This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101136829

# D7.5 Data Management Plan and Ethics v1

| | | | |
|---|---|---|---|
| **Deliverable No.** | D7.5 | **Due Date** | 30.06.24 |
| **Description** | This is the first version of the Data Management and Ethics Plan of the SUNRISE project. | | |
| **Type** | Report | **Dissemination Level** | PU |
| **Work Package No.** | WP 7 | **Work Package Title** | Project Management |
| **Version** | 1.0 | **Status** | Completed |

# Authors


| Name and Surname | Partner Name | E-mail |
|---|---|---|
| Lucas Javier Segal | PBY | lsegal@predictby.com |
| Nick Dietrich | PBY | ndietrich@predictby.com |
| Justine Fleur Van der Feen | PBY | jvanderfeen@predictby.com |


# History

| Version | Date | Change History | Organisation |
|---|---|---|---|
| 0.1 | 26/03/24 | Creation of first draft document. | PBY |
| 0.2 | 26/04/24 | Input and feedback by Sofia Segkouli (CERTH) incorporated.<br>Major structure changes and addition of Sections and Appendices. | PBY |
| 1.0 | 20/06/24 | Input and feedback by internal reviewers incorporated, adaptation to SUNRISE project's template. | PBY |

# Internal Review History

| Name | Institution | Date |
|---|---|---|
| Sofia Segkouli | CERTH | 21/05/24 |
| Kleio Koutra | HMU | 23/05/24 |
| Haridimos Kondylakis | HMU | 23/05/24 |

# Abstract

This public report is the first iteration of Deliverable 7.5 “Data Management Plan and Ethics” (DMP), which is the first deliverable of Task 7.3 “Ethics and Data Management” in the context of the SUNRISE project. The main objective of the current DMP is to define SUNRISE’s project data governance and ethics policy. It outlines the nature of the data to be handled in the context of the activities and the methodologies by which that data is to be collected, processed, stored, shared and protected, during and after the end of the project. In addition, it presents the project's understanding of the ethical and legal obligations concerning data management and consent acquisition in research activities and provides consortium partners with guidance tools to comply with privacy and security requirements in the management of the research activities and its related data. This DMP is a “living” document, which means it will be continuously updated with input from the consortium partners based on the Data Management questionnaire and the Ethics Management questionnaire (Appendix A and Appendix B), following the development of the project’s activities and any updates in data management requirements.

# Legal Notice

# Table of contents

# List of tables

# List of figures

# List of Terms, Acronyms and Definitions

Table 1 – List of Terms, Acronyms and Definitions

| Term / Acronym | Definition |
|---|---|
| **Access rights** | Rights to use results or background (GA, Annex 5, Art 16) |
| **AI Act** | Regulation (EU) laying down harmonised rules on artificial intelligence (approved 13/03/2024, pending publication and numeration as of 15 June 2024) |
| **APIs** | Application Programming Interfaces |
| **CDM** | SUNRISE's Communication & Dissemination Manager |
| **CERN** | European Organisation for Nuclear Research |
| **Consent** | Any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her [GDPR, Art 4(11)] |
| **CSV** | Comma-separated values |
| **Data** | Any digital representation of acts, facts or information and any compilation of such acts, facts or information, including in the form of sound, visual or audiovisual recording [Data Governance Act, Art. 2(1); Data Act, Art. 2(1)] |
| **Data Act** | Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act) |
| **Data concerning health/ Health data** | Personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status [GDPR, Art 4(15)] |
| **Data controller** | The natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data [GDPR, Art 4(7)] |
| **Data Governance Act** | Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act) |

| | |
|---|---|
| **Data minimisation** | Personal data which shall be processed to be "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed" [GDPR, Art 5 (1)(c)] |
| **Data processing** | Any operation or set of operations which is performed on personal data or on sets of personal data, or on non-personal data or on sets of non-personal data in electronic format, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction [GDPR, Art 4(2) and FFDR, Art 3(2)] |
| **Data processor** | The natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller [GDPR, Art 4(8)] |
| **Data security** | Taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons, the controller and the processor shall implement appropriate technical and organizational measures to ensure a level of security appropriate to the risk. |
| **Data subject** | Any identified or identifiable natural person 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 [GPDR, Art. 4(1)] |
| **DMP** | Data Management Plan |
| **DOAJ** | Directory of Open Access Journals |
| **DOI** | Digital Object Identifier |
| **DPIA** | Data Protection Impact Assessment |
| **EDPB** | European Data Protection Board |
| **EDPS** | European Data Protection Supervisor |
| **ENISA** | European Agency for Cybersecurity |
| **EOSC** | European Open Science Cloud |
| **EU** | European Union |
| **FAIR** | Findable, Accessible, Interoperable and Reusable data |

| | |
|---|---|
| **FFDR** | Regulation (EU) 2018/1807 on a framework for the free flow of non-personal data in the European Union |
| **GA** | Grant Agreement |
| **GDPR** | Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) |
| **IPR** | Intellectual Property Rights |
| **JSON** | JavaScript Object Notation |
| **Non-personal data** | 'Non-personal data' means data other than personal data |
| **Open access** | Online access to research outputs provided free of charge to the end-user (GA, Annex 5, Art 16) |
| **Open science** | An approach to the scientific process based on open cooperative work, tools and diffusing knowledge (GA, Annex 5, Art 16) |
| **OpenAIRE** | Open Access Infrastructure for Research in Europe |
| **Personal data** | Any information relating to an identified or identifiable natural person ('data subject'); 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 [GDPR, Art 4(1)] |
| **Personal data breach** | A breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed [GDPR, Art 4(12)] |
| **Privacy by design** | *The term "Privacy by Design" means nothing more than "data protection through technology design." Behind this is the thought that data protection in data processing procedures is best adhered to when it is already integrated in the technology when created.*[1] |

[1] Intersoft Consulting. (2021). Privacy by design. Available at: https://gdpr-info.eu/issues/privacy-by-design/ [last accesed: 14/05/2024].

| | |
|---|---|
| **Pseudonymisation** | *Pseudonymisation* means "the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person" [GDPR, Art. 4(5)]. What differentiates pseudonymisation from anonymisation is that the latter consists of removing personal identifiers, aggregating data, or processing this data in a way that it can no longer be related to an identified or identifiable individual.[2] |
| **QREM** | SUNRISE's Quality, Risk and Ethics Manager |
| **R&I** | Research and innovation |
| **Re3data** | Registry of Research Data Repositories |
| **School-as-a-living-lab** | The digitally-enhanced programme and its components will be developed through evidence-informed co-creation with school-as-a-living-lab methods (https://www.schoolsaslivinglabs.eu/) from day one, with the participation of socially disadvantaged populations such as migrants and people with low socioeconomic status, to ensure equity in access. |
| **Sensitive data** | Special categories of personal data which are, by their nature, particularly sensitive in relation to fundamental rights and freedoms and merit specific protection as the context of their processing could create significant risks to the fundamental rights and freedoms. These categories include genetic data, biometric data, data concerning health, data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and data concerning a person's sex life or sexual orientation [GDPR, Art 4(13, 14, 15) and Art 9 (1)] |
| **SM** | SUNRISE's Scientific Manager |
| **SUNRISE** | Sustainable Interventions and healthy behaviours for adolescent primary prevention of cancer with digital tools |
| **TOMs** | Technical and organisational measures |

[2] Zerdick, T. (2021). Pseudonymous data: Processing personal data while mitigating risks. European Data Protection Supervisor. https://www.edps.europa.eu/press-publications/press-news/blog/pseudonymous-data-processing-personal-data-while-mitigating_e [last accessed: 14/06/2024].

| | |
|---|---|
| **URI** | Uniform Resource Identifier |
| **WP** | Work package |
| **Zenodo** | Zenodo is a general-purpose open repository developed under the European OpenAIRE program and operated by CERN. |

# Executive Summary

This public report is the first iteration of deliverable 7.5 “Data Management Plan and Ethics” (DMP), which is the first deliverable of Task 7.3 “Ethics and Data Management” in the context of the SUNRISE project, funded by the EU Horizon Europe Research and Innovation Programme under Grant Agreement (GA) No. 101136829. It complies with the provisions on data management, data protection and ethical compliance requirements set out in the Grant Agreement (Article 14 and Annex 5) and in the Consortium Agreement (Section 4.5), as well as in the Horizon Europe Programme Guide and applicable EU, international and national law.

The main objective of the current DMP is to define SUNRISE’s project data governance and ethics policy. It outlines the nature of the data to be handled in the context of the activities and the methodologies by which that data is to be collected, processed, stored, shared and protected, during and after the end of the project. In addition, it presents the project's understanding of the ethical and legal obligations concerning data management and consent acquisition in research activities and provides consortium partners with guidance tools to comply with privacy and security requirements in the management of the research activities and its related data.

Later versions of this initial DMP, updated properly with input from the consortium partners based on the Data Management questionnaire and the Ethics Management questionnaire (Appendix A and Appendix B), will include a more detailed description of specific data management procedures as well as of the nature of the data to be processed and handled throughout the project. This DMP is a “living” document, which means it will be continuously updated following the development of the project’s activities and any updates in data management requirements.

This deliverable also provides both Data Protection Information Notice and Informed Consent templates for partners to adapt and use in the respective tasks (see Appendix C and Appendix D).

This report aims to cover all the questions provided in the Horizon Europe Data Management Plan (HE) template (published: 01-04-2022).[3]

In all, this first version of SUNRISE’s Data Management Plan and Ethics establishes a foundational framework for data governance and ethics. It outlines the types of data, methodologies for collection, processing, storage, sharing, and protection, and addresses ethical and legal requirements. Future versions will enhance these guidelines based on feedback from the Data Management and Ethics Management questionnaires, providing detailed procedures and ensuring compliance with privacy and security standards. This DMP includes adaptable templates for Data Protection Information Notices and Informed Consent. SUNRISE is committed to upholding responsible data management and high ethical standards throughout the project.

[3] European Commission. "Data Management Plan (HE):V1.1 – 01.04." EU Grants. Last modified April 1, 2022. Available at: https://ec.europa.eu/docs/temp-form/report [last accessed: 11/03/24].

# 1 Introduction

## 1.1 The SUNRISE project

The burden of cancer incidence and mortality is rapidly increasing and is responsible for an estimated 9.9 million deaths worldwide in 2020.[4] However, a significant proportion of cancers can be prevented by implementing evidence-based strategies to promote healthy behaviours.[5] Adolescence plays a crucial role in this, as it is a critical phase in which many of the preventable risk factors arise.

Considering this substantial challenge, SUNRISE project's aim is to co-create, implement and evaluate an innovative digitally-enhanced life-skills programme for primary prevention of cancer through health behaviour change in adolescents. To tackle the health and societal challenge of primary prevention of cancer in Europe, SUNRISE will combine an established, evidence-based digital solution for smoking prevention, with novel intervention approaches such as peer social media campaigns, advertising literacy training, educational games, and social robot platforms, to take cancer prevention approaches for adolescents in the EU to the next level. The digitally-enhanced programme will be implemented and evaluated at large scale across 154 schools and 7500 students in urban and rural regions of 8 European countries, including socially disadvantaged groups and ethnic minorities.

## 1.2 Purpose and Relevance of the Data Management Plan

This first version of SUNRISE's Data Management Plan and Ethics deliverable (DMP) describes the project's approach, alongside with the strategies and methods used for the collection, processing and storage of personal and non-personal data, in consistency with legal and ethical requirements. It describes the sort of data that is expected to be collected or produced, and how it will be processed during and after the project duration. DMP also outlines data exploitation, data protection and data preservation for future use. Furthermore, it provides the project's understanding of relevant ethical compliance concepts and

[4] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A *Cancer Journal for Clinicians, 71*(3), 209-249. https://doi.org/10.3322/caac.21660.

[5] Kepper, M. M., Walsh-Bailey, C., Zhao, M., Parrish, L., Miller, Z. M., Glasgow, R. E., Fuentes, L. D., Yan, Y., Hayashi, R. J., Brownson, R. C., & Foraker, R. E. (2024). Satisfaction and effectiveness of a digital health tool to improve health behavior counseling among adolescent and young adult cancer survivors: A randomized controlled pilot trial. BMC Digital Health, 2(1). https://doi.org/10.1186/s44247-024-00064-1.

practices for research activities. This DMP will comprehensively, regularly and appropriately be updated to comply with Horizon Europe requirements and overall European regulations.

The DMP is to delineate overall the data management processes and how this data fits together to set up required systems and deliver clean, complete, and consistent data for analysis. It also establishes a common approach towards ethical and legal compliance concerning data protection out of research activities. The first version of the document is published at M6 and will be continuously updated until M52. The DMP abides by the templates of the Principal Researcher to the Digital Curation Centre Downline tool[6], which matches the Guidelines on Data Management in Horizon Europe. This DMP has been prepared by taking into account the Horizon Europe Data Management Plan (HE) template V1.1 – 01.04.2022.

## 1.3 Relation to the other work packages, tasks and activities

The present deliverable is an integral part of WP 7 and aims to define common grounds concerning data management and ethical requirements and motivate participation and collaboration between partners within the SUNRISE consortium in that regard. For this reason, D7.5 relates horizontally to all activities of the project as it cuts across different work packages to ensure consistency and compliance in managing data and following ethical guidelines throughout them. To achieve its aims, the DMP sets out collaboration and coordination instances among project partners to establish and use common standards and protocols. In particular, D7.5 defines data management and ethical procedures that each activity within the different work packages is expected to follow.

