SCEWERO

Funded by the European Union

SCEWERO

STRENGHTENING THE RESEARCH CAPACITIES FOR EXTREME WEATHER EVENTS IN ROMANIA

GA 101159497

DELIVERABLE D2.3: DATA MANAGEMENT PLAN (UPDATE 1)

2026

UNIVERSITATEA BABEȘ-BOLYAI

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici

Universiteit Antwerpen

indeco soft

PROJECT INFORMATION	
Programme Call:	HORIZON.4.1 - Widening participation and spreading excellence HORIZON-WIDERA-2023-ACCESS-02
Grant agreement	101159497
Project title	STRENGTHENING THE RESEARCH CAPACITIES FOR EXTREME WEATHER EVENTS IN ROMANIA
Project acronym	SCEWERO
Partners	UBB, CMCC, UANTWERPEN, INDECO
Duration	36 months
Type of action	Coordination and Support Action (HORIZON-CSA)
Work Package	2. Project Management
Deliverable no.	2.3
Deliverable type	Data Management Plan
Version	Second version
Contractual date of delivery	28/02/2026
Title of document	Data Management Plan (Update 1)
Actual date of delivery	28/02/2026
Responsible partner	UBB
Author(s)	Adina-Eliza CROITORU, Sorin POP, Csaba HORVATH
Content of this document	The document represents the second version of the Data Management Plan for the SCEWERO project. The structure follows recommendations provided in the guiding documents Horizon Europe Data Management Plan Template (Version 1.0, 05 May 2021) for Horizon projects of the European Commission. It contains a data summary and detailed information related to FAIR data, other research outputs, allocation of resources, data security, ethics and other data-related issues.
Availability	Public

Dissemination level		
PU	Public	X
CO	Confidential, only for members of the consortium (including the Commission Services)	
CI	Classified, as referred to in Commission Decision 2001/844/EC	

DOCUMENT AUTHORS		
Author	**Content**	**Date**
Sorin POP **Adina-Eliza CROITORU**	First draft for internal review	07/02/2026
Csaba HORVATH **Sorin POP** **Adina-Eliza CROITORU**	Revision document	26/02/2026
Adina-Eliza CROITORU	Final document (Update 1)	28/02/2026

Contents

Disclaimer

The content of this deliverable does not reflect the official opinion of the European Union. Responsibility for the information and views expressed herein lies entirely with the author(s).

Statement of originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and the work of others has been made through appropriate citation, quotation or both.

The document was developed based on the *Horizon Europe Data Management Plan Template* (Version 1.0, 05 May 2021) provided by the European Commission as a guiding document for Horizon projects.

HISTORY OF CHANGES		
VERSION	PUBLICATION DATE	CHANGE
1.0	31.03.2025	Initial version: Data Management Plan (v1.0).
2.0	28.02.2026	Second version: Data Management Plan (Update 1).

How to quote this document

CROITORU AE; POP S; HORVATH Cs; (2026), Data Management Plan (Update 1), SCEWERO-101159497 project

List of abbreviations

CA – Consortium Agreement

CO – Project Coordinator

CSR – Case Study Region

D – Deliverable

DMP – Data Management Plan

ECAD – European Climate Assessment and Datasets

EFAS – European Flood Awareness System

EDO – European Drought Observatory

EWS – Early Warning System

FAIR - Findability, Accessibility, Interoperability, and Reusability

GA – Grant Agreement

GDPR – General Data Protection Regulation

GIES/IGSU – General Inspectorate for Emergency Situations

HE – Horizon Europe

INS – National Institute of Statistics in Romania

INMSS – National Institute for Health Services Management

LST – Land Surface Temperature

M – Month

NMA – National Meteorological Administration in Romania

OA – Open Access

PI – Principal Investigators

P1/P2/P3 – Partner 1/Partner 2/Partner 3 in the project consortium

RCM – Regional Climate Model

WP – Work Packages

1. Data Management Plan and Updates

According to the SCEWERO project proposal and the GA, the Data Management Plan (DMP) will be periodically updated as relevant additional information becomes available during project implementation and as activities progress. The DMP is planned to be a 'living' document, updated with new information to reflect the current state of the data. The final version will be delivered at the end of the project.

The first version of the DMP was prepared in M6 (D2.1 - Data Management Plan), and it was updated in M17 (D2.3 - Data Management Plan - Interim) and delivered as a final version in M36 (D3.1 - Data Management Plan - Final). The final version will outline the procedure for collecting, storing, evaluating, and utilising the data generated by the project upon its completion.

