinion on this, whilst others emphasised that this should be done collectively between the PI, the university administration and the researching team. Other entities mentioned as possibly being responsible for updating the DMP were the data steward, a dedicated person or department at university and the funding body. Some researchers are set against this, claiming that the PI does not have the time or expertise to do this.

Scale	1	2	3	4	5
No. of respondents	<b>104</b>	<b>113</b>	<b>77</b>	<b>24</b>	<b>19</b>
Percentage	31	34	23	7	6

Table 19: Q19

With only 9% of respondents claiming that the PI should not be responsible for any of the tasks listed by the survey creators (see Table 20), most researchers do agree that PIs should be responsible for certain tasks related to the DMP. The majority of respondents consider written agreements for ownership, use and preservation of research data with external research partners as top of the list of responsibilities for the PI, followed by attending courses on RDM and associated issues and working with the University to refine such courses.

Additional Roles and Responsibilities of the Principal Investigator	No. of respondents	Percentage
The principal investigators should deposit the Data Management Plan into an Institutional Data Repository (or equivalent).	119	35
The principal investigators should be aware of all suitable platforms for depositing research data and the Data Management Plan, in case the University does not offer its own service.	149	44
The principal investigators should attend training and courses on research data management and associated issues and work with the University on refining these courses.	179	53
Where a research project is conducted in collaboration with external research partners, the principal investigators should ensure that suitable agreements for the ownership, use, and preservation of research data are established and agreed in writing by the parties concerned before the project starts.	204	61

None.

31

9

Table 20: AQ2

### 3. Researchers

With 90% of the population agreeing that the researcher should ensure that research data is accurate, complete, authentic and reliable, it can be concluded that the vast majority are consenting with this statement. One needs to bear in mind that, at times, the PI and the researcher are the same person, as outlined in one of the comments. Whilst some participants commented on research integrity, thus affirming their responsibility for ensuring good standards in their data, others used the comments section to reiterate their opinion on having administrative human resources to do this. One comment also mentioned that Universities should have an 'Office of Research Integrity' that will be responsible for ensuring data standards and investigating any complaints.

Scale	1	2	3	4	5
No. of respondents	<b>265</b>	<b>36</b>	<b>12</b>	<b>6</b>	<b>18</b>
Percentage	79	11	4	2	5

Table 21: Q20

As can be observed in Table 22, most academics believe that the researcher should ensure that research data is managed and stored with appropriate security whilst protecting confidentiality. This being said, it is suggested that such responsibility should be assumed by the whole team, rather than just the researcher, possibly with guidance by IT services who have expertise in the system.

Scale	1	2	3	4	5
No. of respondents	<b>203</b>	<b>64</b>	<b>33</b>	<b>16</b>	<b>21</b>
Percentage	60	19	10	5	6

Table 22: Q21

The general feedback collected from Q21 (Table 22) and Q22 (Table 23) is that whilst researchers do agree that their research data should be organised in a specific manner, some of them specified that this should only serve the researcher until the research output is completed; not everyone agrees with the concept of reusability.

Scale	1	2	3	4	5
-------	---	---	---	---	---

No. of respondents	<b>163</b>	<b>84</b>	<b>51</b>	<b>17</b>	<b>22</b>
Percentage	48	25	15	5	7

Table 23: Q22

Table 24 shows the survey respondents' opinions on which additional roles and responsibilities should be assumed by researchers. The comments following this question were effectively duplicated comments which were reported with previous questions.

Additional Roles and Responsibilities of Researcher	No. of respondents	Percentage
The researcher should provide a statement describing how and on what terms any supporting research data may be accessed either in an associated research article or in the descriptive metadata.	198	59
When depositing research data in an external repository (in the case the University does not have its own suitable platform), the researcher should opt for such a repository that equips the records with persistent identifiers (e.g. DOIs).	177	53
The researcher should create a back-up of their research data upon completion of every stage of the research project.	236	70
The format the researcher uses to publish research data should be open-source and non-proprietary.	147	44
None.	22	7

Table 24: AQ3

## Section C: Data Management Planning

Respondents were asked whether a DMP should be created and maintained for every research undertaking. As can be observed in Table 25, less than 20% of respondents disagree with this, whilst 28% are indifferent. Over half the population agrees with having a DMP and updating it. Whilst appreciating the fact that the majority agrees, one has to point out that there seems to be a percentage of academics who do not see the need and importance of having a DMP.