Table 2 – Deliverable context

| Project item | Relationship |
|---|---|
| **Objectives** | The deliverable provides the data management plan and ethical requirements for the project's research activities and is applicable to all the different tasks and activities of each Work Package. |
| **Exploitable results** | The deliverable presents a model for data management and ethics and will serve as a reference for the consortium in its exploitation and dissemination efforts. |
| **Work plan** | The deliverable is a living document and will constantly be updated according to the DoA. |
| **Related deliverables** | The deliverable is to be read in conjunction with the Projects' Handbook, Publication Policy, Quality Assurance and Risk Management Plan and IPR |

[6] Donnelly, M., Jones, S., & Pattenden-Fail, J. W. (2010). DMP online: The digital curation centre's Web-based tool for creating, maintaining and exporting data management plans. *International Journal of Digital Curation*, *5*(1), 187-193.

| | |
|---|---|
| | Policy, as well as in line with both the Grant Agreement and the Consortium Agreement. |
| **Risks** | The constant update of information from tasks and activities within the different work packages will need to be monitored. Particular attention will need to be devoted to data sharing and ethical issues. |

## 1.4 Structure of the Data Management Plan

The data management is structured in 12 subsequent sections, which are complemented with 5 Appendixes. After Section's 1 introduction, Section 2 provides an overview of the various roles and responsibilities for data management and ethical compliance within the SUNRISE consortium. Sections 3 to 10 are organised following the official Horizon Europe Data Management Plan (HE) template. Under each of these sections, the corresponding questions according to that template are considered, and preliminary answers are provided where applicable. After summarising the types and formats of data to be collected or generated in SUNRISE per work package (Section 3), Open Science principles and requirements of FAIR data processing are outlined (Sections 4 and 5). Other research outputs and the allocation of resources are dealt with in Sections 6 and 7. Data security, ethical considerations, and data protection aspects are addressed in Sections 8 and 9. In Section 11, the data management commonalities among the projects that constitute the "Prevention & early detection (behavioural change)" cluster are presented. The data management plan concludes with a chapter summarising how the DMP composes both general standards and principles as well as implementable measures that all project partners in SUNRISE must adhere to (Section 12).

# 2 Roles and Responsibilities in Data Management and Ethics

SUNRISE will follow a structured approach to data management in order to identify, assess, and resolve any data management, ethics or data protection-related issues. Importantly, **all consortium partners have equal responsibility for meeting the organisational, ethical, and legal requirements** in the context of the work they undertake in the project. Moreover, it is **every researcher's responsibility** to identify the appropriate and applicable ethical and legal provisions and ensure compliance when dealing with data collection, handling, processing, publishing and/or storing.

However, there are certain specific tasks and responsibilities with respect to data management and ethics that should be led by partners occupying specific roles within the SUNRISE project.

**The Project Data Management Leader (T7.3 Leader - PBY) is responsible for:**

- Developing the Data Management Plan and Data Protection Impact Assessment at the project level in cooperation with the Project Manager and Technical Coordinator (WP 7 leader).
- Overseeing, requesting, and gathering responsible partners' data collection, data processing and data protection measures information, to foster alignment with the Data Management Plan.
- Writing and updating the Data Management Plan (D7.5, D7.6, D7.7).
- Providing support to the Technical Coordinator, and the Work Package and Specific Activities' Data Managers to implement and comply with the data management and ethics requirements.
- Monitoring the data management activities at the project level, including the project's implementation-related data management.
- Assist the Ethics Committee in its oversight duties on the data management and ethical compliance within the project.

**The Technical Coordinator (WP 7 leader - CERTH) is responsible for:**

- Ensuring the alignment of data management practices with the projects' objectives and the other general project policies.
- Providing support along with the Project Data Management Leader to all partners to implement data management and ethical requirements at both the project and the work package level.
- Providing solutions to specific technical issues in accordance with technical coordination and project management duties.
- Communicating data management requirements to all project members and eventual collaborators.
- Coordinating and ensuring technical aspects related to safe data collection, storage, and analysis, including setting, managing and updating the shared cloud environment and the data repositories, as well as for keeping the data for the duration of the project.
- Establishing security measures to ensure that the data stored in the shared cloud environment for the project remains safe and is located within the EU borders.
- Facilitating collaboration between project teams for data sharing and integration.

- Monitoring data management activities (both collection and publication) and deadlines and sending reminders to WP data managers.
- Ensuring alignment between the DMP, the publications policy and the quality assurance and risk management policy.
- Supporting Work Package Data Managers in the identification and appointment of Activity-specific Data Managers.

**The Work Package Data Managers (all WP leaders) are responsible for:**

- Coordinating data collection and the implementation of the data management and ethics policy within the WP they lead.
- Ensuring data quality and accuracy, as well as the availability of metadata with regard to the activities conducted within the WP.
- Ensuring data protection and ethics compliance measures are taken within the WP activities.
- Identifying activities within the WP where an Activity-specific Data Manager needs to be appointed or a specific measure needs to be taken due to the nature of the data in itself, the data collection process, or the data management or ethical requirements, and inform about it to the Project Data Management Leader.
- Collaborating with the Technical Coordinator and the Project Data Management Leader to monitor data management plan implementation and deadlines, including sending reminders to partners where appropriate.
- Providing input to the data management plan by analysing and summarising the WP-specific data-related activities, gathering the appropriate information out of the WP activities and the partners involved and regularly updating the Data Management questionnaire.
- Contacting and reporting to the Technical Coordinator, the Project Data Management Leader, the QREM and/or the Ethical Advisor Committee on any critical issues identified concerning the data management and ethics compliance within the WP activities, including questions on ethical and privacy issues that may forbid the publication of the data.

**The Activity-specific Data Manager (as appointed by WP leaders or the Technical Coordinator) is responsible for:**

- Managing the data management specific to a particular task or activity as agreed with the corresponding Work Package Data Manager or the Technical Coordinator.
- Ensuring the data collection, processing and storage for the specific activity follow the data management requirements provided for at the project's level.
- Providing activity participants with support and guidance on the data management processes.
- Monitoring data collection, usage and access permissions within the activity.
- Ensuring the activity complies with context-specific requirements (such as national laws governing consent acquisition for certain data collection and ethical approval).

**The Communication & Dissemination Data Manager is responsible for (CDM):**

- Assisting all partners in choosing the right data publication path (green or gold open access, to be aligned with publication policy) and providing further guidance for publishing scientific or other types of publications, with the support of the project's Scientific Manager.
- Ensuring that the open access policy of the journal complies with the Horizon Europe open data requirements before the researcher submits a manuscript.
- Monitoring that green access (self-archiving) publications are deposited in repositories and sending reminders to partners.
- Monitoring that metadata about publications is made available in the R&I Participant Portal and on the SUNRISE website.
- Monitoring that open results (data and software) are deposited in the appropriate repository and properly linked to SUNRISE project.
- Monitoring that research data related to a publication is made available in repositories and linked to the respective publication.
- Monitoring possible embargo periods and sending reminders to partners.
- Monitoring that publications available in public repositories are properly linked with SUNRISE.

**The Quality, Risk and Ethics Manager (QREM) is responsible for:**

- Performing a quality assurance and ethics assessment of open data before their publication, when required.
- Coordinating with the Technical Coordinator, the Project Data Management Leader and the Ethical Advisor Committee to assess critical issues arising from data management and ethics compliance.

**The Researchers (all partners and researchers) are responsible for:**

- Informing the Activity-specific, Work Package and the Communication & Dissemination Data Manager when new open data / papers ready for publication are available.
- Describing the data (by means of appropriate metadata) or scientific publication in accordance to the SUNRISE data management policy (e.g. according to the chosen metadata standard) and with help of the tools (e.g. questionnaire, template, web form, etc.) provided by the project.
- Depositing (publishing into a repository) the data or scientific publication in accordance to the SUNRISE data management policy.
- Complying and enforcing compliance with data protection and ethical compliance requirements in relation to the management of data in research activities.
- Promptly adopting risk mitigation measures for any incident or potential incident related to the processing of personal data, its security or other ethical compliance issues and reporting to the Technical Coordinator, the Project Data Management Leader and/or the QREM.

Table 3 – Data management roles per SUNRISE partner

| Data Management Role/s | Related Project Role/s | Partner/s |
|---|---|---|
| Project Data Manager/QREM | Task 7.3 Leader / QREM | PBY |
| Technical Coordinator | WP 7 leader / Project Coordinator | CERTH |
| Communication & Dissemination Data Manager | CDM / Responsible for Dissemination and Communication activities, including scientific publications | PASYKAF |
| WP 1 Data Manager | WP 1 leader | HMU |
| WP 2 Data Manager | WP 2 leader | PASYKAF |
| WP 3 Data Manager | WP 3 leader | CERTH |
| WP 4 Data Manager | WP 4 leader | OSA |
| WP 5 Data Manager | WP 5 leader | ISGF |
| WP 6 Data Manager | WP 6 leader | PBY |
| Activity-specific Data Manager | Responsible for pilot use case and related activities in Greece | HMU |
| Activity-specific Data Manager | Responsible for pilot use case and related activities in Switzerland | ISGF |
| Activity-specific Data Manager | Responsible for pilot use case and related activities in Cyprus | PASYKAF |
| Activity-specific Data Manager | Responsible for pilot use case and related activities in Belgium | UGENT |
| Activity-specific Data Manager | Responsible for pilot use case and related activities in Romania | IOCN |
| Activity-specific Data Manager | Responsible for pilot use case and related activities in Slovenia | AMEU |
| Activity-specific Data Manager | Responsible for pilot use case and related activities in Spain (Valencia) | FISABIO |
| Activity-specific Data Manager | Responsible for pilot use case and related activities in Italy | FAVO |

| | | |
|---|---|---|
| Activity-specific Data Manager | Responsible for pilot use case and related activities in Spain (Basque) | BBHRI |
| Activity-specific Data Manager | Responsible for execution of the multi-site implementation study | OSA |

**All partners with a specific data management role** assigned to them (as per Table) that have under their scope activities involving the collection or processing of data **should complete the questionnaire** provided in Appendix A as required and keep it updated as appropriate upon changes in circumstances or at the request of the Project Data Management Leader, the Technical Coordinator, the QREM or the Ethical Advisor Committee.

# 3 Data Summary

The SUNRISE project involves carrying out several data collection activities, including personal data and metadata and a set of demonstration tests to assess the technology and effectiveness of the proposed framework. The consortium is committed to following Open Science principles and will also produce Findable, Accessible, Interoperable and Re-usable (FAIR) research data. The management of research data will be conducted in accordance with related soft law instruments governing scientific research such as the European Code of Conduct for Research Integrity, the Guidelines to rules on Open Access to Scientific Publications & Open Access to Research Data in Horizon Europe, and the Guidelines on Data Management in Horizon Europe. Moreover, all processing of personal data will be conducted in compliance with the applicable provisions of a) the GDPR (Regulation (EU) 2016/679), the Universal Declaration of Human Rights and the Convention 108+ for the Protection of Individuals with Regard to Automatic Processing of Personal Data, b) the EU Data Governance Act, Data Act and AI Act, and c) the relevant national regulations, including those governing the acquisition of valid consent.

The data summary section of the DMP outlines the following information:

- The purpose of the collected/generated data and its relation to the overall objectives of the SUNRISE project.
- The types and formats of data already or foreseen to be collected/generated, wherever known at this stage of the project.
- Any existing data that is to be re-used in the SUNRISE project and what for.
- The origin/provenance of the data, either generated or re-used.
- The data utility of the data collected/generated for projects outside of SUNRISE.
- Where available, the expected size of the data.

## 3.1 Purpose of data generation

The purpose of data generation is linked to the achievement of the overall objectives of SUNRISE's original research. The data will serve as a basis for the co-creation of a novel and highly engaging digitally-enhanced programme for primary cancer prevention among adolescents in different socio-economic and cultural contexts in eight European countries. The data collection activities will allow for exploring, monitoring, and evaluating cancer prevention knowledge and identify unmet needs, barriers, and facilitators for the sustainable implementation of the programme. The large-scale multi-country implementation study will enable to comprehensively investigate the impact of the programme on the sustainability of health behaviours and attitudes among adolescents while evaluating the implementation strategy for broad uptake and sustainable use. Ultimately, the data will support policymaking, dissemination activities and exploitation actions aimed at improving awareness, literacy and uptake of digitally-enhanced school-based cancer prevention programmes among educators, health professionals, and policymakers to maximise societal impact. As part of SUNRISE, data is therefore collected at all implementation sites in order to evaluate the effectiveness of the measures in terms of changing young adolescent's health behaviour.

## 3.2 Description of the data per work package

We understand data as provided for in Article 2 of the Data Governance Act[7] as "**any digital representation of acts, facts or information and any compilation of such acts, facts or information, including in the form of sound, visual or audiovisual recording**". SUNRISE will collect, process, and generate various types of data, including personal data and non-personal data,[8] as part of the development of the project and the deployment of its activities. Indicatively, the types of data will mainly concern measurements of behavioural and socio-demographic information.

SUNRISE's data can be divided into three categories: **research-related data, dissemination-related data,** and **project implementation-related data**. WP 1 to 5 predominantly concern research-related data, while WP 6 deals with dissemination-related data and WP 7 handles the data on project implementation. The description of the data in this first version of the DMP is mainly indicative and will be continuously enriched with input from the consortium partners in the following versions. Specifically, information will be collected from the consortium partners through a questionnaire (see Appendix A) and the aggregated inputs will be implemented to the DMP in its corresponding sections.

In the sub-sections below, data summaries and tables per work package provide information on the following questions as per the Horizon Europe Data Management Plan (HE) template:

- *Will you re-use any existing data and what will you re-use it for?*
- *What types and formats of data will the project generate or re-use?*
- *What is the purpose of the data generation or re-use and its relation to the objectives of the project?*
- *What is the expected size of the data that you intend to generate or re-use?*
- *What is the origin/provenance of the data, either generated or re-used?*
- *To whom might your data be useful ('data utility'), outside your project?*

The data summary tables per WP below were filled in by the respective WP leaders.

### 3.2.1 Research-related data

**Research-related data** concerns the data collected for the design, implementation, and evaluation of the SUNRISE project. This will include both personal and non-personal data. The data will be summarised **per work package** based on the input from the responses provided to the questionnaire under Appendix A, as described in the section Roles and Responsibilities in Data Management and Ethics above.

[7] Data Governance Act, available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868. See also Data Act, Art. 2, available at: https://eur-lex.europa.eu/eli/reg/2023/2854.

[8] On personal data, see Section 9.2 below.