Each version of the DMP will reflect the project's data status. The generation of new data, not foreseen in the project proposal or in previous versions of the DMP, will be presented in the updated DMP.

The project coordinator (CO), UBB, is responsible for writing and updating all versions of the data management plan throughout the project implementation period, but input from all partners is requested and integrated into each version.

2. Data Summary

During the project lifetime, the consortium partners will decide whether to include any new robust datasets for reuse, if needed. The data flow chain is synthetically presented in Figure 1.

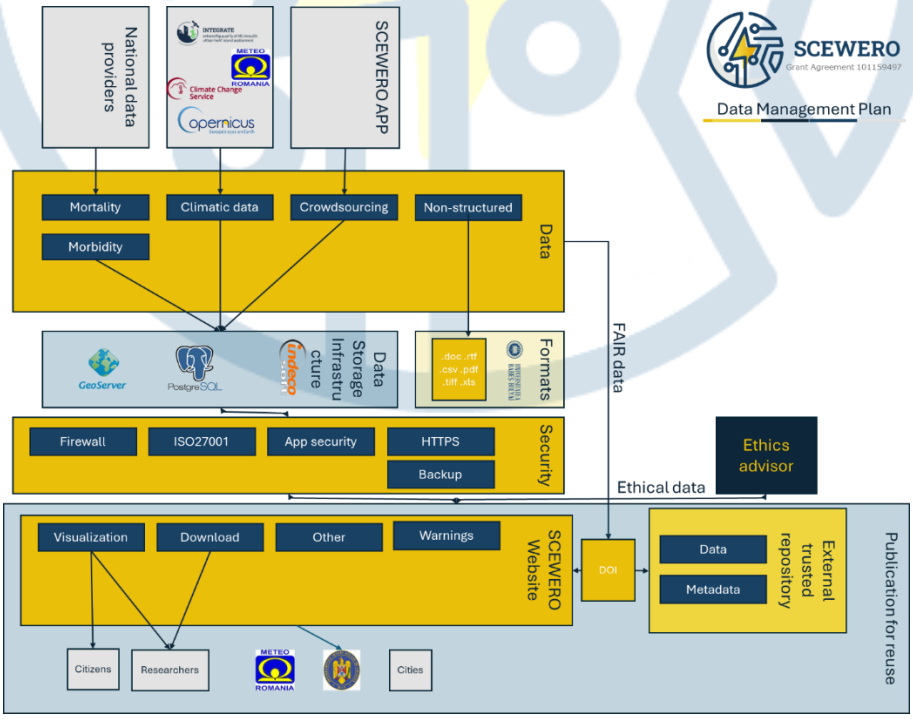

Figure 1. SCEWERO data flow chain

2.1. Types and Formats of Data Generated and Reused in the SCEWERO Project

2.1.1. Types of data reused

Several types of data have been planned to be reused to achieve the SCEWERO project objectives:

- ***Weather, climate and hydrological data***

SCEWERO will re-use existing data such as historical weather, climate and hydrological data, climate projection data derived from RCMs, remote sensing data (LST), health data, available from open sources and from national data providers for practical applications during training sessions (for CO researchers) and summer school (for national and international students) (WPs 4 and 5), as well as for the research activities to be developed during the project implementation (WP7).

For the evaluation of heatwave, hydrological extremes (floods and drought) and compound events over the past, we mainly build on existing gridded datasets such as ERA5 reanalysis (25 km as horizontal resolution) for the historical period, and data from the Coupled Model Intercomparison Project Phase 6 (CMIP6) global climate models and the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) global hydrological models for future projections. ERA5 (Hersbach et al., 2020) was produced by ECMWF within the Copernicus Climate Change Service (C3S); additional gridded datasets for comparison, such as MSWEP (Beck et al., 2019) for precipitation, will be used. The detection methods and analyses to be presented during training sessions and summer schools are primarily derived from the ERA5 reanalysis.

To assess the impact of heat events on human health, historical weather observation data, freely available from the NMA or international databases (ECA&D, Meteomanz, Oggimet, Allmetsat) for locations within the CSRs, will be used.

