



Climate-Resilient Development Pathways in Metropolitan Regions of Europe

Deliverable 5.6: IDSS architecture
July 2025

Grant Agreement Number 101137851

Deliverable No:	D5.6
Deliverable nature:	Report
Work Package and Task:	WP5, Task 5.1
Dissemination level:	Public
Authors:	Fabiola Ramelli, Stefanie Kieferle (MM), Vasile Crăciunescu, Irina Onțel (MeteoRo)
Reviewers:	Panagiota Koltsida, Eleni Toli (ATHENA-RC), Sorin Cheval, Laurențiu Ciuhu, Angela Cristina Pleșcan, Dana Micu (MeteoRo), Mirabela Marin, Nicu Constantin Tudose (INCDS), Mihai Răzvan Niță, Cristian Iojă (UB), Emma Ferranti (UoBir), Latifa Yousef (WMO), Steffi Urhausen, Debbie Hemming (MOHC), Andrés Carrancá (UFR), Eleftheria Exarchou (MITIGA), Marianna Adinolfi (CMCC), Aphrodite Bouikidis, Iphigenia Keramitsoglou (NOA), Christos Spyrou (AoA), Elena Xoplaki (JLU)
Contractual submission date:	31 July 2025
Actual submission date:	31 July 2025

Technical References

Project title	Climate-Resilient Development Pathways in Metropolitan Regions of Europe
Project acronym	CARMINE
Grant agreement number	101137851
Instrument	HORIZON-RIA
Call	HORIZON-CL5-2023-D1-01
Starting date of the project	01 February 2024
Project duration	48 months

Document history

Version	Date	Description
0.1	18 June 2025	Initial structure
1.0	15 July 2025	First draft submitted for consortium review
1.1	22 July 2025	Second draft after internal review
1.2	30 July 2025	Final version

Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Climate, Infrastructure and Environment Executive Agency (CINEA). Neither the European Union nor the granting authority can be held responsible for them. Every effort has been made to ensure that all statements and information contained herein are accurate, however the CARMINE Project Partners accept no liability for any error or omission in the same.

UK participants in this project are co-funded by UK Research and Innovation (UKRI).

The work of Meteomatics partner (Switzerland) has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI).

Table of Contents

List of Figures	5
Abbreviations	5
1 Introduction	8
1.1. Purpose and scope	8
1.2. Relation to other deliverables	8
1.3. Structure of the document	8
2 IDSS core principles and objectives	9
2.1. Scope and purposes	9
2.2. Guiding technical principles	9
3 The open architecture of the IDSS	10
3.1. High-level architectural overview	10
3.2. Data acquisition and integration	11
3.3. Key architectural components	11
3.4. Interaction with the Digital Twin Platform	12
3.5. Core Visualization Component: The MetX Solution	12
3.6. Decision support engine	14
3.7. Interoperability and collaboration layer	14
3.8. Security and data privacy	15
3.9. Co-Development, validation and sustainability	15
4 Conclusions	16
5 Apendix 1	17
5.1. Introduction	17
5.2. Data acquisition	17
5.3. Sources of data	17
5.4. Licensing and access constraints	18
5.5. Data formats	18
5.6. Data transformation (ETL - Extract, Transform, Load)	18
5.7. Publication and access	20
5.8. Data catalog and metadata	20
5.9. Maintenance and updates	21
References	21

List of Figures

Figure 1: IDSS general architecture and data flow 10

Figure 2: Example of CARMINE MetX graphical interface integrating data from Brasov CSA... 13

Abbreviations

Abbreviations	Full wording
ANCPI	Romanian National Agency of Cadastre and Land Registration
API	Application Programming Interface
AWS	Amazon Web Services
C-A	European Climate-ADAPT
C3S	Copernicus Climate Change Service
CAMS	Copernicus Atmosphere Monitoring Service
CC 0	Creative Commons Public Domain Dedication
CC BY	Creative Commons Attribution
CDSE	Copernicus Data Space Ecosystem
CEMS	Copernicus Emergency Management Service
CINEA	European Climate, Infrastructure and Environment Executive Agency
CLMS	Copernicus Land Monitoring Service
CMEMS	Copernicus Marine Environment Monitoring Service
COG	Cloud Optimized GeoTIFF
CSA	Case Study Area
CSV	Coma Separated Values
DEM	Digital Elevation Model
DestinE	Destination Earth
DT	Digital Twin
DXF	Drawing Exchange Format
EC	European Commission
EEA	European Environment Agency
EPSG	European Petroleum Survey Group
ESA	European Space Agency
ETL	Extract, Transform and Load
FDMS	Federated Data Management Solution
GDAL	Geospatial Data Abstraction Library