Table 4 – Data Summary WP level (Example)

| Data Summary – WP level (Example) | |
|---|---|
| **Type of data collected/generated** | ***Describe the type of data used or generated within the project.*** *Example: Audio, Video, Geospatial, Statistical, Image, Table, Web, Text, Data bases…* |
| **Format and form of data collected/generated** | ***Specify the form and format of the data.*** *Examples of data formats: Text (Type) could be: XML, PDF, Doc, txt, HTML, JSON, RTF, and/or encrypted, compressed. Examples of data forms: basic demographics, physiological data, behavioural/social data (e.g. QoL questionnaires etc.).* |
| **Data origin/provenance of data** | ***Define and describe the origin/source of your data. Data can be gathered from different sources.*** *Examples:*<br>*Observational: Data captured in real time - often not reproducible i.e. sensor readings, images.*<br>*Experimental: Data from lab equipment.*<br>*Derived/Compiled: Data coming from analysis or compilation* |
| **Re-use of existing data** | ***Indicate****: No/Yes (if yes, then specify what kind)* |
| **Expected size of data** | *n (MB, GB)* |
| **Data utility** | ***Explain****: Specific purposes of the data collection/generation or analysis related to the objectives of the project, and possible beneficiaries outside the project* |
| **Related task(s)** | ***Indicate:*** *Tasks of WP involving data collection, processing and/or storage* |
| **Personal data** | ***Indicate:*** *Yes/no (If yes, specify what kind)* |
| **Responsible Data Manager(s)** | ***Indicate****: Partner(s) in charge of the data collection, processing, and storage in each of the WP activities, if applicable* |
| **Data Security & Storage** | ***Explain****: Type of data storage, describe data security policies (for example, only authorized access will be permitted to sensitive data).* |
| **Data value** | ***Explain****: Long and short term of data collection and generation benefits (for example, describe how the data collected and generated will provide short- and long-term benefits to specific target groups of users etc.).* |

**WP 1 –** Data collection and processing in WP 1 activities will relate to the school-as-a-living-lab co-creation methods to engage stakeholders to address barriers and solutions for implementing the cancer prevention program tailored to adolescents. The co-creation phase will focus on identifying pathways for sustainable behaviour change, engaging all societal actors, and prototyping ideas for real-world testing. A total of at least 16 co-creation sessions will be conducted. To inform the co-creation phase, a number of surveys, workshops and (semi-structured) interviews will be conducted. The identified data to be gathered

is expected to comprise information about the (digital) health literacy, cancer literacy, cancer risk behaviours and other environmental factors of adolescents. Additionally, information about feasible cancer prevention pathways and digital tool requirements will be gathered to inform the development of the program. The resulting deliverables concern two written reports on 1) the multi-actor requirements including psycho-social, ethical, and legal requirements, and 2) a cancer prevention programme design document. Table indicatively outlines the types and formats of data to be generated/re-used in WP 1 of the SUNRISE project.

Table 5 – Data Summary Work Package 1

| Data Summary – WP 1 | |
|---|---|
| **Type of data collected/generated** | Answers on structured questionnaires and textual notes and reports. |
| **Format and form of data collected/generated** | PDF, Doc for the text, Database records for the questionnaire answers |
| **Data origin/provenance of data** | Observational: Data captured using structured questionnaires and textual reports<br>*Derived/Compiled: Data coming from analysis or compilation* |
| **Re-use of existing data** | No. |
| **Expected size of data** | ~100MB |
| **Data utility** | Identifying pathways for sustainable behaviour change, engaging all societal actors, and prototyping ideas for real-world testing. Capture information about the (digital) health literacy, cancer literacy, cancer risk behaviours and other environmental factors of adolescents. Additionally, information about feasible cancer prevention pathways and digital tool requirements will be gathered to inform the development of the program. |
| **Related task(s)** | T1.1 Situation analysis in urban and rural schools of Europe and definition of risk-based cancer prevention pathways<br>T1.2 Multi-stakeholder requirements: Children, teacher, family, and community<br>T1.3 Psycho-social, ethical and legal requirements<br>T1.4 School-as-a-living-lab research activities for co-creation<br>T1.5 Design of cancer prevention programme facilitating sustainable health behaviour change |
| **Personal data** | Anonymized data including demographics, behavioural and social data. |
| **Responsible Data Manager(s)** | Data Managers and controllers will be all involved partners in the various activities, including also the WP leader and the project coordinating organisation<br>T1.1 - FAVO, PASYKAF, FISABIO, IBBHR, AMEU, UGENT, ISGF, IOCN, HMU, YCE |

| | |
|---|---|
| | T1.2 - HMU, AMEU, AUTH, FISABIO, CERTH, PARTICLE, GEZOND<br>T1.3 - PBY, AMEU<br>T1.4 – ALL<br>T1.5 – ALL |
| **Data Security & Storage** | Type of data storage<br>• Secure data storage and communications based on encryption<br>• Authorized access with username/password to the data through an identity management system for the structured data<br>• Secure repository of the project for the textual notes and reports<br>Data security policies<br>• Storing only de-identified data<br>• Signed informed consent for stakeholders participating in the council workshops |
| **Data value** | All gathered data will set the requirements and guide the development of the methods and tools delivered through the program. |

**WP 2 –** Data collection and processing in WP 2 will mainly concern thorough desk research activities complementary to the co-creation efforts in WP 1, to inform the design and implementation of the program. To this aim, market research, literature reviews, and a Delphi-survey will inform the development of implementation strategies tailored to the needs of adolescents and diverse societal actors. The resulting data will concern an inventory of evidence-based cancer prevention behaviours and strategies, a repository of multimedia content for health promotion targeting preventable behaviours, an implementation plan to enhance sustained implementation and cost-effectiveness, and an action plan on multicultural incentives for sustained health behaviour adoption. Table indicatively outlines the types and formats of data to be generated / re-used in WP 2 of the SUNRISE project.

Table 6 – Data Summary Work Package 2

| **Data Summary – WP 2** | |
|---|---|
| **Type of data collected/generated** | Audio, video, Statistics, Images, websites, links, texts, databases |
| **Format and form of data collected/generated** | Text will be in the form of: excel, PDF, word Doc, txt, HTML, PPT, encrypted, compressed when needed.<br>Data forms: demographics, physiological data, behavioral/social data including questionnaires and surveys. |
| **Data origin/provenance of data** | Data coming from analysis and from literature resources. |

| | |
|---|---|
| **Re-use of existing data** | Yes, data and information will be gathered/used from previous studies or publications, as well as national/international guidelines and recommendations |
| **Expected size of data** | Not available now |
| **Data utility** | The information that will be collected in the WP 2 will be links from trustworthy resources that will include data from other studies and/or publications. Any data collection that will be collected through questionnaires will be anonymous for the purposes of the project and the deliverables only. |
| **Related task(s)** | The tasks of WP involving data will be related to all tasks: T2.1, T2.2, T2.3 and T2.4. |
| **Personal data** | Not personal data will be required |
| **Responsible Data Manager(s)** | To be confirmed. |
| **Data Security & Storage** | Not applicable |
| **Data value** | The information that will be collected in the WP 2 will be links and files from trustworthy resources that will include data from other studies. Any data collection that will be collected through questionnaires will be anonymous and for the purposes of the project and the deliverables only. |

**WP 3 –** Data collection and processing in WP 3 relate to further developing the digital tools for improvement, tailoring and scaling-up of evidence-based interventions toward sustainable health behaviour change. In terms of the SmartCoach intervention programme, developed by partner ISGF, data will be selected through mobile apps to prevent the problematic substance consumption by promoting life skills. In addition, data will be gathered in terms of social media campaign, social robot platforms, educational games, and digital tools for the authoring and monitoring of the interventions. The SUNRISE authoring and monitoring tools will enable educators and health experts to tailor and personalise the digital tools according to the needs of the target populations in each country. Additionally, monitoring tools will provide implementers with insights into program participation, adherence rates, and intervention outcomes, allowing for timely adjustments. Application Programming Interfaces (APIs) of the digital solutions will enable secure data storage and communication. The digital platform with the authoring and monitoring tools in SUNRISE will employ an identity management system for the secure and consented access of educators and health experts. To protect sensitive data, data de-identification techniques according to the European Agency for Cybersecurity (ENISA) will be used (see Section 8.1.1). Table indicatively outlines the types and formats of data to be generated / re-used in WP 2 of the SUNRISE project.

Table 7 – Data Summary Work Package 3

| **Data Summary – WP 3** |
|---|

| | |
|---|---|
| **Type of data collected/generated** | *Statistical, Images, Videos, Tables, Web, Text, Databases.* |
| **Format and form of data collected/generated** | Data format: Service providers will return data in JSON or CSV format according to specified requirements.<br>Data form: Basic demographics, behavioral, and social data. |
| **Data origin/provenance of data** | *Observational: Data captured in real time - often not reproducible i.e. sensor readings, images.*<br>Self-reports on behavioural/social aspects, images and videos created by SUNRISE partners or free to use images/videos, data coming from user interaction within the SmartCoach mobile app, social media influencer campaign, digital games, educational packages, social robots and conversational assistants.<br>*Experimental: Data from lab equipment.*<br>Recommendations for education activities to the students (games, videos, educational content, etc.) by exploiting collaborative filtering and context-based recommendation techniques.<br>Data coming from an advanced visual analytics dashboard enabling the interactive visualization and clustering of the collected data.<br>*Derived/Compiled: Data coming from analysis or compilation* |
| **Re-use of existing data** | *No.* |
| **Expected size of data** | *<10 GB* |
| **Data utility** | The main purpose of data collection in the context of WP 3 is to provide interactive digital tools for the improvement, tailoring and scaling-up of evidence-based and novel interventions toward the life-skills training of adolescents and the sustainment of healthy behaviours for primary prevention of cancer. Possible beneficiaries outside the project: adolescents from different countries, parents, educators, healthcare professionals, associations, local authorities, policy makers etc. |
| **Related task(s)** | • T3.1 Prevention of the onset and escalation of smoking and promotion of healthy eating habits through influence in popular social media<br>• T3.2 Health promotion through social robots and educational games<br>• T3.3 Interactive module for adolescent, family and educator literacy on cancer prevention<br>• T3.4 Security and privacy module<br>• T3.5 Development of digital platform for creation, adaptation, monitoring, and evaluation of health behaviour change activities |
| **Personal data** | Yes - Demographics, behavioural and social data. |

| | |
|---|---|
| **Responsible Data Manager(s)** | Data Managers:<br>T3.1 – ISGF for SmartCoach, UGENT for social media influencer campaign<br>T3.2 – CERTH for social robot and conversational assistant, UGENT for advertising literacy training games, AUTH for games based on interactive simulation scenarios<br>T3.3 – GEZOND LEVEN for educational package, PARTICLE for mobile app<br>T3.4 – HMU<br>T3.5 – BRIDG |
| **Data Security & Storage** | Type of data storage<br>• Secure data storage and communications based on encryption<br>• Authorized access with username/password to sensitive data through an identity management system, with different access rights for health experts, educators, parents and adolescents<br>• Deployment and data storage of SUNRISE digital interventions to be defined: It could be Microsoft Azure, Amazon AWS, other cloud providers chosen by the schools of the implementation countries, in-house cloud systems of partners in the SUNRISE consortium, or local infrastructures at the schools. All data storages will reside in EU and Switzerland.<br><br>Data security policies<br>• Digital consent mechanism for a) participation in the implementation study, b) access to data in the mobile app (T3.3) and digital platform (T3.5) along with user authentication mechanisms, according to GDPR.<br>• At M5-M9, HMU will proceed with data de-identification techniques according to ENISA guidelines to protect sensitive data.<br>• At M9-M16, HMU and PARTICLE will develop an identity management system providing robust access rights to children, educators, families, and public health experts, using encryption.<br>At M30, the security and privacy module will be adapted in the different countries according to evolving needs. |
| **Data value** | The collected data will provide adolescents with useful insights on how they can improve and sustain their health behaviours. Furthermore, the data will provide insights for parents, educators, health experts and policy-makers, on how they can provide an appropriate environment for adolescents toward the adoption of healthy habits and the primary prevention of cancer. |

**WP 4 & WP 5** – Data collection and processing activities will involve conducting a multi-country implementation study at European urban and rural schools to analyse the effectiveness of the programme on the sustainability of health behaviours and change attitudes of adolescents, as well as to assess the implementation strategy for adoption and sustained use of the programme. To this aim, data on health

behaviours, knowledge, and overall well-being of adolescents will be collected through online questionnaires, social media metrics, and mobile and platform usage logs. Focus groups and (semi-structured) interviews will further inform the perceived impact of the programme through thematic analysis. Additionally, various data on usage, engagement, acceptability, and perceived usefulness of the programme will be assessed to continuously monitor the impact of the interventions and provide feedback for refinements. Table indicatively outlines the types and formats of data to be generated / re-used in WP 2 of the SUNRISE project.

Table 8 – Data Summary Work Packages 4 & 5

| Data Summary – WP 4 & WP 5 | |
|---|---|
| **Type of data collected/generated** | Data within WPs 4 and 5 will be collected via study participants' self-reports and automatically saved user data of the program participants. Typically, self-reported data, will be assessed online via smartphone, personal computer or tablet computer. In sum, the total dataset will include approximately 300 variables (columns) from approximately 7500 study participants and should not exceed 10 MB. The dataset will be thoroughly described by metadata. These metadata, including a detailed study protocol, document the scope of the data, limitations about the data that other users should be aware of, and whether it is raw or processed data. The variable names are explained within the SPSS data file or self-explanatory. |
| **Format and form of data collected/generated** | The data includes numeric and text variables with approximately 200 self-reported variables assessed at each of the two (Study2) or three (Study1) assessment points (baseline, six- and eighteen- months follow-up) and approximately 100 variables on user interactions within the digital intervention programs provided (SmartCoach in Study 1 and a platform with several digital tools for cancer prevention in adolescents in Study 2).<br>The derived Integer and String-data will be password-encrypted and ZIP-compressed and stored as CSV-files as well as texts in JSON-format. |
| **Data origin/provenance of data** | Observational: Data captured in real time - often not reproducible i.e. sensor readings, images.<br>Experimental: Data from lab equipment.<br>Derived/Compiled: Data coming from analysis or compilation<br>Data within WPs 4 and 5 will be collected via study participants' self-reports and automatically saved user data of the program participants. Typically, self-reported data, will be assessed online via smartphone, personal computer or tablet computer. |
| **Re-use of existing data** | We will not re-use data of external or existing sources. |
| **Expected size of data** | The total dataset will include approximately 300 variables (columns) from approximately 7500 study participants and should not exceed 10 MB. |

| | |
|---|---|
| **Data utility** | Data will be collected and analysed in order to test the reach, adoption and effectiveness of several newly developed digital tools for cancer prevention in adolescents. |
| **Related task(s)** | Data collection is administered via 128bit encrypted and password protected SSL-connections by the different digital intervention programs. The online assessments allow various forms of data validation, like accepting only numbers for number fields, ensuring that users do not mark multiple choices for single choice question, or instating on mandatory fields to be filled out. On a monthly basis, data is extracted from the databases of the digital interventions provided and stored on fileservers of the responsible institutions providing the interventions for further processing and archiving. The hosting institutions provides automatic daily, weekly and monthly backups. During data collection, data access is restricted to the institution responsible for each intervention and related scientific personnel.<br><br>After completion of data assessments, all de-identified self-report and user data will be available as SPSS files for all partner institutions of the SUNRISE consortium. These data files include definitions of the variable names and whether a variable contains raw or processed data. Metadata include SPSS syntax files, which allow researchers to understand how processed data was generated and to re-calculate processed data.<br><br>During and after the publication phase within the SUNRISE project, the data will be made available on the Dryad Digital Repository. This is a curated resource that makes the data underlying scientific publications discoverable, freely reusable, and citable. Dryad provides a general-purpose home for a wide diversity of datatypes. Other researchers using these data and publishing on these data, should cite the original publication and the Dryad data package. Other researchers are informed in the respective publications that the data are accessible. |
| **Personal data** | Only few personal data (mobile phone number and a username) will be assessed. These data will be linked temporarily to the other data to be scientifically analysed (self-reported data, user data) by means of lists which link the names of the participants to their test subject-code (ID). The personal data will be deleted as soon as possible, i.e., after completion of the follow-up assessment for each study participant. This allows for the scientific data to be irreversibly anonymized. |
| **Responsible Data Manager(s)** | The institutions that provide the intervention programs in WP 4 will be responsible for data collection, processing and storage. |
| **Data Security & Storage** | On a monthly basis, data are extracted from the databases of the digital interventions provided and stored on fileservers of the responsible institutions providing the interventions for further processing and archiving. The hosting institutions provides automatic daily, weekly and monthly backups.<br><br>Access to personal data during data collection will only be permitted to scientific personnel providing the interventions and conducting the assessments. These are subject to the duty of confidentiality. The personal data will be deleted as soon as possible, i.e., after completion of the follow-up assessment for each study participant. This allows for the scientific data to be irreversibly anonymized. |

| | |
|---|---|
| **Data value** | The results of the evaluations conducted in WP 5 will inform the scientific community and prevention workers on barriers and facilitators of innovative digital programs for cancer prevention in youth, their perceived usefulness, acceptance, effectiveness and cost-effectiveness. |