For the hydrological extremes (floods and drought) training and session in summer school, as well as for publication related to hydrological extremes, publicly available datasets and resources for research and educational purposes, are used: European Meteorological Observations (EMO-1) (Thiemig et al., 2022), the Global Runoff Data Centre (GRDC), the high-resolution pan-European hydrological reanalysis (HERA), the European Commission, Joint Research Centre (EC-JRC) LISFLOOD, Global Flood Awareness System (EFAS or GloFAS) and various drought indicators datasets in the European Drought Observatory (EDO).

To make it usable, the collected data from various sources were/will be homogenised to achieve common/compatible file formats and standard spatiotemporal resolutions, put in ready-to-use formats (e.g., raster-based) to be further shared within the consortium partners to be used during the project implementation period and beyond for both training sessions and summer schools to allow hands-on activities and for research activities.

- ***Remote sensing data***

LST data derived from Landsat missions and available from the United States Geological Survey (www.usgs.gov) will be reused to integrate with air temperature, humidity, and wind data for advanced analysis of heat stress in the CSAs in the WP dedicated to research.

- ***Health microdata***

Health data to be reused includes anonymised datasets on daily mortality and emergency morbidity data provided by national data providers in Romania: INS (https://insse.ro/cms/) and INMSS (https://inmss.ro/en/). They will be employed only for research activities and will be required from the national data providers on a request basis.

Mortality microdata provided by the INS covers the reference period 1995-2024 and includes the following information:

- Home city (as per ID)
- County
- Date of death
- Gender
- Age or year of birth
- Diagnosis according to ICD-11 standard
- Cause of death (ICD-11)
- Secondary causes of death, if available
- Metadata

According to the data-providing contract signed between the INS and UBB, all data received and all confidential data derived from them will not be shared. The raw data received will be destroyed by the UBB after project completion. Research results containing only non-confidential data will be published through the project repositories. Also, according to the contractual conditions, access to and data raw data processing are limited to the UBB researchers nominated in the contract who have signed NDAs as Annexes to the data-providing contract.

The medical emergency interventions data provided by the INMSS include:

- Date of emergency
- Home city of the patient (as per ID)
- County of the patient
- Gender
- Age or year of birth
- Diagnosis according to the ICD-11 standard
- Cause of death (ICD-11)
- Secondary causes of death, if available
- Comorbidities
- Metadata

According to the contractual conditions, the data provided by the INMSS cannot be published or further shared, except for aggregate reports not directly linked to the data.

➢ ***Other data***

During the project implementation period, various data for project management are developed, including partner information, meeting agendas and minutes, project reports, conference or other scientific event presentations, tutorials, specification documents, videos, and photos from project activities and meetings.

2.1.2. Types of Data Generated

(1) ***Research-derived data***: data obtained from the research activity will include four types of datasets (WP7):

- *(Bio)Climatic data:* consist of datasets for heat-related events and thermal stress with specific indicators covering the historical period and the following decades; they will include various indicators (e.g., frequency, intensity, duration, cumulated duration etc.),
- *Thermal comfort/stress perception data* (crowdsourced data) collected through the mobile app from the users.
- *Early warning system and heat warnings perception* data collected through a sociological survey.
- *Impact data* will include the relative risk of mortality and morbidity for specific weather conditions.
- *Critical threshold data for health impact* will contain a set of weather variable combinations that lead to thermal discomfort and increased morbidity and mortality.

(2) ***Software development***:

- A *mobile application* for collecting citizen perception on thermal comfort/stress;
- *AI models* for the heat/weather conditions on health;
- A *platform* to integrate physical science data, health data, and socio-demographic data to provide early warnings for heat events.

(3) ***Stakeholder-collected information*** will become available after organizing the workshop for stakeholders (public authorities and health system employees) and the webinars for media. Their requirements and priorities will be analysed and integrated into the impact scenarios for enhancing the EWS in Romania.

(4) ***Personal data of participants*** to various events organised: summer schools, workshops, trainings, and scientific events organised during the project will generate derived datasets, including names, contact details (including email addresses), organisational affiliation, education background, personal identifiers, and financial information (e.g., bank account details). They are collected and processed in line with EU regulations. This data are used exclusively for project-related purposes, such as issuing invitations, participation certificates, processing payments, and fulfilling other administrative or legal obligations. Access to this data is strictly limited to authorised consortium members and will not be disclosed to third parties without a valid legal basis.

(5) **Specific research results derived from hands-on activities** during summer schools and training activities. The activities during the training sessions and summer schools are not expected to produce any new datasets, but the results obtained can be further used for scientific publications, theses or other professional reports of the participants in the training and summer school events.