Scale	1	2	3	4	5
No. of respondents	<b>90</b>	<b>89</b>	<b>93</b>	<b>44</b>	<b>21</b>
Percentage	27	26	28	13	6

Table 25: Q23



Constructive feedback received by respondents included the suggestion to avoid unnecessary bureaucracy, to make sure that the policy differentiates between areas of study, nature of the research, scope, duration and depth. Thus, it should not be a one-size fits all policy, but different procedures need to be in place depending on the project in question. It is also important to consider research done with no financial support; asking for a DMP in such cases may add to the financial burden incurred by the researcher.

Participants were asked whether they agreed with the idea that a DMP should address the creation, management, documentation, storage and sharing of research data, and the production of descriptive metadata in order to ensure adherence to the FAIR Data Principles. As can be observed in Table 26, only 10% of the population disagrees with this.

Scale	1	2	3	4	5
No. of respondents	<b>118</b>	<b>112</b>	<b>76</b>	<b>19</b>	<b>12</b>
Percentage	35	33	23	6	4

Table 26: Q24

The majority of the population also agrees that the DMP should specify where the research data is deposited once the research is concluded (see Table 27).

Scale	1	2	3	4	5
No. of respondents	<b>130</b>	<b>107</b>	<b>64</b>	<b>21</b>	<b>15</b>
Percentage	39	32	19	6	4

Table 27: Q25

Feedback received on this question included the idea of allowing researchers to choose a data repository after the data is collected, or at least allowing them to change the chosen repository. The argument here is that since research is a dynamic process, the kind of data gathered may not be straightforward, thus the right kind of repository should be selected. This argument is not considering the possibility of having a dynamic university repository which caters for all types of data.

Asked whether there is anything else that should be addressed in the DMP, participants brought forward the following salient points:

- The DMP should be brief
- It should specify who is responsible for what
- Access levels should be outlined
- It should clearly state who owns the data
- Academics could be involved in creating guidelines for DMPs

## Concluding Remarks

The results and comments obtained through the survey have shown that on one hand there are a number of academics who are in favour of open research data and of

having a DMP. Some have expressed willingness to learn more about the subject in order to be able to better handle their data. On the other hand, there are a number of academics who are completely against sharing their data and feel burdened by the administrative work incurred for preparing and maintaining DMPs. The latter have used the survey to communicate negative remarks by stating that the University is asking for too much administrative work from researchers. Although they may be few in number, such cohorts of academic may hinder the implementation process for policies and procedures concerning research data. As a result, more awareness and targeted communication are needed. Awareness and education on open research data should include legal training, as well as training on how to handle data and how to anonymise sensitive data.

Additionally, not all researchers understand what is meant by raw data and which data should be shared at what stage in the research process.

Universities who will implement a Research Data Management policy may consider appointing a team of people who will be responsible for assisting researchers in their data management. Ideally the team would comprise various entities within universities, including contributions coming from librarians, IT support, experts in data management, and legal support, who could offer guidance on issues regarding data privacy and sensitivity. Additionally, the team should have IT representatives who will assist with the technical aspect of data handling.

When choosing a data repository, one should ensure ease of use and compatibility with a wide range of data types / data formats in order to cater for all subject areas. Clear concise guidelines need to be provided, including which types of data should be included, how to anonymise and whether data should be curated or raw (adhering to FAIR data principles). Researchers should be aware of the FAIR data principles and of what is meant by data or research data; a few survey participants stated that they do not know what is meant by research data.

Feedback received through the questionnaire has also suggested that the University should provide a template for the DMP so as to minimise the extra burden on the PI and researchers. In conclusion, it is believed that researchers may be more willing to collaborate if given the necessary assistance and if the responsibility and the creation of the DMP is shared between the researcher, PI, librarians and IT experts.