### 3.2.2 Dissemination-related data

SUNRISE will have a specific dissemination and exploitation strategy, addressed under WP 6. Accordingly, **dissemination-related data** (various information about the project's activities) will be used to publicise its activities. This includes the production of scientific publications, press releases, podcasts, promotional videos, brochures, infographics, a logo, as well as a report on recommendations for scalable and easily implementable cancer prevention pathways in schools. All scientific publications and their data management will also follow the rules set in the project's Publications policy.

SUNRISE partners involved in the dissemination strategy will attend several conferences as part of their dissemination activities and ensure compliance with open-access practices in scientific publications. The dissemination plan includes organising two project-dedicated public events in the last project year to share findings and boost awareness in the field. Where dissemination activities involve the use of data of identifiable individuals, such as for the publication of attendance lists for the project-dedicated events, this will be done by obtaining explicit consent and will be processed and stored in a GDPR-compliant manner.

Table 9 – Data Summary Work Package 6

| Data Summary – WP 6 | |
|---|---|
| **Type of data collected/generated** | Text data (scientific publications, press releases, white paper on recommendations, exploitation strategy plan, health technology assessment study).<br>Audio data (podcasts)<br>Video data (promotional videos, recordings from webinars and seminars if applicable).<br>Web data (online content from public events and webinars, online dissemination materials).<br>Image data (graphical elements in scientific publications, visual aids in webinars and promotional videos). |
| **Format and form of data collected/generated** | Data format: PDF, DOCX, HTML, PPT, MP3/MP4/MOV, CSV, PNG, XLSX.<br>Data forms: The forms of data include research articles, policy recommendations, press releases, informational podcasts, promotional videos, event summaries, webinar recordings, charts, diagrams, logo, website and infographics. |
| **Data origin/provenance of data** | Data coming from compilation of findings within SUNRISE. Data coming from dissemination, exploitation and communication activities. |

| | |
|---|---|
| **Re-use of existing data** | Data coming from other WPs may be used for dissemination and/or exploitation activities. |
| **Expected size of data** | To be determined. |
| **Data utility** | Dissemination and exploitation including communication activities. Dissemination activities are to build effective awareness of the results, creating understanding and aiming for action among the target audience identified. Exploitation activities include identifying key exploitable results, conducting preliminary market and competition analysis and business modelling. Communication strategy includes specific measures to promote the project itself and the results attained, with the mission to reach out to a critical mass. |
| **Related task(s)** | All the tasks of this WP may involve data management: T6.1, T6.2, T6.3 and T6.4. |
| **Personal data** | Only few personal data is expected to be gathered, potentially in relation to educational webinars, conferences, and networking activities and participation (e.g. participants lists, recordings). |
| **Responsible Data Manager(s)** | Activity-specific Data Managers and data controllers will be the Task leader partner of each activity.<br>T6.1 - PASYKAF<br>T6.2/T6.3 - PBY<br>T6.4 - IOCN |
| **Data Security & Storage** | All data generated or used will be stored, when possible, in the project's Drive environment, access to which is restricted to participating partners. Where personal data is involved, appropriate data protection measures will be taken, including safekeeping on secure local servers and restricting access. The data generated in the framework of dissemination or exploitation activities and intended for publication will not involve classified or sensitive information. |
| **Data value** | The dissemination, exploitation and communication activities of SUNRISE are central to reach, engage and synergise key target audiences and stakeholders, maximising the potential short-term outcomes and long-term impacts of the project and the wide scale roll-out of Key Exploitable Results (KERs). Target groups aimed to be reached through the various measures and channels include citizens, including people at high risk of developing cancer, cancer patients, survivors, and their families; international healthcare and cancer organisations; research and academics; as well as local and national policymakers and authorities. |

### 3.2.3 Project implementation-related data

**WP 7** is aimed at monitoring the successful implementation of research activities from WP 1-WP 6 with the agreed time, costs and quality metrics. In order to implement SUNRISE and achieve its objectives, the project collects data about the consortium partners and their work (**project implementation-related data**). This may include personal data when reporting work tasks, as well as e-mail traffic, the drafting of

reports and the co-creation of other data trails. The implementation of SUNRISE involves various tasks, including maintaining records of activities, deliverables, tasks, and associated individuals' contact details. When dealing with any information identifying individuals, collection and processing will adhere to consent or contractual agreements outlined in the Consortium Agreement (CA).

The operation of the project implementation-related data relies primarily in the Google Drive environment managed by CERTH. Data such as draft and final versions of reports, deliverables, minutes, presentations, surveys, workshops, as well as meeting minutes, will be securely stored and preserved there. SUNRISE may also involve the collection of personal data from consortium members such as email addresses and contact details for the creation of mailing lists and the organization of both in-person and virtual meetings (e.g. video & audio recordings, pictures, field notes, signatures, IP addresses, etcetera).

All project implementation-related collected data will only be used for SUNRISE purposes and will not be used beyond this scope without explicit consent. Recorded data may concern important information about the project and shall not be shared outside the consortium without specific notice.

The personal data collected as part of the project implementation-related activities will be stored in a specified Google Drive folder to which only the project members directly involved in the SUNRISE project have access. The information will only be available for internal project use under the following best practices and conditions:

- Mailing lists, recordings, images, or field notes of meetings will be made available in the SUNRISE Google Drive folder "*SUNRISE_Implementation-folder*".
- The data will be made available for internal project use only. No project partner or employee of a project partner may share the recordings on any public platform without prior and explicit authorization.
- Where personal data is collected from individuals, it should be recorded based on consent in accordance with the privacy notice provided for each recording activity.
- CERTH is the data controller for the personal data stored in the Google Drive environment.

Table 10 – Data Summary Work Package 7

| Data Summary – WP 7 | |
|---|---|
| **Type of data collected/generated** | In WP7 the types of data which will be used/generated are the following:<br>-Personal data when reporting work tasks (e.g. emails, addresses, contact details)<br>-Video, audio recordings, pictures, IP addresses<br>-Images, pictures, field notes<br>-Text (e.g. reports, deliverables) |
| **Format and form of data collected/generated** | Data format: XML, PDF, Doc, txt, HTML, JSON |

| | |
|---|---|
| **Data origin/provenance of data** | Observational data: Data captured in real time - often not reproducible i.e. readings, images.<br>Derived/Compiled: Data coming from analysis or compilation |
| **Re-use of existing data** | All data from the other WPs may be used for the proper management and coordination of the project activities |
| **Expected size of data** | To be determined. |
| **Data utility** | The implementation-related data will be used to: a) monitor the successful implementation of research activities (WP1-WP6) within the agreed time, costs and quality metrics, with continuous management of risks and corrective actions, b) coordinate and manage all administrative, financial and contractual aspects related to the project, c) establish and update the data management plan of the project. |
| **Related task(s)** | Tasks of WP7 that will involve data collection, processing and/or storage:<br>- T7.1 Project Management<br>- T7.2 Technical and Scientific coordination<br>- T7.3 Ethics and Data Management |
| **Personal data** | Yes<br>-Personal data when reporting work tasks (e.g. emails, addresses, contact details)<br>-Video, audio recordings, pictures, IP addresses<br>-Images, pictures, field notes |
| **Responsible Data Manager(s)** | Activity-specific Data Managers and data controllers will be the Task leader partner of each activity.<br>T7.1 Project Management (CERTH)<br>T7.2 Technical and Scientific coordination (CERTH)<br>T7.3 Ethics and Data Management (PBY) |
| **Data Security & Storage** | All data generated or used will be stored in alignment with EU data security policies and guidelines, when possible, in the project's Google Drive environment. Only authorised access restricted to participating partners will be permitted.<br>In case of personal data handling, appropriate data protection measures are provisioned to be taken, including physical mitigation measures such as safekeeping on secure local servers and limited access. |
| **Data value** | In the context of WP7, data is collected about the consortium partners and their work (project implementation-related data). This implemented data is valuable as refers at SUNRISE various tasks, activities, deliverables, tasks, and associated individuals/partners. To this end this data collection is anticipated to bring short benefits for the purposes of the project and form the other hand long term benefits as will ensure the sustainability of the project. |

# 4 Open Science practices

SUNRISE will follow thoroughly all the required actions to be aligned with the Open Science practices as defined in the Horizon Europe guidelines[9]. In particular, SUNRISE will address the following Open Science practices:

**Early and open sharing of research**: During the project, all partners will follow methods and steps that assure the early and open sharing of the project outcomes. More specifically, preregistrations of the research plans in advance of the study implementation, as well as registered reports will be submitted to Open Access Repositories. The principle "as open as possible, as closed as necessary" will be followed throughout the project.

**Research output management and measures to ensure reproducibility of research outputs**: SUNRISE will pay special attention to ensure the reproducibility of the outputs by covering the three main research processes that reproducibility is based on: reproduction, replication, and re-use.

**Open access to research outputs and participation in open peer-review**: All the research outputs that will be produced from the implementation of SUNRISE will be aligned with the Open Access and Open Science regulations of the EU. More specifically, the publications of the project will be published in **Open Access Journals** which will be initially checked from the SHERPA/RoMEO platform and DOAJ in order to confirm the open access and the copyright policies of specific international journals, and **Open Access Repositories** (e.g. PubMed, Zenodo, arXiv, etc.) which will be identified through platforms as ROAR, OpenDOAR, OpenAIRE or OAD. Moreover, to ensure the reusability of data and other research outputs, SUNRISE partners will be encouraged to release their work, when feasible, under Creative Commons licenses such as the CC BY 4.0 (Attribution) or the CC0 1.0 (Public domain).

**Involvement of relevant knowledge actors including citizens, civil society and end users in the co-creation of research and innovation (R&I) agendas and contents**: During all SUNRISE project's phases the active involvement of the general public and the non-professional scientists will be encouraged. In particular, in WP 1 and WP 6 the involvement of the end-users (citizens, professionals, associations, local authorities, etc.) as well as of other actors such as policymakers, businesses and non-professional researchers, will be crucial for the best evaluation of the projects' results.

[9] European Commission, Open science in Horizon Europe, available at: https://rea.ec.europa.eu/open-science_en [last accessed: 14/06/2024].

# 5 The FAIR requirements

SUNRISE's general data management plan complies with the Open Science provisions under the Grant Agreement (Annex 5, Art 17) and is aligned with the overall EU Open Science policy.[10] The project partners should thus produce **findable**, **accessible**, **interoperable,** and **reusable** research data (aligned with the **FAIR** principles)[11] and ensure that it is soundly managed. Researchers shall be able to find, re-use and access as much data as possible, thereby maximising the effectiveness and reproducibility of the research conducted within the SUNRISE project. All data and metadata involved in SUNRISE project will comply with Zenodo Data repository FAIR data principles.[12]

The following sub-sections are organised according to each of the pillars of the FAIR principles to address the corresponding questions in the Horizons Europe DMP template.[13] Where available, the corresponding answers are provided; alternatively, it is indicated how they will be addressed at a later stage of the project.

## 5.1 Findable data

Making data findable includes the description of the discoverability of data, the identifiability of data and reference to standard identification mechanisms, the naming conventions used, the approach to keywords and clear versioning, and the specification of standards for the creation of metadata.

### 5.1.1 Will data be identified by a persistent identifier?

All open data, publications and open-source software produced in SUNRISE will be identifiable and locatable by means of a persistent Uniform Resource Identifier (URI). For publicly available datasets, publications, and reports (deliverables), the SUNRISE consortium is encouraged to assign a Digital Object Identifier (DOI). Open SUNRISE results that are deposited in the SUNRISE default Open Access repository (i.e. Zenodo) will be assigned a DOI automatically and will benefit also from Zenodo's DOI versioning support.

### 5.1.2 Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case

[10] https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en. Last accessed: 14/06/2024.

[11] https://www.go-fair.org/fair-principles/. Last accessed: 14/06/2024.

[12] https://about.zenodo.org/principles/. Last accessed: 14/06/2024.

[13] https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/temp-form/report/data-management-plan_he_en.docx. Last accessed: 14/06/2024.

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

Based on the premise that it is essential to provide metadata as accurate as possible, since rich metadata significantly improves the searchability and re-usability of data, the SUNRISE project will ensure that every effort is made to this end. There are many different metadata standards for many different types of data, and it may not be possible to find a rule that suits all purposes. For this reason, the preferred option is to take a pragmatic approach, by agreeing on a common, minimum catalogue metadata schema for those datasets that are published in public catalogues and data repositories, and to use specific schema extensions for different types of data, if necessary.