All personal data collected as presented in (1), (3), and (4) will be processed in compliance with the GDPR. Participants are fully informed about the nature and purpose of data collection, and explicit, informed consent will be obtained prior to their participation in any project-related activities. Although the project does not directly collect sensitive data from participants, it processes anonymised health-related data provided by national authorities, which falls under Article 9 GDPR (health data). These datasets are handled under strict confidentiality agreements and processed exclusively for scientific research purposes under Article 9(2)(j) GDPR. A clear and transparent consent form will be provided to participants at the initial point of contact and must be signed before involvement in any project activities. Data subjects are also informed of their rights under the GDPR, including the right to access, rectify, or erase their data, as well as to withdraw consent at any time.

2.1.3. Format and Size of Data

SCEWERO uses various digital data formats that include, but are not limited to:

- Multidimensional datasets (*.nc; *NetCDF);
- Compressed files (*.rar, *.zip);
- Images: *.jpg and *.png;
- Model scripts and data processing codes, in open-source coding languages (e.g. Python, R), version control software (Git), and trusted repositories (e.g., GitHub.com, Zenodo);
- Multimedia files for video recordings – interviews for dissemination and communication, as well as for online webinars and workshops: *.mp4;
- Spatial data: vector data in ESRI shapefile/geodatabase format; raster data in GeoTIFF format;
- Tabular data: ASCII format (*.txt, *.csv), MS Excel (*.xlsx), Postgres databases;
- Text data/information (for project reports, newsletters, articles, tutorial for the mobile app, presentations, and other documents, containing text): Rich Text Format (*.rtf), Text (*.txt), PDF (*.pdf), MS Word (*.docx), MS PowerPoint (*.pptx).

The expected size of SCEWERO datasets reused and generated is in the range of several gigabytes (GB), including all types of data needed for or derived from project implementation.

2.2. Purpose of Data Generation or Reuse in Relation to the Project Objectives

For the training activities for CO's researchers and summer schools organised during the project implementation, we reuse climatic and hydrological data to allow participants to practice during hands-on activities and sessions. They are associated with the project SOs to raise the research capacity and scientific reputation of the CO's researchers and to increase CO's international visibility through organising summer schools. We consider reusing open data from external resources such as international databases (e.g., Copernicus, ECA&D project, EFAS, EDO, Meteomanz, Oggimet, Allmetsat).

For research activities, the data reuse aims to support advanced statistical analysis of heat-related events and calculate the relative risk of mortality and morbidity associated with heat events in the two CSRs in Romania.

During the research activity, we collect crowdsourced data via a mobile application developed in the project (WP7) and through a sociological survey, which includes user demographics (gender, age), location for each dataset, and the thermal perception and the heat EWS opinion of the user at the reporting moment/location. We will also produce datasets regarding heat event features (frequency, intensity, duration, etc.) and their related health impacts (e.g., on mortality and morbidity), as well as a set of critical thresholds for the occurrence of these impacts.

To ensure a holistic approach, the collected health and perception data will be integrated with official, location-specific meteorological records available from the NMA, Meteomanz, Oggimet, etc.

During research activity (WP7), advanced statistical techniques (regression analysis and clustering) combined with AI algorithms will be employed to extract meaningful insights from the datasets and to ascertain associations between recorded temperatures and other weather variables that contribute to thermal stress and resultant health outcomes, based on daily data.

Potential correlations and interactions among user perceptions, meteorological data, and health in the Romanian context, particularly regarding the simultaneity of available data, will be further explored. Regarding extreme heat conditions, in order to better characterise heat stress (human perception of stress), we plan to build on composite indices that also consider humidity and wind speed alongside temperature. Moreover, it will allow the identification of a new set of weather condition thresholds for the two CSRs and the vulnerable groups, enhancing the reliability and applicability of the findings.

2.3. Origin/Provenance of the Generated or Reused Data

(1) ***Origin of the generated datasets***:

- *(Bio)Climatic data* consisting of datasets for heat-related events and thermal stress with specific indicators (e.g., frequency, intensity, duration, cumulated duration, etc.) will be derived from historical and RCM output data (ERA5 and CMI).
- *Thermal comfort/stress perception* data will be collected from citizens in Romania through the mobile app and through the sociological survey.
- *Health impact data (*relative risk of mortality and morbidity) will be derived through processing historical observation and gridded weather/climate data and health data.
- *Critical threshold data for health impact* will be derived based on weather and health impact data.
- S*takeholder-collected information* will be collected during the workshop from the stakeholders (public authorities, health system employees) and from journalists during the webinars for media.