## APPENDIX III

# SEA-EU Research Data Management Policy Framework

Title of the document	SEA-EU Research Data Management Policy Framework
Document type	Policy Guidelines
Version of the document	1.0
Approved by	SEA-EU Alliance
Approval date	



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017454

Last review date	
------------------	--

## List of Abbreviations

DMP	Data Management Plan
DR	Data Repository
RDM	Research Data Management
SEA-EU	European University of the Seas (A 'European University' Alliance)

## 1. Preamble

The rapid development of information technology has fundamentally altered the parameters of academic research. Research is enabled due to the possibilities for collecting, processing, analysing and exchanging large quantities of data. In the wake of this, foundations were laid for an Open Science culture reflected by the establishment of principles for the open access publishing of research data in a FAIR (findable, accessible, interoperable, reusable) manner. To achieve this, awareness has to be raised amongst the scientific community, accompanied by the implementation of professional research data management services and sustainable information infrastructures.

## 2. Introduction

### 2.1 Purpose

The SEA-EU Alliance recognises research data as a valuable asset, pivotal for academic research and its contribution to society. To this effect, the implementation of Research Data Management Policies within the SEA-EU partner universities is fundamental to ensure that research data is organised in a harmonised fashion throughout the entire research lifecycle which supports the protection, archiving and sharing of data, as and where appropriate.

The purpose of this Policy Framework is to provide common principles and guidelines to policymakers responsible for research management within the SEA-EU partner universities. This facilitates the process of how research data should be managed, preserved and disseminated in order to maximise the potential of the research output in support of core values and missions. In practical terms, this Policy Framework aims to support SEA-EU partner universities with the adoption and implementation of institutional Research Data Management Policies that: facilitate the appropriate curation and management of data; secure its longevity; and support its potential to be shared and re-used.

### 2.2 Scope of the document

This Policy Framework serves as a guideline that is expected to be adopted by the SEA-EU Alliance for the implementation of research data management policies. These policies will facilitate dissemination, visibility and impact of research data generated by the SEA-EU Alliance. The policy builds upon the different national policies applicable to each SEA-EU partner university and any relevant European policies already in place.

### 3. Definitions

#### 3.1 Data Management Plan

The Data Management Plan (DMP) is a plan that outlines how data is managed from the point of collection at the start of a research undertaking, all the way through to its analysis and elaboration of results and how it will be used beyond the original research undertaking. Typically, a DMP will cover areas such as data types, formats and volumes of data collected, metadata, quality control, scientific integrity, specifics concerning access and information concerning publications (as may be applicable).

#### 3.2 Metadata

Metadata refers to additional informative data that explains and describes the characteristics, context and provenance of a dataset. To allow for findability, traceability and (re)usability, the metadata should follow the FAIR Data Principles and be generous and extensive<sup>1</sup>. This should at least include: the name of the dataset's creator(s)/contributor(s); its name or title; its date of collection or generation; its date of publication; a unique and persistent identifier; a description of what the data contains, an explanation of how the data has been created, collected or generated; a description of how it has been analysed, as well as details of any licensing information, where applicable. This metadata provides other researchers with the information needed to understand, reuse and build further on the data, as well as making the data more retrievable.

#### 3.3 Principal Investigator

A Principal Investigator is a researcher responsible for a research undertaking, of any size, conducted for, on behalf of, or in association with the university; on university premises; or using university facilities.