### 5.1.3 Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

All open SUNRISE results deposited in a repository will provide search keywords together with their metadata. Keywords for open data will be selected from controlled vocabularies that are suitable for the specific type of the data (see 5.3.1).

### 5.1.4 Will metadata be offered in such a way that it can be harvested and indexed?

Offering metadata in a way that it can be harvested and indexed means making metadata available in a format and structure that allows automated systems (such as search engines or data repositories) to gather, organize, and store it for easy retrieval and analysis. This enables efficient searching and indexing of information, improving its accessibility and usability for users who rely on metadata to discover and access relevant resources or datasets.

## 5.2 Accessible data

This requirement will be fulfilled by specifying which data is made openly available and in case any data is kept restricted, to specify the reasons. It will be outlined how the data is made available and what methods or software tools are needed to access the data. Furthermore, it is specified where the data, associated metadata and documentation are stored and how access will be provided.

***Repository**:*

### 5.2.1 Will the data be deposited in a trusted repository?

For the SUNRISE project, a central repository will be set up to safely centralise all project documents. All project data will be deposited in trusted Open Data repository (i.e. Zenodo), which assigns DOIs ensuring their findability.

### 5.2.2 Have you explored appropriate arrangements with the identified repository where your data will be deposited?

Currently, there is no need for such an arrangement. Zenodo is OpenAIRE's recommended "catch-all" repository for projects like SUNRISE without ready access to an organised data centre.

### 5.2.3 Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

See answer to 5.2.1.

***Data**:*

### 5.2.4 Will all data be made openly available?

The principle "as open as possible, as closed as necessary" will be followed. Both publications and datasets will be made openly available via Zenodo – part of the European Open Science Cloud (EOSC) – and, whenever possible, data.europa.eu, the official portal for EU data. However, access restrictions related to the following conditions may apply:

- Copyright and permissions for reusing third-party datasets – as it may lead to unclear IPR situations.
- Personal data treatment, consent, and confidentiality – as it may trigger privacy, ethical or security issues.
- User-generated content – it may only be made open with explicit permission from the end user.

In case there are restrictions in place on data needed to validate the results presented in scientific publications, access to individuals with legitimate interests may be granted on request and under a controlled process.

More restrictions may apply. This response will be further elaborated in the next stages of the project.

### 5.2.5 If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Open access to articles is often delayed (embargo period), however it must be granted within six months of publication. On the contrary, an article published under Gold Open Access is immediately available in open access mode by the scientific publisher. The associated costs are shifted away from readers and are covered instead by the university or research institute to which the researcher is affiliated, or to the funding agency supporting the research. Publication fees will be charged on the institution of the first and submitting author. The publication of the research related data or its access on request (when open publication is not possible) will follow the process provided for in the DMP.

### 5.2.6 Will the data be accessible through a free and standardized access protocol?

Access will be granted based on the rules agreed upon in the GA and CA.

### 5.2.7 If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

Where a restriction on open access to research data is necessary, attempts will be made to make data available under controlled conditions to other individual researchers. In the case where restricted or embargoed data is stored in the Zenodo repository, information about the restricted data will be published in the repository, and details of when the data will become available will be included in the metadata. Restricted access allows a researcher to upload a dataset and provide the conditions under which he/she grants access to the data. Researchers wishing to request access must provide a justification for how they fulfil these conditions. The owner of the dataset gets notified for each new request and can decide to either accept or reject the request.

### 5.2.8 How will the identity of the person accessing the data be ascertained?

The identity of the person accessing restricted data stored in SUNRISE's default repository (i.e. Zenodo) will be ascertained according to the process described in section 5.2.7.

### 5.2.9 Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

In case there are any issues regarding the restricted access to research results, SUNRISE's Quality, Risk and Ethics Manager, the Project Coordinator and the Ethical Advisor Committee can act as data access committee and seek clarification.

***Metadata**:*

### 5.2.10 Will metadata be made openly available and licensed under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?

Metadata of **deposited publications** must be open under a Creative Common Public Domain Dedication (CC 0) or equivalent, in line with the FAIR principles (in particular machine- actionable) and provide information at least about the following: publication (author(s), title, date of publication, publication venue); Horizon Europe or Euratom funding; grant project name, acronym and number; licensing terms; persistent identifiers for the publication, the authors involved in the action and, if possible, for their organisations and the grant. Where applicable, the metadata must include persistent identifiers for any research output, or any other tools and instruments needed to validate the conclusions of the publication.

Metadata of **deposited data** must be open under a Creative Common Public Domain Dedication (CC 0) or equivalent (to the extent legitimate interests or constraints are safeguarded), in line with the FAIR principles (in particular machine-actionable) and provide information at least about the following: datasets (description, date of deposit, author(s), venue and embargo); Horizon Europe or Euratom funding; grant project name, acronym and number; licensing terms; persistent identifiers for the dataset, the authors involved in the action, and, if possible, for their organisations and the grant. Where applicable, the metadata must include persistent identifiers for related publications and other research outputs.

### 5.2.11 How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

Data will be accessible for at least ten years following the end of the project, although longer periods may be granted depending on the nature of the data and its potential future relevance. As for metadata, ensuring its availability even after the data is no longer accessible is crucial. Metadata will be maintained and stored separately from the primary data in trusted repositories or archives to ensure its long-term preservation and accessibility.

### 5.2.12 Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open-source code)?

Documentation of (open source) software needed to access the data and developed by SUNRISE will be made available on the SUNRISE website and the respective source code and release repositories.

## 5.3 Interoperable data

As defined by the European Data Protection Supervisor, interoperability of data refers to the functionality of information systems to exchange data and to enable sharing of information.[14] Data from different sources can thereby be accessed, understood, and used by various systems without requiring significant data transformation or other manual intervention. The interoperability of data thus facilitates the sharing of data, integration and collaboration between different platforms or organisations. To this end, the data and metadata vocabularies, standards, or methods to be used will be specified.

### 5.3.1 What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?

Vocabularies for keywords and other metadata properties have yet to be selected and will be reported in future versions of the data management plan.

### 5.3.2 In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more

[14] EDPS. *Interoperability*. Available at: https://www.edps.europa.eu/data-protection/our-work/subjects/interoperability_en . [last accessed: 14/06/2024].

### commonly used ontologies? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?

Currently, SUNRISE does not intend to introduce new project-specific ontologies or vocabularies.

### 5.3.3 Will your data include qualified references to other data (e.g. other data from your project, or datasets from previous research)?

Where appropriate, there will be qualified references to other data from the SUNRISE project as well as to data from previous research.

## 5.4 Re-usable data

To enable the re-use of data generated by the SUNRISE project, it will be specified how the data will be licensed to permit the widest reuse possible, when and to whom the data is made available for re-use, including after the end of the SUNRISE project. A data quality assurance process will be provided if applicable.

### 5.4.1 How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?

Accessibility of Data will be validated using the platform of re3data (https://www.re3data.org/). To ensure reusability of data/research outputs, the data will be well-documented, and they will have clear license and provenance information.

### 5.4.2 Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

In accordance with Article 16 of the GA, results including data, know-how or software are owned by the beneficiary that generates them. However, beneficiaries owning open results arising from the SUNRISE activities are encouraged to release their work under a Creative Commons license, preferably Creative Commons Attribution 4.0 (CC-BY-4.0[15]) (see **Open Science practices** above).

[15] https://creativecommons.org/licenses/by/4.0/. Last accessed: 14/06/2024.

### 5.4.3 Will the data produced in the project be useable by third parties, in particular after the end of the project?

The outcomes generated by the project and stored in an appropriate repository become available for use by third parties post-project conclusion. In instances where confidentiality, security, protection of personal data, or IPR pose constraints on open access to specific research data essential for validating a scientific publication, said data might be placed in a restricted repository. Access could then be provided upon request and subject to the terms of a restricted license (see Article 16 of the Grant Agreement).

Data archiving and sustainability will be guaranteed by the Zenodo digital repository. As a European Commission supported initiative, and technically supported by CERN, Zenodo can ensure access to the generated data continues after the project ends.

### 5.4.4 Will the provenance of the data be thoroughly documented using the appropriate standards?

See response at 5.4.1 above.

### 5.4.5 Describe all relevant data quality assurance processes.

Data quality refers to the set of principles provided for in Article 5 of the GDPR. The corresponding processes are provided for in this DMP section 0 below and in the Quality Assurance and Risk Management Plan.

# 6 Other research outputs

Further to the FAIR principles and according to the Horizon Europe Data Management Plan, the DMP should also address research outputs other than data. Such outputs may be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.). Therefore, in addition to data, the management of such other research outputs that may be generated or re-used throughout the projects will be addressed as required and in line with complementary project's policies, such as the IPR policy. Sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles, will be duly provided.

# 7 Allocation of resources

### 7.1.1 What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.)?

The estimation of costs to comply with the FAIR principles, if any, will be further elaborated in subsequent versions of the DMP.

### 7.1.2 How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)

See answer 7.1.1 above above.

### 7.1.3 Who will be responsible for data management in your project?

The data management plan's first iteration will serve to identify roles and responsibilities for data management in the project and its activities, including how this responsibility is to be assumed by the members of the consortium (see Section **Roles and Responsibilities in Data Management and Ethics** above). The data management activities need to be coordinated and monitored both at the project and work package level. Moreover, some activities under the work package level (such as country-level pilots) require specific data management roles and rules to be set.

### 7.1.4 How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?

No immediate costs are anticipated for open data that is stored for long-term preservation in the Zenodo repository. Additional details will be reported, as needed, in future versions of the DMP.

# 8 Data security

### 8.1.1 What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?

The project will implement all necessary security guarantees for handling the project's data, complying with GDPR and other legal requirements, provided for both in European as well as in national and regional frameworks. The software and data security tools used at SUNRISE are integrated with software engineering best practices, GDPR compliance mechanisms and state-of-the-art security measures.

As mentioned in Section 3.2.3 above, SUNRISE project's coordinating institution CERTH hosts all information related to the management and implementation of the project in a secure environment on the Google Drive cloud platform, which contains security features such as encryption, two factor authentication, and phishing and malware detection tools. The access to such Drive is restricted to consortium partners. To ensure the security of the data stored on this server, CERTH regularly monitors the last elements uploaded and access credentials.

In addition, all SUNRISE partners have experience in data management and data protection, both from their participation in multiple Horizons projects and from their respective institutional tasks and obligations in the European context. Each has (and is responsible for) its internal data protection, privacy, security and ethics policies and for the correct and secure maintenance of its respective servers.

Moreover, datasets made publicly available will have robust recovery and preservation mechanisms – for instance, Zenodo conducts nightly backups and replicates the data files and its metadata in multiple copies in the online system.[16]

On top of that, a number of technical and organisational measures (TOMs) will be implemented to ensure the level of security is commensurate to the risks faced by data subjects in the context of the project's activities. Such TOMs include data encryption, concealment, and data de-identification requirements. To these aims, SUNRISE partners will follow the recommendations and techniques provided for by ENISA in several publications.[17]

For more details on the efforts to be made, including specific TOMs, for personal data protection in particular, see section 9.2.

[16] https://about.zenodo.org/policies/ [last accesed: 14/06/2024].

[17] On pseudonymisation scenarios, best practices and techniques, see https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices and https://www.enisa.europa.eu/publications/deploying-pseudonymisation-techniques; on personal data sharing practices in the health sector. On security of personal data processing, see https://www.enisa.europa.eu/publications/handbook-on-security-of-personal-data-processing. On personal data sharing, see https://www.enisa.europa.eu/publications/engineering-personal-data-sharing.

### 8.1.2 Will the data be safely stored in trusted repositories for long term preservation and curation?

See responses at 5.2.1 and 5.2.4 above.

# 9 Ethics and Data Protection

The SUNRISE consortium understands that the research it will conduct may raise ethical and data protection questions. Therefore, the consortium confirms its strict commitment to the ethical standards, data protection rules and principles of Horizon Europe (in particular Article 19 of Regulation (EU) 2021/695) in order to minimize unnecessary harm to the participating individuals, which will be rigorously applied regardless of the country in which the research takes place.

As part of the Ethics Appraisal Procedure during the preparation phase of the project, the partners conducted the Ethics Self-Assessment and identified some salient issues that would require special consideration during the implementation phase of the research activities. These potential ethical challenges mainly regard the obtention of proper consent and data protection, and are related to the involvement of:

- Human participants, which may be:
    - Volunteers for non-medical studies
    - Potentially vulnerable individuals or groups
    - Children/Minors
- Non-invasive interventions on study participants
- The processing of personal data
- A non-EU country as the location of activities (Switzerland)

In view of these potential issues, the consortium is specifically committed to undertake all activities in compliance with European, international, and national legislation relevant to the country where the data collection activities are taking place. This includes, among others:

- The Universal Declaration of Human Rights and the Convention 108 for the Protection of Individuals with Regard to Automatic Processing of Personal Data;
- The General Data Protection Regulation (GDPR) (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data;
- The Directive 2002/58/EC of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) [2002] OJ L201/37, as modified by the 2009/136/EC Directive, and
- The European Charter of Fundamental Human Rights.

**All partners have equal responsibility for meeting ethical and legal requirements in the context of the work they undertake in the project.**

To protect the rights of the participants while taking into account the general commitment to make ethics an integral part of the Project's management, this DMP helps to create a common understanding of the key concepts and requirements for ensuring the ethical processing of data collected in research, as well as technical and organisational strategies for risk mitigation. To that end, in the following section, the main principles, concepts and procedures for identifying and addressing ethical issues are outlined (Section 9.1). Then, the key concepts of data protection and lawful data processing are presented and recommended TOMs to ensure compliance are listed (Section 9.2). As this is the first version of this

deliverable, it is expected that the approaches will be adapted according to the needs and possibilities of the partners in the development of the tasks.

This section deals with the following questions from the Horizon Europe Data Management Plan template:

- *Are there, or could there be, any ethics or legal issues that can have an impact on data sharing?*
- *Will informed consent for data sharing and long-term preservation be included in questionnaires dealing with personal data?*

## 9.1 Ethics Principles and Procedures

In order to ensure compliance with ethical principles and in light of the identified risks of the project, SUNRISE will implement a series of standards and best practices. To help the consortium address potential risks and enhance mitigation measures, while also facilitating the project's compliance with ethics and data protection requirements, several guidance templates and documents are provided under Appendix C and Appendix D. Beyond complying with the principles and procedures described in the following, SUNRISE relies on the Ethics Advisor Committee, composed of external experts, which will perform professional and specialised ethical oversight tasks to identify and address the risks and needs that may arise.