(2) ***Reused data*** are provided by the national data providers in Romania or from online available authorised sources, as follows:

- *Demographic and daily mortality data*: INS - not shareable, to be destroyed after project completion.
- *Daily morbidity data:* INMSS - not shareable.
- *LST* – United States Geological Survey (www.usgs.gov)
- *Surface Urban Heat Islands* and associated hot/cool spots – P4 and CO, from INTEGRATE project (www.integrate.indecosoft.ro)
- *Weather/climate observation* data:
    - NMA (www.meteoromania.ro): via DATA.GOV.RO portal (https://data.gov.ro/) or on a request basis;
    - Meteomanz (www.meteomanz.com) – available online;
    - Oggimet (www.oggimet.com) – available online;
    - Allmetsat (https://ro.allmetsat.com/metar-taf/romania-bulgaria.php) – available online;
- *Gridded climate historical data* (1 km spatial resolution) – NMA via DATA.GOV.RO portal (https://data.gov.ro/) and from ERA5 (Hersbach et al. 2020) dataset for the evaluation of past heatwaves and hydrological extremes;
- *Projected climate data:* we build on the 6th Coupled Model Intercomparison Project (CMIP6) data to derive future projections on the long term and on the C3S Copernicus seasonal forecast dataset for the comparison of dynamical seasonal forecast results to the Machine Learning based approach we'll propose during the training sessions;
- *Hydrological data*:

- European Meteorological Observations (EMO-1) (Thiemig et al., 2022): this dataset consists of daily spatiotemporal meteorological data at a 1-arc-minute resolution; the data set is freely available;
- The Global Runoff Data Centre (GRDC): the centre provides time series of in situ river discharge data shared by National Hydrological Services;
- The high-resolution pan-European hydrological reanalysis (HERA) (Tilloy et al., 2025): the dataset is the result of running the LISFLOOD hydrological model at a 1-arc-minute resolution with a 6-hourly time step, forced with downscaled and bias-corrected reanalysis ERA5-land;
- The European Commission, Joint Research Centre (EC-JRC) LISFLOOD: this organisation provides open access to tools for hydrological modelling, the Global Flood Awareness System (EFAS or GloFAS) and various drought indicators datasets in the European Drought Observatory (EDO).

- Apart from the data mentioned in the previous section, projections of socioeconomic factors (e.g., population, GDP) in CSV and NetCDF formats form the backbone of the risk analysis presentation on hydrological extremes during summer schools.

The sources of *Meteomanz* and *Oggimet* data are original SYNOP messages collected from weather stations in Romania; the source of *Allmetsat* data is the original METAR messages issued by Romanian Airports.

Datasets to be reused will combine fine-scale data and information from physical sciences (e.g., weather, climate and hydrology datasets) with data derived from social sciences (e.g., demography, health, and communication sciences) to obtain health impact data.

Outside the SCEWERO project, the data will be useful to several key stakeholders in the Romanian Emergency Management context, to professionals and researchers in the health sector, to the local authorities in the two CSRs, and to the NMA and the GIES/IGSU to improve the national Early Warning System in Romania for heat-related events at the country level.

3. FAIR Data

The data used in trainings or research activities of the SCEWERO project will follow the FAIR principles (Findability, Accessibility, Interoperability, and Reusability), EU standards and regulations.

3.1. Making Data Findable, Including Provisions for Metadata

Since the persistent identifiers will allow efficient and rapid identification of data locations and access to the data by researchers, practitioners, professionals, and stakeholders, as well as students, in a very efficient manner, all newly created datasets will be registered for DOI. As our first option is to share the SCEWERO-produced data through the *Zenodo* repository, all datasets uploaded will be registered for a DOI.

Metadata submission is a requisite for all research data, whether newly generated or reused and must be recorded at the time of dataset creation or collection. The metadata will be aligned with established standards from data archive centres (e.g., INSPIRE, Pangaea), stored in a centralised metadata catalogue, and made open and accessible to all targeted audiences (researchers, students, professionals, stakeholders, etc.). This approach enhances the project's impact and

ensures research transparency. Datasets metadata will be freely available under a Creative Commons Public Domain Dedication (CC0) or an equivalent framework. In accordance with the FAIR principles, particularly ensuring machine-actionability, the metadata will include, at a minimum, the following information: dataset description, date of deposit, Horizon Europe funding details, grant project name, acronym, and number, licensing terms, and persistent identifiers for the dataset, authors and their affiliations. Additionally, where applicable, the respective metadata will be linked to related publications and other research outputs that contain persistent identifiers as well.