#### 3.4 Research

Research is the creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of humankind, culture and society, and the use of this stock of knowledge to devise new applications.<sup>2</sup>

#### 3.5 Research Data

Research data refers to the evidence that underpins the answer(s) to research question(s) and hypothesis testing, and validates findings and reproducibility regardless of its form (e.g., print, digital or physical). These might be quantitative measurements and information, or qualitative

---

<sup>1</sup> FAIR Data Principle F2: Data are described with rich metadata <https://www.go-fair.org/fair-principles/>

<sup>2</sup> OECD (2015), Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development. <http://oe.cd/frascati>

statements collected by researchers in the course of their work by experimentation, observation, modelling, interviews or other data-collection methods, or information derived from existing evidence. Data may be: raw or primary (e.g., direct from creation, measurement or collection); derived from primary data for subsequent analysis or interpretation (e.g., following quality checks, gap filling or as an extract from a larger data set); or derived from existing sources where the rights may be held by others. Data may be defined as a 'relational' or 'functional' component of research, thus signalling that its identification and value lies in whether and how researchers use it as evidence for claims.<sup>3</sup> Some examples of types of research data include measurements, videos, surveys, interviews, photos, samples, transcriptions, recordings, translations, models, algorithms, protocols and standards.

### 3.6 Research Data Management

Research Data Management (RDM) is a term that describes the organisation, storage, documentation, preservation, and sharing of data collected and used in a research undertaking. It involves the everyday management of research data during the lifetime of a research undertaking (e.g., using consistent file-naming conventions which describe the type of data within the file, the initials of the Principal Investigator and date). It also addresses collection strategies, backup and storage of data, data documentation, and ethical and legal requirements related to data, data protection, data sharing, data archiving and data destruction.

### 3.7 Research Undertaking

Any type of research undertaken, supported or conducted by or within the University.

### 3.8 Researcher

A researcher is a member of staff of the University who has an appointment of employment and who performs research as defined in this document. It includes students or researchers registered at the University who are undertaking research as part of their studies.

## 4. Policy Guiding Principles

This Policy Framework sets out the fundamental guiding principles which SEA-EU partner universities are encouraged to adhere to and promote while compiling their institutional Research Data Management Policies. These guiding principles include:

- 4.1 Quality:** It should be ensured that research data is accurate, complete, authentic, and reliable.

---

<sup>3</sup> Higher Education Funding Council for England, Research Councils UK, Universities UK and Wellcome (2016), UK Concordat on Open Data <https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-ConcordatonOpenResearchData.pdf>



- 4.2** Open Data: Research Data should be published on a designated discipline-specific or institutional Data Repository<sup>4</sup> for consultation and reuse by external parties, as quickly as possible with an open data licence<sup>5</sup>. Data access should be as open as possible and as closed as necessary.<sup>6</sup>
- 4.3** FAIR Data Principles: Published research data should adhere to the FAIR Data Principles (i.e., data should be findable, accessible, interoperable and reusable).
- 4.4** Compliance: Rules and guidelines on complying with statutory, ethical and contractual requirements relating to research data should be defined. Necessary measures should include, amongst others, respect for privacy and confidentiality, copyright provisions, and the safeguarding of Intellectual Property Rights.
- 4.5** Planning approach: A Data Management Plan should be created and maintained for every research undertaking.
- 4.6** Responsibility: The principal investigator should be responsible for the proper handling and publication of the research data.
- 4.7** Availability: When feasible, research data should be made available for consultation and reuse as quickly as possible. Published research outputs should include a statement on how to access and use any supporting research data, and in the case of embargos, the timelines when this data becomes available, if and when applicable.

## 5. Roles and Responsibilities

With the objective of facilitating Research Data Management, there are a number of principles that researchers are encouraged to embrace and follow at the level of the SEA-EU Alliance. On their part, the Universities endeavour to incentivise Research Data Management practices within their research community by providing advice, facilities and support to enable and enhance research data exchange in a manner consistent with international conventions, where applicable.

---

<sup>4</sup> To facilitate interoperability of repositories and to harmonise research data, it is recommended that the Data Repositories of the SEA-EU partner universities should be OpenAire compliant.