### 9.1.1 Informed consent

**Consent** is fundamental to both engaging individuals in research activities and to lawfully collecting personal data (see 9.2.1 below). Under these two premises, SUNRISE partners will ensure that this consent is **freely given, informed, specific and unambiguous**.

The participation of individuals in the research activities shall be **fully voluntary**. This means that no data will be collected without explicit informed consent, which shall be obtained before the study begins. SUNRISE partners are expected to disclose information about the corresponding activity's purpose, risks and procedures, as well as the measures to be taken in the case of harm resulting from participation. Participants should be able to fully understand and agree to the procedures/research being undertaken by giving their explicit consent. SUNRISE's consortium acknowledges that the process of obtaining informed consent begins when initial contact is made with a prospective subject and continues throughout the entire course of the study. Whenever children or adolescents are involved, SUNRISE partners will request the appropriate **informed parental consent** in advance. If required, other societal actors' consent will also be obtained (e.g. educators, health professionals).

To these aims, participants will be provided with both an **Informed Consent Sheet and a Data Protection and Information Notice**, for which guidance templates are provided in **Appendix C and Appendix D.** These documents will be drafted in a fully understandable language for the participants and will clearly describe the aims, methods, and implications of the activities (e.g., co-creation, digital tool testing, interviews, workshops, focus groups and online surveys), the nature of their participation, and any benefits or risks (e.g., to privacy) that might be involved. They will also outline how and for what purposes data will be collected, and how it will be used and stored. All participants will be allowed to ask questions and receive understandable answers from the interviewers and test organizers about their participation.

SUNRISE partners shall ensure that appropriate data protection and information notices and consent forms are duly attached to all documents concerning activities that would require the explicit consent of external participants or regard personal data collection activities.

### 9.1.2 Ethical Approval and Ethics Management

SUNRISE partners are well aware of the ethical challenges associated with research involving humans, given their training and extensive experience in research activities. All partners will thus follow the best practices of ethical research and their consequently required procedures.

Whenever needed, research protocols will be submitted for consideration, comment, guidance, and approval to the concerned research ethics committee before an activity begins. Independent or institutional ethics committees should review and consider whether the research is ethical and lawful and provide appropriate safeguards.

Furthermore, to enhance internal monitoring WP leaders will have to report the ethical management of the activities under their WP. For this purpose, a questionnaire will be distributed to all consortium partners, which should be filled in and duly sent by those **partners leading any specific Activities carried out within the SUNRISE project which collects external participants' data** (see Appendix B). The resulting documents should be properly archived in the project's Google Drive space and made available to the granting authority when required.

In case of questions regarding the need for ethical approval for an activity or the related procedures, the partners may raise the issue to the QREM, the Project Coordinator and/or to the Ethics Advisors Committee.

### 9.1.3 Purpose Limitation

SUNRISE's research objective is limited to its scientific and social contribution to protecting and promoting people's health. No data collected in the context of its activities will be sold or used for any purposes other than the current project.

### 9.1.4 Data Minimisation

SUNRISE commits to strictly stick to the principle of data minimisation by avoiding the collection and processing of any unnecessary personal data. Personal data should be kept in a form permitting the identification of data subjects for no longer than is reasonable, proportionate, and necessary for the purposes for which the personal data are processed. Only relevant personal data will be collected and processed, and it will be anonymized as soon as viable. Any shadow (ancillary) personal data obtained during the research activities will be immediately cancelled.

### 9.1.5 Reimbursement and Compensation for Research Participants

Compensation –if provided– will correspond to a simple reimbursement for working hours lost as a result of participating in the study; special attention will be paid to avoid any form of unfair inducement.

### 9.1.6 Research Involving Children and Adolescents

SUNRISE is committed to ensuring that research activities involving children and adolescents are conducted in a manner entirely consistent with the International Ethical Guidelines for Health-related

Research Involving Humans (in particular Guideline 17)[18], and the Commission's Guidance note on Informed Consent.[19] Under this framework, it is recognised that such subjects have specific needs and capacities that place them in particular risk situations that must be specifically considered by researchers. Considering that children and adolescents may need adequate support to protect their own interests, SUNRISE partners commit to obtaining prior to the start of research activities the informed consent of the child's parents or legal representatives, as well as to inform the child or adolescent of the scope of the activity and its intervention in a manner especially tailored to their capacity of understanding and level of maturity. Finally, SUNRISE researchers shall respect the decision of children or adolescents who refuse to participate or to continue to participate in the activities.

### 9.1.7 Privacy and confidentiality

All SUNRISE partners will take every measure to safeguard the privacy and confidentiality of research subjects' personal information. In accordance with GDPR (Art. 4(1)), no personal data should be included in any information disclosed or shared among SUNRISE partners during the project implementation or exploitation activities. Each partner shall remove, anonymize, or otherwise render all personal data in the shared information inaccessible to other parties before it is shared.

## 9.2 Data Protection

Data management is a continuous process and the DMP is intended to be a dynamic document. The current DMP adheres to the EDPB's recommendations and best practices to ensure that all researchers adhere to the principles of lawful and ethical data management, and particularly to the principle of 'Data protection by design and by default', as provided by Article 25 GDPR.[20] These principles are mandatory for all data controllers within the scope of the SUNRISE project.

Data Management questionnaires (Appendix A) have been distributed to the consortium members. However, given the nascent phase of the research, it is not yet possible to definitively describe the precise specifications and handling procedures for the data that will be collected, processed, and stored during the project.

[18] International Ethical Guidelines for Health-related Research Involving Humans, Fourth Edition. Geneva. Council for International Organizations of Medical Sciences (CIOMS) 2016. Available at: https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf) [last accessed 29/04/2024].

[19] European Commission. Guidance for Applicants Informed Consent. Available at: https://ec.europa.eu/research/participants/data/ref/fp7/89807/informed-consent_en.pdf [last accessed: 14/06/2024].

[20] EDPB Guidelines 4/2019 on Article 25 - Data Protection by Design and by Default Version 2.0, adopted on 20 October 2020, available at: https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_en.pdf; EDPB Guidelines 07/2020 on the concepts of controller and processor in the GDPR Version 2.1, adopted on 7 July 2021, available at: https://www.edpb.europa.eu/system/files/2023-10/EDPB_guidelines_202007_controllerprocessor_final_en.pdf.

### 9.2.1 Personal data: key concepts and processing principles

Personal data must be processed in a lawful, appropriate, and transparent manner, and be correct and up to date. For the purposes of providing a common understanding of certain key concepts and roles and to facilitate the implementation of related data protection processes, it is useful to consider the following definitions:

- **Personal data** is 'any information relating to an identified or identifiable natural person'. An 'identifiable natural person', or 'data subject', 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' (Article 4(1) GDPR).
- **Sensitive data** is that belonging to special categories of personal data, including 'racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade-union membership, genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural's sex life or sexual orientation' (Article 9(1) GDPR).
- **Data processing** is 'any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction' (Article 4(2) GDPR).
- **Data controller** is the 'natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data […]" (Article 4(7) GDPR). The actual data processing may be delegated to another party, called the data processor. The data controller is responsible for the lawfulness of the processing, for the protection of the data, and for respecting the rights of the data subject. The controller is also the entity that receives requests from data subjects to exercise their rights.
- **Data processor** is the 'natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller' (Article 4(8) GDPR). The data processor must provide sufficient guarantees in respect of the technical security measures and organisational measure governing the processing to be carried out, and must ensure compliance with such measures.
- **Joint data controllers** are two or more controllers that jointly determine the purposes and means of processing data. These 'shall in a transparent manner determine their respective responsibilities for compliance with the obligations under this Regulation [GDPR], in particular as regards the exercising of the rights of the data subject and their respective duties to provide the information referred to in Articles 13 and 14, by means of an arrangement between them unless, and in so far as, the respective responsibilities of the controllers are determined by Union or

Member State law to which the controllers are subject. The arrangement may designate a contact point for data subjects' (Article 26(1) GDPR).[21]

Moreover, it is essential to take into account that to ensure data quality, all personal data processing shall comply with the set of principles provided for in Article 5 of the GDPR:[22]

**a) Lawfulness, fairness and transparency**

For the processing of personal data to be **lawful** in accordance with Article 6 of the GPDR it is necessary that the data controller identifies a legal basis, amongst which are the consent of the data subject, a contractual obligation, a legal obligation, a vital interest of the data subject, the public interest or duly justified legitimate interests. **Fairness** means that personal data should not be processed in a way that is unjustifiably detrimental, unlawfully discriminatory, unexpected, or misleading to the data subject. **Transparency** is about being clear and open with the data subject about how the data controller will collect, use and share personal data.

**b) Purpose limitation**

The design of the personal data processing should be shaped by what is necessary to achieve the purposes. The purposes must therefore be predetermined, specified, explicit and legitimate.

**c) Data minimisation**

Only personal data that is adequate, relevant, and limited to what is necessary for the limited purpose shall be processed. The controller shall delete or anonymize personal data as soon as identification is no longer needed.

**d) Accuracy**

Every reasonable step shall be taken to ensure that personal data that is inaccurate or outdated is erased or rectified without delay.

**e) Storage limitation**

The controller must ensure that personal data is kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data is processed.

**f) Integrity and confidentiality**

The principle of integrity and confidentiality includes protection against unauthorised or unlawful processing and against accidental loss, destruction, or damage, using appropriate technical or

[21] For more details on controller, processor and joint controllership, see EDPB Guidelines 07/2020 on the concepts of controller and processor in the GDPR Version 2.1, adopted on 07 July 2021, available at: https://www.edpb.europa.eu/system/files/2023-10/EDPB_guidelines_202007_controllerprocessor_final_en.pdf [last accessed: 14/06/2024].

[22] All descriptions are based on GDPR and EDPB, Guidelines 4/2019 on Article 25 - Data Protection by Design and by Default Version 2.0, adopted on 20 October 2020, available at: https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_en.pdf [last accessed 29/04/2024].

organisational measures. The security of personal data requires appropriate measures designed to prevent and manage data breach incidents; to guarantee the proper execution of data processing tasks, and compliance with the other principles; and to facilitate the effective exercise of individuals' rights.

**g) Accountability**

The data controller shall be responsible for and be able to demonstrate compliance with all the above-mentioned principles.

### 9.2.2 The legal processing of personal data

To ensure the proper implementation of such principles and compliance with ethics and data protection in general, each consortium partner needs to clarify *who is in charge of what*. This entails identifying **who is/are the data controller(s) and the data processor(s) for each activity**, and thus responsible in different ways for the proper management of the personal data to be collected. To facilitate the self-identification of such roles (which may vary for the different activities partners are involved in) and what obligations it entail, partners should follow the European Data Protection Board (EDPB)Guidelines 07/2020 on the concepts of controller and processor in the GDPR (Version 2.1).[23] A flowchart to assess what role corresponds to each partner is provided in Appendix E.

### 9.2.3 Data subject rights

Recital 59 of the GDPR mandates the establishment of mechanisms that allow data subjects to easily exercise their rights. These include the right to access, correct, or delete their personal data free of charge. Data controllers are required to enable these rights to be exercised electronically and to respond promptly to the data subject's requests.

Participants in the SUNRISE research activities will be informed of their right to withdraw from the study at any time without needing to provide a reason, and this right is maintained throughout the study. Should any research activities include audio recordings or electronic note-taking, participants will be informed that they may request the interviewer to pause or delete all or part of the recorded material at any moment. Participants also have the right to request the deletion of their data retrospectively. These rights and procedures are outlined during the informed consent process, as described above (9.1.1) and following the guidance provided in Appendix C and Appendix D. Participants may also verbalize their decision to withdraw.

### 9.2.4 Data sharing

Each partner providing shared information to another partner shall guarantee that: (i) they have the **authority** to disclose the information shared with the parties; (ii) where necessary and relevant, they have secured the appropriate **informed consents** from all individuals involved or any other relevant institution,

[23] EDPB, Guidelines 07/2020 on the concepts of controller and processor in the GDPR, Version 2.1., adopted on 07 July 2021. Available at: https://www.edpb.europa.eu/system/files/2023-10/EDPB_guidelines_202007_controllerprocessor_final_en.pdf [last accessed: 14/06/2024].

in compliance with applicable regulations; and (iii) **no restrictions exist** that would prevent any other partner from using the shared information.

### 9.2.5 Data Protection Impact Assessment

A Data Protection Impact Assessment (DPIA) is a process designed to describe the data-processing, assess its necessity and proportionality, and help manage the risks to the rights and freedoms of natural persons by determining the measures to address them. The DPIA is required when the processing is "likely to result in a high risk to the rights and freedoms of natural persons" (Article 25 GDPR). A DPIA for the cancer prevention programme operation in the different countries is foreseen in the context of Task 1.3 which will be informed by and assess the practices of each of the partners in their activities. If data controllers detect changes in the risk represented by the processing operations, they should review their compliance and if necessary consult with the QREM, the Project Coordinator or the Ethics Advisor Committee.

### 9.2.6 Technical and Organisational Measures

Following Art. 32(1) GDPR, the data controller and the data processor should implement appropriate technical and organisational measures (TOMs) to ensure a level of security appropriate to the risk, "[t]aking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons".

SUNRISE's partners are encouraged to implement the following TOMs, to the greatest possible extent:

a) Enhancing awareness and training among researchers on data protection and data security requirements and rules;
b) Using reliable information systems (hardware and software);
c) The pseudonymisation (when possible, anonymisation) and encryption of personal data;
d) Ensuring the ongoing confidentiality, integrity, availability and resilience of processing systems and services;
e) Restoring the availability and access to personal data in a timely manner in the event of a physical or technical incident;
f) Regularly testing, assessing and evaluating the effectiveness of technical and organisational measures for ensuring the security of the processing.

Subsequent versions of the DMP will specify the TOMs adopted by each partner in order to safeguard the rights of data subjects.

# 10 Other issues

### 10.1.1 Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

Currently, SUNRISE does not make use of procedures for data management other than those described in this data management plan.

# 11 Prevention & early detection (behavioural change) cluster

A common chapter on the commonalities among the projects taking part in the “Prevention & early detection (behavioural change)” cluster is to be developed in next DMP updates. The chapter will address commonalities between projects in data standards, data validation and data protection practices, as well as foster data exchange among the cluster participants, in particular when it comes to sharing project data with a future federated UNCAN.eu data platform.