Search keywords will be provided in the metadata to optimise the possibility of data discovery and re-use, including author, year, research topic, or other free text such as heat waves, impact, heat stress/thermal perception, climate indices, health impact, risk, and health emergency.

Moreover, metadata will be provided in a way that allows it to be harvested and indexed. This will be ensured by adopting a metadata standard.

For multiple-version datasets, metadata will be developed for each version.

3.2. Making Data Accessible

3.2.1. Repository

The data, along with the metadata, will be deposited in the Zenodo repository and on the project portal available at www.scewero.eu. The first option for SCEWERO data and metadata sharing is the Zenodo portal (an account for SCEWERO project has already been created) because a DOI is automatically assigned to both datasets and metadata, and the project team should not arrange it separately. A list of data types, along with a detailed description of the datasets (metadata), will be prepared for submission.

3.2.2. Data

Project-generated datasets will be made available openly in specific formats, depending on the data types. Generated maps will be shared in various formats through the Zenodo repository (see section 3.2.1) and on the project website.

Reused data and received from third-party institutions (e.g., health data, weather/climate/hydrological data) will not be shared directly.

In case of restrictions imposed by the health data providers regarding the use of raw data solely by the party in the contract (which is the project coordinator, UBB), the project partners, in particular, will have access only to processed datasets produced by researchers of the CO, based on discussions in the consortium meetings.

Crowdsourced raw data collected through the app developed during the project implementation, as well as sociological data collected through the survey, including user gender, age, health, and location, will be anonymised before publication to prevent user identification.

No data embargo is currently envisaged. If necessary, such restrictions will be provided and detailed in the revised DMP at the end of the project.

Project-generated scientific data will be provided under a CC0 licence. However, depending on the various pathways of the thermal comfort/stress perception research, data will be collected during the project implementation, and certain data will be subject to restrictions on use. Access

will be granted or conditioned upon meeting specific access conditions. In addition, Non-Disclosure Agreements will be used when appropriate and if the SCEWERO consortium considers outsourcing certain operations or sharing data with third parties.

No personal data collected from participants in the organised events will be shared.

At this stage, we do not ascertain the identity of persons accessing our data. Once available in a public repository, the project-generated thermal perception data will be available for reuse without prior approval or an access committee, in compliance with the Creative Commons license mentioned above (CC0).

3.2.3. Metadata

Metadata of the newly generated dataset will be made openly available under CC0.

Data will be available as long as the chosen repository permits it. Furthermore, the data will be available on the project website for at least 10 years. The 10-year retention period applies exclusively to project-generated, non-confidential datasets. Confidential health microdata received under contractual agreements will be securely deleted upon project completion, in accordance with provider requirements.

In the final version of the DMP to be delivered at the end of the project, the consortium partners will agree on the procedures for data usage beyond the project's completion.

As previously mentioned, the data will utilise various standard formats, thereby third parties can use their tools or open-source tools for data usage (e.g. QGIS, ESRI ArcGIS, Geoserver for georeferenced data and standard editing tools such as MS Word, Excel, or their open-source alternatives for text or table data) (see data formats in section 2).

3.3. Making Data Interoperable

The data and metadata vocabularies used will ensure standardised formats that enable exchange and re-use among researchers, institutions and third parties. Data will be unencrypted and usable by open-source tools. In this respect, SCEWERO will use generally accepted standards and formats to organise and structure the data and metadata. Geographical data will comply with Open Geospatial Consortium standards.

The SCEWERO project will establish a new national set of thresholds for heat-related early warnings, which will be published as a vocabulary to ensure that values are understood. In terms of data formats, they will maintain full compliance with OGC standards.

Our analysis on extreme events is mainly based on existing datasets such as ERA5 (Hersbach et al. 2020), C3S seasonal forecast data (https://climate.copernicus.eu/seasonal-forecasts) and CMIP6 projections (Jukes et al., 2020), as already stated before.

3.4. Increase Data Reuse

As the project progresses and data is identified and collected, further information on increasing data reuse will be outlined in the final version of the DMP.

Open-source software will be utilised as much as possible, while reusing the same infrastructure as in previous projects carried out by consortium members. Reused data will be properly cited according to the rules established for existing work.