<https://guidelines.openaire.eu/en/latest/>

<sup>5</sup> To explicitly allow for using, reusing and redistributing of the research data, the licence should conform to the Open Definition of the Open Knowledge Foundation. <https://okfn.org/>

<sup>6</sup> Balancing “as open as possible” and “as closed as necessary”

<https://www.cesaer.org/news/balancing-as-open-as-possible-and-as-closed-as-necessary-758/>

## 5.1 The University

- The University is responsible for disseminating information amongst its staff on the requirements of its Research Data Management Policy. Faculties and departments should be proactive in disseminating these requirements within their respective academic communities, as well as, encouraging and facilitating compliance.
- The University should provide advice and support to researchers on data management practices and the compilation of DMPs.
- The University should provide advice and support to researchers on associated issues, such as data protection, research integrity, research ethics, FAIR data principles and Intellectual Property Rights.
- The University should provide an institutional Data Repository (DR) that collects, preserves and provides access to research data. Access to data should be managed by the University.
- The University should offer support and training to researchers on how to deposit and access research data uploaded on the DR.

## 5.2 Researchers

### 5.2.1 Principal Investigators

- Principal Investigators hold day-to-day responsibility for the effective management of research data generated within or obtained from their research, including their research groups and research undertakings. This should include understanding and complying with the requirements of any relevant contract or grant agreement with the University that includes provisions regarding the ownership, preservation and dissemination of research data.
- Principal Investigators are to ensure that a DMP is written before research undertakings commence.
- Principal investigators are to determine if and when a DMP needs to be updated.

### 5.2.2 Researchers

- Following a professional approach, researchers should make every reasonable effort to keep an accurate and comprehensive record of their research, including documentation of clear procedures for the collection, storage, use, reuse, access and retention or deletion of the research data associated with their undertaking. Where appropriate, this

approach should also include defining and documenting protocols and responsibilities in collaborative research undertakings.

- Researchers should ensure that research data is managed and stored with appropriate security. This includes protecting confidential, personal and sensitive research data in accordance with legal and ethical requirements related to the research they conduct.
- Researchers must ensure that they abide by licences or terms of use when using or sharing third party data.
- Researchers are encouraged to publish or deposit data in an appropriate digital format (i.e. in a non-proprietary format) that is suitable for long-term retention, along with sufficient descriptive metadata on the DR, in order to facilitate data findability and re-use.
- Researchers should ensure that published research outputs include a statement on how to access and use any supporting data.
- Researchers should ensure that the research data is published under an open data licence that conforms to the Open Definition (See Clause 4.2).
- Where researchers supervise students, postdocs or other research staff, they should be aware of supervisor responsibilities with regards to ensuring that data is being managed in conformity with the DMP.
- Students, postdocs, researchers and their supervisors should ensure that data management is planned and documented at the outset of the research undertaking.

## 6. Data Management Planning

- 6.1. A Data Management Plan (DMP) should be developed at the outset of any research undertaking. The DMP should form the basis of data management throughout the various stages of the research lifecycle. Relevant support services within each SEA-EU partner university should advise whether a research funder requires a DMP to be included in the grant application.
- 6.2. DMPs are intended to address the creation, management, documentation, storage, protection and sharing of research data, and the production of descriptive metadata to aid discovery and re-use.
- 6.3. DMPs evolve with research undertakings and thus may require updating throughout the duration of a research undertaking. It is the responsibility of the Principal Investigator to determine if and when a DMP needs to be updated. Moreover, the DMP should specify where the data is to be deposited after the research undertaking has been concluded, and any conditions that may apply.

## **7. RDM Policies in Relation with existing Institutional Policies**

While complying with this overarching SEA-EU Policy Framework, the specific Research Data Management Policies implemented by the individual SEA-EU partner universities should fall within the specific university's regulatory framework and be interpreted in conjunction with any other existing institutional policies and guidelines pertaining to research (e.g., Open Access Policies, Intellectual Property Policies, Privacy Policies, Copyright Provisions, Research Code of Practice, Ethics and Integrity).

## **8. Support for the Implementation of Research Data Management Policies**

The implementation of Research Data Management policies involves the collaboration of various entities within the respective partner universities of the SEA-EU Alliance (e.g. Library Services, Computer Centre Services, Legal Services, Corporate Research & Knowledge Transfer, Project Support Office). Consequently, it is highly recommended that these entities work closely together to provide the necessary support and advice on the diverse aspects of research data management principles and practice.

14.06.2023