# 12 Conclusion

In conclusion, this first version of SUNRISE's DMP lays the foundational framework for the project-specific data governance and ethics policy. It specifies the types of data involved in the project, alongside methodologies for their collection, processing, storage, sharing, and protection. The DMP also addresses the ethical and legal imperatives related to data management and consent procedures in research activities, equipping SUNRISE partners with the necessary tools to meet privacy and security standards.

Future iterations of this DMP will incorporate detailed enhancements based on inputs received through the Data Management and Ethics Management questionnaires (Appendix A and B). These versions will elaborate on the specific data management procedures and clarify further the nature of the data handled throughout the project and the measures taken to ensure the appropriate level of protection and compliance with ethical standards. The current DMP also includes guidance documents/templates for Data Protection Information Notices and Informed Consent, to be found in Appendix C and Appendix D, which partners can adapt for their respective activities.

SUNRISE is committed to continuously implementing comprehensive measures and guidelines that ensure fair and secure data management practices and uphold high ethical standards across all phases of the project. Responsible data management and the protection of every participant's rights are core elements of the SUNRISE project's success.

# References

Council for International Organizations of Medical Sciences (CIOMS). (2016). I*nternational Ethical Guidelines for Health-related Research Involving Humans*. Fourth Edition. Geneva. https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf

Creativecommons.org. (n.d.). *Creative Commons — Attribution 4.0 international — CC BY 4.0*. https://creativecommons.org/licenses/by/4.0/

Donnelly, M., Jones, S., & Pattenden-Fail, J. W. (2010). DMP online: A demonstration of the digital Curation centre's web-based tool for creating, maintaining and exporting data management plans. *Research and Advanced Technology for Digital Libraries*, 530-533. https://doi.org/10.1007/978-3-642-15464-5_74

European Commission, Open science in Horizon Europe, available at: https://rea.ec.europa.eu/open-science_en [last accessed: 14/06/2024].

European Commission. "Data Management Plan (HE):V1.1 – 01.04." EU Grants. Last modified April 1, 2022. Available at: https://ec.europa.eu/docs/temp-form/report [last accessed: 11/03/24].

European Commission. (n.d.). *Guidance for Applicants Informed Consent*. https://ec.europa.eu/research/participants/data/ref/fp7/89807/informed-consent_en.pdf.

European Commission. (n.d.). *Open science*. Research and innovation. https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en.

European Data Protection Board (EDPB). (2020). *Guidelines 4/2019 on Article 25 - Data Protection by Design and by Default Version 2.0*. https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_en.pdf.

European Data Protection Board (EDPB). (2021). *Guidelines 07/2020 on the concepts of controller and processor in the GDPR Version 2.1*. https://www.edpb.europa.eu/system/files/2023-10/EDPB_guidelines_202007_controllerprocessor_final_en.pdf.

European Data Protection Supervisor (EDPS). (n.d.). *Interoperability*. https://www.edps.europa.eu/data-protection/our-work/subjects/interoperability_en

European Data Protection Supervisor. (2021). Pseudonymous data: processing personal data while mitigating risks. Available at: https://www.edps.europa.eu/press-publications/press-news/blog/pseudonymous-data-processing-personal-data-while-mitigating_en [last accessed: 14/06/2024].

European Union Agency for Cybersecurity. (2017). Handbook on security of personal data processing. Publications Office. https://data.europa.eu/doi/10.2824/569768.

European Union Agency for Cybersecurity. (2019). Pseudonymisation techniques and best practices: Recommendations on shaping technology according to data protection and privacy provisions. Publications Office. https://data.europa.eu/doi/10.2824/247711.

European Union Agency for Cybersecurity. (2022). Deploying pseudonymisation techniques: The case of health sector. Publications Office. https://data.europa.eu/doi/10.2824/092874.

European Union Agency for Cybersecurity. (2023). Engineering personal data sharing: Emerging use cases and technologies : January 2023. Publications Office. https://data.europa.eu/doi/10.2824/36813.

GO FAIR. (2022). *FAIR principles*. https://www.go-fair.org/fair-principles/.

Intersoft Consulting. (2021). *Privacy by design*. https://gdpr-info.eu/issues/privacy-by-design/.

Kepper, M. M., Walsh-Bailey, C., Zhao, M., Parrish, L., Miller, Z. M., Glasgow, R. E., Fuentes, L. D., Yan, Y., Hayashi, R. J., Brownson, R. C., & Foraker, R. E. (2024). Satisfaction and effectiveness of a digital health tool to improve health behavior counseling among adolescent and young adult cancer survivors: A randomized controlled pilot trial. *BMC Digital Health, 2*(1). https://doi.org/10.1186/s44247-024-00064-1

Regulation 2854/2023. *Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act)*. https://eur-lex.europa.eu/eli/reg/2023/2854.

Regulation 868/2022. *Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act)*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians, 71*(3), 209-249. https://doi.org/10.3322/caac.21660.

Zenodo. (n.d.). *Principles*. https://about.zenodo.org/principles/.

Zerdick, T. (2021). *Pseudonymous data: Processing personal data while mitigating risks*. European Data Protection Supervisor. https://www.edps.europa.eu/press-publications/press-news/blog/pseudonymous-data-processing-personal-data-while-mitigating_e.

# Appendix A Data Management Questionnaire

**Instructions**

The following questionnaire is to be responded to **by all SUNRISE partners conducting research activities that entail the collection, processing and/or management of data (including personal and non-personal data).** All partners with a specific **data management role** assigned to them (as per DMP's Table) should complete the questionnaire as required and keep it updated as appropriate upon changes in circumstances or at the request of the Project Data Management Leader, the Technical Coordinator, the QREM or the Ethical Advisor Committee.

The questionnaire builds on the Horizon Europe Data Management Plan (HE) template and leverages GDPR requirements, FAIR principles, and complementary data protection and ethical requirements. **The questionnaire asks about all data collected, queried, and maintained during the lifetime of the SUNRISE project.** The answers will be used for the continuous update of the SUNRISE Data Management Plan. The survey consists of two parts, which you should complete depending on your role in data management:

**Part A: Data Summary (WP Level)** – please fill in Part A **ONLY if you are a WP leader**.

**Part B: Data Summary, FAIR Data, Data Security, Ethics and Data Protection (Activity Level)** – please fill in Part B if you **are leading any task/activity involving the collection or processing of data**.

If you hold **both roles**, **please complete both Part A and Part B**. As regards the Data Summary in Part A, please consider all the activities that fall under your WP. In Part B, the Data Summary shall be limited to the specific Activity you are responsible for.

The questions under the heading "**Ethical and Data Protection Aspects" should be answered in cooperation with your organisation's Data Protection Officer (DPO)**. If you are not sure what information to include, please reference to the European Commission's Ethics and data protection guidelines published on 05 July 2021: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-and-data-protection_he_en.pdf.

You may also take a look at the official EU Grants template for Data Management Plan (HE): V1.1 – 01.04.2022: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/temp-form/report/data-management-plan_he_en.docx

If you are still uncertain how to answer a question, please ask the Project Coordinator (CERTH) and/or Project Data Management Leader (PBY). Please save your responses in the dedicated folder within the Project's Google Drive environment (under WP 7).

**Part A – Data Summary (WP level)**

Please fill in the table below, keeping in mind **all the data types** that are to be generated, processed, or re-used as part of the activities in your WP. Please be as precise as possible and consult the responsible Activity specific Data Managers where necessary.

| Data Summary – WP level | | |
|---|---|---|
| **Type of data collected/generated** | ***Describe the type of data used or generated within the project.*** *Example: Audio, Video, Geospatial, Statistical, Image, Table, Web, Text, Data bases…* | |
| **Format and form of data collected/generated** | ***Specify the form and format of the data.*** *Examples of data formats: Text (Type) could be: XML, PDF, Doc, txt, HTML, JSON, RTF, and/or encrypted, compressed. Examples of data forms: basic demographics, physiological data, behavioural/social data (e.g. QoL questionnaires etc.).* | |
| **Data origin/provenance of data** | ***Define and describe the origin/source of your data. Data can be gathered from different sources.*** *Examples:*<br>*Observational: Data captured in real time - often not reproducible i.e. sensor readings, images.*<br>*Experimental: Data from lab equipment.*<br>*Derived/Compiled: Data coming from analysis or compilation* | |
| **Re-use of existing data** | ***Indicate****: No/Yes (if yes, then specify what kind)* | |
| **Expected size of data** | *n (MB, GB)* | |
| **Data utility** | ***Explain****: Specific purposes of the data collection/generation or analysis related to the objectives of the project, and possible beneficiaries outside the project* | |

| | | |
|---|---|---|
| **Related task(s)** | ***Indicate:*** *Tasks of WP involving data collection, processing and/or storage* | |
| **Personal data** | ***Indicate:*** *Yes/no (If yes, specify what kind)* | |
| **Responsible Data Manager(s)** | ***Indicate****: Partner(s) in charge of the data collection, processing, and storage in each of the WP activities, if applicable* | |
| **Data Security & Storage** | ***Explain****: Type of data storage, describe data security policies (for example, only authorized access will be permitted to sensitive data).* | |
| **Data value** | ***Explain****: Long and short term of data collection and generation benefits (for example, describe how the data collected and generated will provide short- and long-term benefits to specific target groups of users etc.).* | |

**<u>Part B</u> – Data Summary, FAIR Data, Data Security, Ethics and Data protection (Activity level)**

Please answer the following questions as precise as possible, keeping in mind **all the data types** that are to be generated, processed, or re-used as part of your Activity.

| **Data Summary – Activity level** | |
|---|---|
| What type of data will you produce, generate or collect during the Activity? *Please identify the respective Activity and task next to the type of data. Deliverables are not included in this question.* | |
| Will you re-use any existing data and what will you re-use it for? | |
| What is the origin/provenance of the data collected/generated? | |
| What is the expected size of the data that you intend to generate or re-use? | |
| What is the purpose of the data collection/generation and its relation to the objectives of the SUNRISE project (for each data type)? | |
| To whom might your data be useful ('data utility'), outside your activity and outside of the SUNRISE project? | |
| **Findable data** | |
| Will data be identified by a persistent identifier? | |
| What metadata will be created to allow for discovery? | |
| What disciplinary or general standards will be followed for documentation and metadata? | |

| | |
|---|---|
| Will search keywords be provided in the metadata to optimise the possibility for discovery and then potential re-use? | |
| Will metadata be offered in such a way that it can be harvested and indexed? | |
| **Interoperable data** | |
| Which data produced and/or used in the activity will be made openly available? | |
| How will the data be made accessible (e.g. by deposition in a trusted repository)? | |
| What methods or software tools are needed to access the data? | |
| Where will the data and associated metadata, documentation and code be deposited? | |
| If there are restrictions on use, how will access be provided, both during and after the end of the project? | |
| How will the identity of the person accessing the data be ascertained? | |
| **Reusable data** | |
| How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)? | |
| Will your data be made freely available in the public domain to permit the widest re-use possible? | |
| Will the data produced in the activity be useable by third parties, in particular after the end of the project? | |

| | |
|---|---|
| Will the provenance of the data be thoroughly documented using the appropriate standards? | |
| Describe all relevant data quality assurance processes. | |
| Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement? | |
| **Data Security** | |
| What are the major risks to data security? | |
| What steps will be taken to protect privacy, security, confidentiality, intellectual property or other rights? | |
| What other security measures do you anticipate being required for safe data storage and management? | |
| What, if any, data recovery protocols have you developed? | |
| What, if any, protocols have you developed or outlined for the safe transfer of personal data? | |
| Who decides what data or what categories of data will be kept and for how long and on what basis? How long should it be retained (e.g., 3-5 years, 10-20 years, permanently)? | |
| Where will such data be maintained? Are there data archives that are appropriate for your data (subject-based or institutional)? | |
| The GDPR requires personal data not be kept longer than necessary for the purpose for which it was stored. What protocol(s) will you put in place to ensure you delete personal data that is no longer required to be stored? | |

| Ethics and data protection | |
|---|---|
| What types of personal data do you intend to collect, generate or process? | |
| What types of sensitive data do you intend to collect, generate or process? | |
| Will any of the data subjects be children or vulnerable people? | |
| If the data is personal data, as defined by the GDPR, which of the six Art. 6 legal bases will you rely on for the processing of each category of personal data? https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679#d1e1888-1-1 | |
| If the data is sensitive data, as defined by the GDPR, which of the ten Art. 9 legal bases will you rely on for the processing of each category of sensitive data? https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679#d1e2051-1-1 | |
| Will any entity (including any service provider) outside of the E.U. have access to personal or sensitive data? If so, who? For what purpose? | |
| Do you plan to use pseudonymisation and/or anonymization for any data category? If so, how do you plan to implement this? | |

# Appendix B Ethics management questionnaire

The current questionnaire is **mandatory** to be filled out ***by the partners leading any specific Activities carried out within the SUNRISE project which collects external participants' data***. Please attach any required forms either to this document or as a separate attachment in an email to the corresponding Work Package leader, along with the returned completed questionnaire. SUNRISE Work Package leaders shall store all questionnaires and related documents in the dedicated folder within the Project's Google Drive environment (under WP 7).

The European Commission has identified that for all activities funded by the European Union, ethics is an integral part of research from beginning to end, and ethical compliance is pivotal to achieving real research excellence. Regardless of the discipline, research involving human participants, and especially such involving personal data processing, requires an ethical review assessment and an ethical approval. To comply with these obligations, researchers and organization representatives need to apply for ethical approval at their local Ethics Committees (i.e., University, Research Institute, Hospital, etc.) and follow their local guidance and procedures for the application.

The **Ethics and Data Protection Decision Tree of the European Commission** can help you navigate through the process of Ethics Management and inform the following:

1. Support the identification of potential ethics risks related to the data processing activities of your project;
2. Facilitate compliance with the data ethics requirements aimed at safeguarding the fundamental human rights and freedoms of the research participants;
3. Foster the application of the "ethics by design" principles.