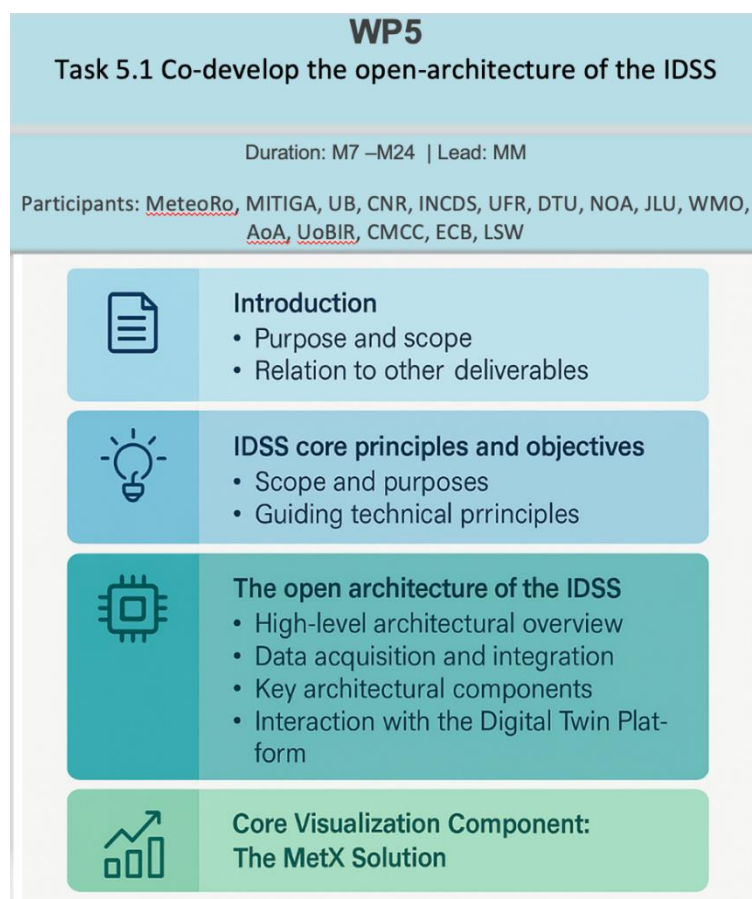
GDPR	General Data Protection Regulation
GeoJSON	Geographic JavaScript Object Notation
GeoTIFF	Georeferenced Tagged Image File Format
GIS	Geographic Information System
GML	Geography Markup Language
GRIB	GRIdded Binary
GUI	Graphic User Interface
IDSS	Impact-Based Decision Support Service
IETF	Internet Engineering Task Force
INSPIRE	Infrastructure for Spatial Information in Europe
ISO	International Standardization Organization
IT	Information Technology
JPEG	Joint Photographic Experts Group
KML	Keyhole Markup Language
LiDAR	Light Detection and Ranging
Living Labs	LL
MeteoRo	Romanian National Meteorological Administration
netCDF	Network Common Data Form
OGC	Open Geospatial Consortium
OSM	OpenStreetMap
PNG	Portable Network Graphics
REST	Representational State Transfer
SC-HUB	Stakeholder Community Hub
SERI	Swiss State Secretariat for Education, Research and Innovation
TIFF	Tagged Image File Format
UKRI	UK Research and Innovation
W3C	World Wide Web Consortium
WCS	Web Coverage Service
WFS	Web Feature Service
WMS	Web Map Service
WMTS	Web Map Tile Service
WP	Work Package

Executive Summary

This deliverable outlines the open architecture of the CARMINE Impact-Based Decision Support Service (IDSS). The IDSS is a critical component of the CARMINE project, designed to act as a forecast and advisory service that translates complex climate and socioeconomic data into actionable insights for metropolitan decision-makers and policymakers.