The data generated within the project will be made available under a Creative Commons license (CC0) through a public repository and the project website, and can be used indefinitely after the end of the project.

As an exception to the previous paragraph, data provided by INS and INMH cannot be reused or shared due to data sensitivity and imposed by contractual confidentiality clauses imposed by the data providers.

The provenance of data will be documented using metadata, and, where appropriate, a *readme* file describing the data will be provided.

Data quality assurance is to be performed by research team members, as well as through automated validation by the AI-driven algorithms for data generation and error checks. AI-based algorithms will be developed in accordance with the principles of transparency, traceability, and reproducibility. Model architecture, training data sources, validation procedures, and performance metrics will be documented in agreement with international ethics requirements.

4. Other Research Outputs

Research articles will be published in open-access journals. Where appropriate, datasets will be provided as supplementary materials (e.g., survey data, thermal perception data, etc).

Given the project's specificity, numerous training materials (e.g., presentations and lecture recordings, where applicable) are developed for the dedicated WPs (4, 5 and 6). They are/will be freely available through the project website and in the Zenodo repository, according to the GA. For instance, all presentations delivered during training sessions and summer schools on extreme events will be made available on the project website (www.scewero.eu).

The mobile application developed in the project will be made available in major stores for Android (Google Play) and iOS (App Store) as free software.

AI-based algorithms developed during research on extreme heat events and their impacts, when applied to Romania, will be made available on the project website and in a public repository.

5. Allocation of Resources

The costs for making data or other research outputs FAIR in the SCEWERO project (e.g. direct and indirect costs related to storage, archiving, reuse, security, etc.) are expected to be minimal. The costs for data curation and preservation are estimated at about 2,000 EUR.

The open-access publishing fee for the research papers generated in the project is planned to be covered by the requested EC contribution and is foreseen as a direct cost.

The morbidity data costs were covered by the requested EC contribution, foreseen in the budget as a direct cost.

In the SCEWERO project, each individual partner will be responsible for data management and compliance with its own rules and national/EU regulations.

The project General Assembly will decide at the end of the project what data will be kept and for how long, depending on the GA and the selected repository conditions. According to the GA, P3

(INDECO) guarantees storage for at least 10 years at a very low cost, utilising its own infrastructure. Additionally, because the data is placed in a free repository, it will remain available after the end of the project, subject to the repository owner's policy.

The project platform will be hosted by P3 (INDECO) for at least 10 years and made accessible for National and Local Authorities. GIES/IGSU and NMA will be granted free access to the Platform and data downloads for that time period. Algorithms developed within the SCEWERO project for data processing will be published as open-source software that third-party developers can re-use.

6. Data Security

The SCEWERO project does not envisage generating sensitive data that can raise specific security issues. It does not have potential for military applications or elements that can harm humans, animals or the environment.

To ensure data security, all partners follow best practices for systems management. Data repositories curated by P3 (INDECO) are covered by the company's compliance with ISO 27001 standards. Research data provided through the project website is stored in a secure geospatial environment provided by P3 (INDECO), based on a GeoServer implementation with Postgres/PostGIS support. It is firewall-secured and accessible only via the project website. Project data stored on OneDrive, provided by CO (UBB), is a space for voluntary data exchange from project partners. It is designed to facilitate collaborative work during project implementation at the consortium level and is not public.

SCEWERO data will be stored by P3 for at least 10 years on its own infrastructure, with backup and high availability configurations. Also, research data will be curated and preserved in the project's database, in accordance with data quality, security, and privacy standards, as well as relevant EU legislation. They will be stored in the publicly recognised and trusted repository recommended by the REA (Zenodo).

For each partner in the Consortium, the PIs will ensure that all team members involved in project activities related to data collection are aware of, understand, and adhere to the regulations outlined in the current DMP.

After the project ends, the mobile application and the platform will be updated only if the technical organisational measures are met.

Data collected and reused by each partner should comply with the principles and regulations of the latest version of the SCEWERO DMP. If any practical issues regarding data handling arise during project implementation, they will be discussed during the General Assembly and the Work Package Group Leaders' meetings.

7. Ethics

The primary legal concern is the processing and sharing of personal data collected from participants at summer schools, workshops, training sessions, and other scientific events, as well as the collection of data through the mobile application, the survey and the mortality and morbidity data. As the project involves the collection of personal information (e.g., names,

contact details, organisational affiliations, education, and financial details), all data processing and sharing must fully comply with the *General Data Protection Regulation* (GDPR) (EU 2016/679).