## General

1. **What is your procedure for the identification of participants?** *(i.e., existing contact lists, social media, general public, academics, etc.)*

2. **What is your procedure for the recruitment of participants?**

3. **Do you have an informed consent procedure in place?**

   ☐ Yes, (Please describe it below and attach a copy to the questionnaire of the consent form used in the context of your study)

☐ No (A template will be provided in D7.5. Please utilise it and tailor it to the needs of your project/activity)

1. **Do you need to submit an EA?** [24]

☐ Yes

☐ No (Please describe below why, what is your alternative strategy, and what kind of EA do you need)

*If you have answered "YES" to the question above:*

2. **What type of EA do you have/need?**

3. **Please explain the procedure you have followed in order to obtain EA.**

**Please attach the following documents, if applicable:**

[24] Each responsible researcher should assess whether Ethical Approval is needed for the activities she/he is carrying out, based on local regulations and international guidelines, and always taking Article 19 of Regulation (EU) 2021/695 into account.

a) **Main document** (i.e. summary of the project/activity, study design/protocol, methodology & data analysis, recruitment, etc.)
b) **Participant information sheet/information letter**
c) **Informed (parental) consent form**
d) **Ethical Approval** (if already obtained. If not, please indicate the submission date and the foreseen approval date)

# Appendix C Data protection and information notice

## Introduction

You are invited to participate in [ACTIVITY] in the context of the project "SUNRISE - SUstaiNable inteRventions and healthy behavIours for adoleScent primary prEvention of cancer with digital tools".[25]

The [ACTIVITY] will be conducted by [a researcher] belonging to [PARTNER NAME], which is a member of the SUNRISE consortium and will be responsible for the collection, processing and storing of personal information that you will provide in the context of the [ACTIVITY], in compliance with the General Data Protection Regulation (GDPR).

This **Data Protection and Information Notice** describes the personal data that will be collected in the context of such research activities and the measures taken to protect it. As a participant in the "SUNRISE" project, you shall be informed about the processing activities relating to your personal data and the rights you have as a data subject.

Taking part in this research is **voluntary** and participants **can stop at any moment** in time without providing any reason for stopping. Before you consent to participate, please read carefully this Data Protection and Information Notice and learn why the research is being done and what it will involve. Please do not hesitate to ask questions and make sure that you have a complete understanding of the activities you are participating in.

## Brief description of the project and activity

The overall aim of SUNRISE is to co-create, implement and evaluate an innovative digitally-enhanced life-skills programme for primary prevention of cancer through health behaviour change in adolescents. To tackle the health and societal challenge of primary prevention of cancer in Europe, SUNRISE will combine an established, evidence-based digital solution, with novel intervention approaches such as peer social media campaigns, advertising literacy training, educational games, and social robot platforms, to take cancer prevention approaches for adolescents in the EU to the next level. The digitally enhanced programme will be implemented and evaluated at a large scale across 154 schools and 7500 students in urban and rural regions of 8 European countries, including socially disadvantaged groups and ethnic minorities.

The objective of this [SPECIFIC ACTIVITY] is to [INSERT OBJECTIVES OF THE SPECIFIC TASK]. Personal data or data relatable to a specific individual is needed to conduct [ACTIVITY] in order to [PURPOSES OF PERSONAL DATA COLLECTION]. To this end, we will collect and process personal data as indicated below (see Data processor, data collection and storage). The information gathered will be aggregated and anonymised for the completion of the above-mentioned task.

[ELABORATE ON HOW TO CONDUCT THE ACTIVITY]

[25] https://cordis.europa.eu/project/id/101136829.

### Duration

[INSERT DURATION OF THE ACTIVITY]

### Data processor, data collection and storage

As these activities involve the processing of personal data, they are subject to data protection and processing rules as established by Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). All processing will be based on the principles of lawfulness, fairness and transparency.

The Data Processor will be [ORGANISATION AND CONTACT INFORMATION]

The Data Controller will be [ORGANISATION AND CONTACT INFORMATION]

The purposes of this processing are [ADAPT TO THE SPECIFIC ACTIVITY PURPOSES]:

- *[EXAMPLES]*
- *To register participants of events and meeting*
- *To set up networking activities*
- *To inform of follow-up activities related to the project*
- *Dissemination purposes*
- *Etc.*

The following of your personal data are collected: [ADAPT TO THE REQUIRED DATA] e.g.:

- *[EXAMPLES]*
- *First name;*
- *Last name;*
- *Title;*
- *Function;*
- *Professional e-mail address;*
- *Image/voice (upon explicit consent)*
- *Feedback on the online interviews/surveys*
- *Other personal data possibly contained.*

The following third-party tools might be used [IF APPLICABLE]:

[INDICATE WHICH TOOLS, FOR WHAT PURPOSES, REFER DATA PRIVACY POLICY LINK]

### Recording [if applicable]

[*Please describe here, whether the participation will be recorded, for which purposes, where and to whom will the recording be accessible, and for how long*]. Recordings and photos can only be processed based on your explicit prior consent. Please note that the [ACTIVITY] might be recorded and that your voice/image data might be collected. Upon registration, you can give your explicit consent to have your image/voice recorded. In the absence of your consent, the organiser will try to find suitable alternatives, so that you can fully take part in the [ACTIVITY].

### Data storage

The data collected will be anonymously stored by [NAME ORGANISATION], the Data Processor, in its servers, with the possibility of sharing the anonymised data within the SUNRISE Consortium, strictly for SUNRISE project purposes and only to the necessary extent to fulfil the project's objectives. The data will be stored for up to 5 years after the completion of the project for potential follow-up purposes unless the participant requests to exercise their rights of erasure or retrieval before the finalisation of this period.

### Withdrawal of consent

If you want us to delete your data, please contact us and we will deal with your request without undue delay. Please note that withdrawing your consent does not affect the lawfulness of any processing based on your consent before this consent is withdrawn. Attention is drawn to the consequences of a delete request, which means that all your contact details will be lost.

### Personal rights and legal ground

You have the right to access your personal data and to request your personal data to be rectified, if the data is inaccurate or incomplete; where applicable, you have the right to request a restriction or to object to processing, to request a copy or erasure of your personal data held by the data controller. You have the right to data portability, meaning you may request us to transfer to you any identifiable information we hold about you in a machine-readable format. If processing is based on your consent, you have the right to withdraw your consent at any time, without affecting the lawfulness of the processing based on your consent before its withdrawal. If you have **any queries** concerning the processing of your personal data or wish to exercise any of the rights described above, you can contact [INSERT E-MAIL OF RESEARCHER], responsible for this study.

### Benefits and potential risks

There are no foreseeable risks or benefits associated with your participation. In addition, we do not envisage any potential risk for your rights and freedoms to be caused by your participation. [IF APPLICABLE, DESCRIBE POTENTIAL BENEFITS AND RISKS].

### Reimbursement

There is no reimbursement schema.

### Data Protection Officer (DPO)

The DPO of our institution is: [NAME]

For further questions, please do not hesitate to contact him/her under the following email address/contact number: [CONTACT DETAILS OF THE ORGANIZATION'S DPO]

# Appendix D Informed consent sheet

I,_____(full name), understood the information that was provided to me – in writing and verbally – about the participation in the [ACTIVITY] being conducted by [NAME ORGANISATION] in the context of the SUNRISE project, to which I am being invited to take part of.

| Please place an "X" in the box on the right to consent to the following statements: | | |
|---|---|---|
| 1) | I confirm that I have read and understood both this form and the accompanying Data Protection and Information Notice. I had the time and opportunity to ask questions as needed. | |
| 2) | I understand that I am free to withdraw my consent at any time without giving reason and that my participation in this project is voluntary. | |
| 3) | I am aware that my participation in the [ACTIVITY] will not have any costs for me. | |
| 4) | I am aware of the potential risks and benefits of this research study. | |
| 5) | I consent to participate in the research activities as described in the Data Protection and Information Notice having been fully informed of the potential risks, benefits and alternatives of the research [study/activity]:<br>A) [INSERT RESEARCH ACTIVITY 1, 2, etc.] | |
| 6) | I consent to the processing of my data. My personal data can be gathered to be used, stored, and shared in the ways described on the accompanying Data Protection and Information Notice. The personal data collected will be processed following the GDPR and will be pseudonymised/anonymised to the greatest extent possible. | |
| 7) | I consent to the use of *audio/video recording* for the exclusive purposes indicated in the Data Protection and Information Notice. | |
| 8) | I understand my right to request access to any, and all, personal information that I have voluntarily provided as part of my participation, and that I may ask for that information to be rectified and/or amended if it is inaccurate, or request that all personal information that I have provided be deleted (if not yet already anonymised). | |
| 9) | I understand that the SUNRISE consortium intends on retaining my personal details for a period of up to 5 years following the completion of the project. Information from research activities will be permanently and irrevocably deleted after a maximum of 5 years after the end of the project. | |
| 10) | I would like to receive updates on the progress and findings of the project. | |
| 11) | I have been offered a copy of this consent form. | |

As such, **I freely agree to take part in this study, in the terms selected in this form and as described in the Information Notice attached that was presented to me by the researcher.**

My Signature: _____

Date: __ / __ / 2024

**Statement by the Researcher taking consent**

I have accurately provided the information sheet to the participant and, to the best of my ability, made sure that the participant understood it. I confirm that the participant was given an opportunity to ask and get answers to questions about SUNRISE and the research activities. I confirm that the participant has given consent freely and voluntarily.

**Name of the researcher**: XX

**Organisation**: XX

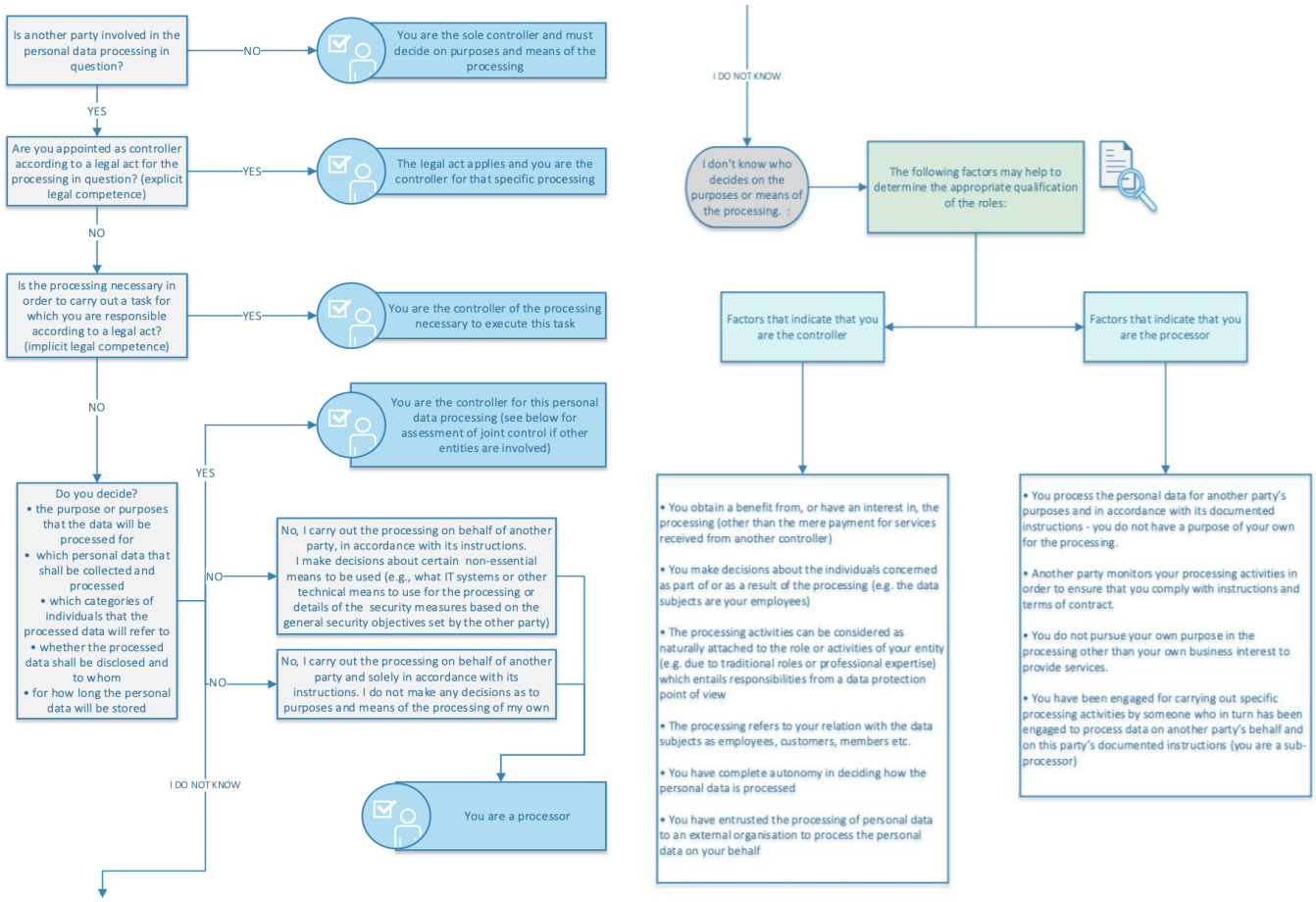
**Email**: XX

Researcher's signature: _____

Date: __ / __ / 2024

# Appendix E Data Controller and Data Processor identification flowchart and Responsibilities checklist

Figure 1 - Flowchart for identifying Data Controller and Data Processor Roles



Source: EDPB Guidelines 2020/07, Appendix E – Flowchart for applying the concepts of controller, processor and joint controllers in practice.

Table 11 – Data Controller and Data Processor responsibilities checklist

| Data Controller or Joint Controller's responsibilities checklist | Data Processor's responsibilities checklist |
|---|---|
| • Complying with data protection principles under art. 5 GDPR<br>• Upholding individuals' data protection rights<br>• Keeping records of processing operations<br>• Ensuring the security of processing<br>• Choosing an appropriate data processor<br>• Detailing in a binding contract the controller-processor relationship<br>• Notifying personal data breaches to the relevant EEA data protection authority and to individuals, where applicable<br>• Being accountable for the processing operations, practising data protection by design & default, carrying out data protection impact assessments when necessary<br>• Appointing a data protection officer when necessary<br>• Complying with the data protection obligations on international transfers of personal data<br>• Cooperating with data protection authorities | • Following the controller's instructions<br>• Keeping records of processing operations<br>• Ensuring the security of processing<br>• Respecting and upholding the binding controller-processor contract<br>• Obtain the authorisation of the controller before engaging a new sub-processor (and give the controller a possibility to object). If applicable, a processor - sub-processor contract must be put in place and equate to the initial contractor- processor contract<br>• Notifying personal data breaches to data controller<br>• Notifying GDPR breaches to the controller<br>• Being accountable for the processing operations: e.g. practising data protection by design & default<br>• Appointing a data protection officer when necessary<br>• Ensuring that international transfers are authorised by the controller and comply with the GDPR<br>• Cooperating with data protection authorities |

Source: EDPB.europa.eu