Deliverable 5.6 is developed within the framework of Task 5.1: *Co-develop the open-architecture of the IDSS*, which is implemented between Month 7 and Month 24. The task is led by MM, with active participation from MeteoRo, MITIGA, UB, CNR, INCDS, UFR, DTU, NOAA, JLU, WMO, AoA, UoBIR, CMCC, ECB, and LSW.

This document details the architectural design, key components, data flows and operational framework of the IDSS. It also describes its important integration with the CARMINE Digital Twin (DT) Platform and the Federated Data Management Solution (FDMS) to support local adaptive capacity and enhance cross-sectoral climate resilience. The architecture has been co-developed considering stakeholder needs and sectoral priorities, ensuring stability, scalability and interoperability with existing European platforms like the Climate-ADAPT (C-A) platform.



Key words

Open Architecture, decision support, .digital twins, climate resilience, data integration, visualisation platform, interoperability, stakeholder co-development, cloud services, ETL pipeline

Introduction

1.1. Purpose and scope

The CARMINE project aims to help metropolitan communities become more climate-resilient by co-producing knowledge-based tools, strategies, and plans. A key output of this project is the IDSS, developed within WP5. The primary goal of the IDSS is to empower stakeholders (including decision-makers and policymakers) by providing them with user-friendly access to high-resolution modeling outputs that support local adaptation assessments and plans.

This deliverable, D5.6, presents the open architecture of the IDSS. It has been designed as an advisory service to reinforce local adaptive capacities and the dynamic management of complex metropolitan systems. The architecture supports the integration of data from multiple sources, including:

- Stakeholder data from the Living Labs (LL) (WP1)
- Outputs from the climate and socioeconomic risk data and related model outputs (WP2 & WP3)
- Simulations from the eight Case Study Area (CSA) DTs (WP4)

The IDSS architecture is designed to be open, stable and elastic, adapting the [ESA-based "Earth Observation Exploitation Platform Common Architecture"](#) to simplify the development of distributed climate services. This document will detail the technical components, data flows and the plan for iterative calibration and validation with stakeholders to ensure the IDSS meets user requirements and promotes better, science-based decision-making (Wenkel et al., 2013).

1.2. Relation to other deliverables

This deliverable closely aligns with Deliverable D2.1 (Report on the modelling framework), D4.1 (Climate and socioeconomic risk data and related model outputs) and D4.2 (Digital Twin Architecture), describing the critical interactions and integration aspects necessary for seamless data and service interoperability.

1.3. Structure of the document

The document follows a logical structure, beginning with an overview and progressively detailing the technical components and implementation methodologies of the open-architecture.

IDSS core principles and objectives

The design of the IDSS is guided by the overarching goal of blending climate and environmental data with socioeconomic impact and risk assessments (Palutikof et al, 2019). Its core principles are rooted in the need to provide quick access and operational guidance to cross-sectoral knowledge.

To clarify its role, we differentiate between the service's high-level purpose and the technical principles that guide its architecture.

2.1. Scope and purposes

The primary purpose of the IDSS is to serve as a user-centric advisory service. Its goals are:

- **To support informed decision-making** - the service is designed to empower decision-makers to identify, visualize and verify the potential effects of transformative adaptation measures, helping them to prioritize cost-effective solutions.
- **To provide interactive and actionable insights** - the IDSS will offer timely and accurate, interactive access to tailored data through a user-friendly interface, using customized maps, charts and dashboards.
- **To ensure relevance and legitimacy through co-development** - the service will be co-developed and iteratively validated with sectoral stakeholders and target user groups to ensure its final form is relevant, usable and trusted by the communities it aims to serve. To serve this diverse audience, the IDSS is being developed with a dual-access approach:
 - A web-based graphical user interface (GUI) - the primary public face of the IDSS will be an intuitive, web-based graphical interface accessible through any standard browser. This interface is specifically designed for non-technical users. It will feature customized dashboards and workflows tailored to specific user groups (e.g., urban planners, emergency service managers), allowing them to explore data, play with predefined scenarios and generate reports without needing specialized software or data science skills.
 - Standardised APIs - for advanced users, such as researchers, data analysts or developers working within larger organizations, the IDSS will also offer programmatic access to its data and analytical functions via a well-documented API. This will allow them to integrate IDSS outputs into their own models, perform custom analyses and build specialized applications on top of the CARMINE infrastructure.