Ethical considerations include the responsibility to protect participant privacy and confidentiality, ensure *informed consent,* and prevent the unauthorised or unintended disclosure of personal information. Moreover, the project commits to the principles of *data minimisation* and *purpose limitation*, ensuring that only necessary data is collected and used solely for predefined project activities (e.g., issuing certificates, managing reimbursements, reporting, and identifying vulnerable groups at risk of heat stress).

Potential **legal risks** involve:

- The risk of data breaches if proper technical and organisational measures are not implemented;
- Limitations on international data transfers, should any non-EU partners or subcontractors be involved;
- Restrictions on sharing personal data with third parties outside the consortium without a lawful basis or explicit consent.

Ethics-related risks include:

- The potential for participants to feel coerced into sharing personal data without fully understanding how it will be used;
- Reusing research data in publications or reports without appropriate anonymisation or deidentification;
- To mitigate these issues, the project will implement strict data management protocols and conduct an internal ethics review before the beginning of data collection. All participants will be informed about their rights under the GDPR, including the right to withdraw consent and request the deletion of their data.

In the questionnaire developed, informed consent for data sharing and long-term preservation will be included in questionnaires dealing with personal data, and participants will be fully informed about:

- The nature and purpose of the data being collected;
- How their data will be processed, stored, and potentially shared (e.g., within the consortium or in anonymised form for scientific publications);
- The duration for which their data will be retained, including provisions for long-term preservation in accordance with institutional, national, and EU data management policies.

The consent form integrated into the questionnaires will clearly explain that:

- Data will only be shared in compliance with GDPR, and when necessary, will be anonymised or pseudonymised to protect participant privacy;
- Participants have the right to refuse data sharing or long-term preservation without any negative consequences on their involvement in the project;
- Participants retain the right to withdraw their consent at any time, as well as the right to access, rectify, or request deletion of their personal data.

Consultations with the Ethics Advisors have been conducted during the project's implementation period to ensure that all ethical requirements are met.

All ethics issues related to personal data protection and AI-based algorithms will be assessed in the two ethical reports delivered by the external independent Ethics Advisor. Both reports are deliverables of the dedicated WP (WP1 - *Ethics requirements*) and are planned to be delivered at the mid-term and at the end of the project.

8. Data Retention and Deletion Policy

The SCEWERO project distinguishes among data categories and applies specific retention rules. They are synthetically present in the following subsections.

8.1. Project-Generated Open Research Data

They consist of project datasets, indicators, maps, thresholds, AI outputs, and anonymised survey datasets.

- Retention period: **minimum 10 years after project completion**, in line with Horizon Europe Open Science policy.
- Storage location: Zenodo repository and project platform (hosted by P3).
- Legal basis: open science principles under HE GA.
- Preservation method: repository-based long-term archiving.
- Responsible entity: Universitatea Babeș-Bolyai (CO) in collaboration with Indeco Soft (P3).

8.2. Health Data

These data consist of mortality and morbidity (medical emergency presentations) data received on a request basis and under contractual conditions from the national data providers in Romania.

- Nature: restricted, non-shareable data under contractual agreement.
- Retention period: only for the duration of the project (until M36).
- Action after project completion: secure deletion in accordance with data provider requirements.
- Legal basis: data processing agreement; Article 9(2)(j) GDPR (research purposes).
- Responsible entity: project coordinator (UBB).

No copies will be stored in open repositories.

8.3. Crowdsourced Perception Data

These data are collected via the mobile app developed in the project or the sociological survey.

- Raw identifiable data: retained only as long as necessary for validation and processing.
- Anonymised datasets: retained for 10 years and made publicly available under CC0 (if no restriction applies).
- Participants may request deletion at any time (GDPR Art. 17).

8.4. Personal Administrative Data of Participants

This type of data refers to the participants in various events organised during the project, such as training sessions, workshops, and summer schools.

- Data category: names, emails, financial details.
- Retention: According to CO UBB administrative and financial regulations (typically 5–10 years depending on fiscal requirements).

- Access: Restricted to authorised administrative staff.
- Not shared or published.

9. Other issues

In addition to the guidelines provided by this document, all partners will handle data according to the applicable European and national laws and regulations and accordingly with any internal procedures, guidelines and rules adopted and in force in their respective cases.