2.2. Guiding technical principles

To achieve the purpose outlined above, the IDSS architecture adheres to the following core technical principles:

- **Integration of diverse and complex data** - the architecture must be capable of integrating data from multiple, complex sources, including the DT platform, local CSA data, physical risk modeling outputs and socioeconomic data.
- **Interoperability and long-term sustainability** - the service is being built with state-of-the-art technologies and standard-compliant interfaces (e.g., OGC API) to facilitate interoperability and to establish sustainable, operational links with key European infrastructures like the C-A platform.

The open architecture of the IDSS

The IDSS architecture is designed in a modular fashion, separating distinct macro-components to ensure flexibility, scalability and security (Anghel et al., 2024). This design simplifies the integration of various data sources and allows for the smooth incorporation of the tools and information developed in other work packages.

3.1. High-level architectural overview

The IDSS functions as an integration and visualization hub, building upon the foundational infrastructure of WP4. As illustrated in the data flow diagram (Figure 1) the IDSS sits downstream from these WP4 components. It is specifically designed to access the outputs from the DTs via the FDMS, which acts as the central gateway for all data coming from the DT platform. Some other connectors are needed to especially integrate with CSAs specific data, general reference data (e.g. statistics, GIS layers, etc.) and the relevant European initiatives.

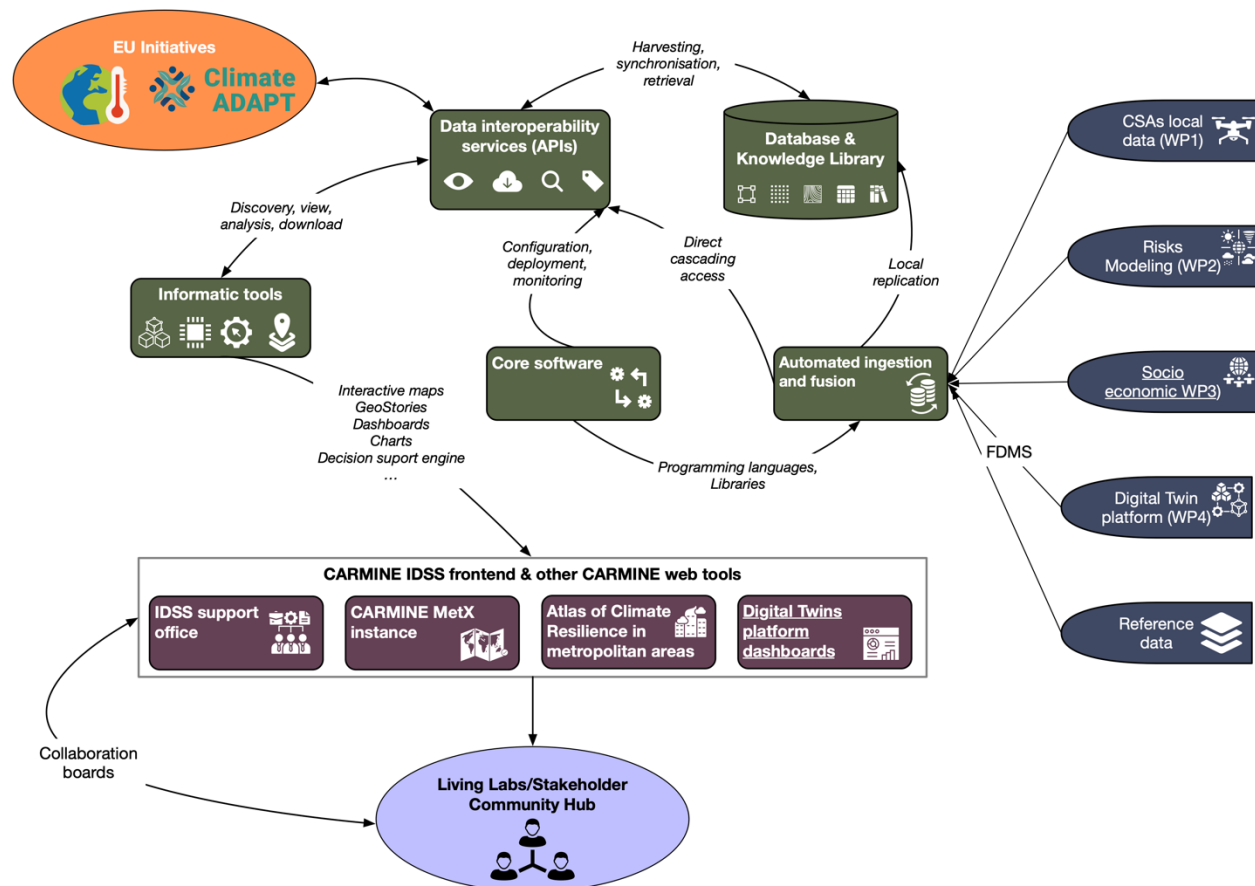


Figure 1: IDSS general architecture and data flow

3.2. Data acquisition and integration

IDSS design accommodates two primary data ingestion pathways:

- A streamlined, centralized flow from the DT FDMS.
- A flexible, direct ingestion pathway for rich, local data sourced and prepared independently by each CSA and/or various open data platforms.

This dual-pathway approach ensures that the IDSS benefits from both the harmonized outputs of the project's core modeling efforts and the detailed, context-specific knowledge held at the local/national/European level.

Data from the project's core infrastructure, including the outputs of the eight DTs, are processed and made available through the FDMS. The IDSS connects to the FDMS via standardized APIs to consume these harmonized outputs.

IDSS is also designed to directly integrate local data that is not managed by the FDMS. This local data, usually sourced by each CSA or relevant institutions (e.g. national mapping agencies), is essential for grounding the analysis in the specific realities of each metropolitan region and for addressing local challenges identified in WP1-3. The acquisition and preparation of this local data follow a structured methodology which includes:

- Data acquisition - CSAs are gathering data from a wide variety of sources, including public open-data portals (e.g. INSPIRE, Copernicus Services), direct requests to governmental authorities (e.g. urban planning departments, environmental agencies, etc.) and potential partnerships with the private sector.
- Data transformation (ETL) - raw local data, which comes in multiple formats (e.g. Shapefiles, GeoTIFF, CSV), undergoes a significant transformation process. This includes cleaning, validation, re-projection to a consistent coordinate system, and enrichment (e.g. joining tabular census data to spatial boundaries) to ensure it is homogenized and ready for integration into IT systems like the IDSS. The infrastructure for this component is assured by MeteoRo for the entities lacking such specialized IT infrastructure/knowledge.
- Publication and access - once prepared, this local data is made available for the IDSS to consume. This can be achieved by publishing the data through OGC-compliant services (e.g. WFS, WMS) or by storing it in cloud-optimized formats (e.g. GeoParquet, COG) that the IDSS Data Access and Discovery Component can efficiently query.

A short guide was prepared for the above-mentioned steps and is presented in Appendix 1.

3.3. Key architectural components

The IDSS is composed of several logical macro-components, each with a specific function:

- Data access and discovery component. This component is responsible for connecting to the FDMS and other external data sources. It utilizes the standard-compliant APIs (e.g., OGC API, IETF, W3C) developed as part of the DT Platform's Data Management Service to pull the required information. It handles the discovery and access of datasets from WP1 (local stakeholder data), WP2 (physical modeling), WP3 (socioeconomic data) and WP4 (DT outputs).
- Processing and analysis component: This component contains the logic for integrating and processing retrieved data streams. It uses advanced algorithms and statistical approaches to

analyze large volumes of data. This is where the synthesis of climate hazards, socioeconomic vulnerabilities and adaptation scenarios occurs to create impact-based insights that are central to the IDSS.

- User interface and visualization component: This is the front-end of the IDSS, providing the interactive experience for the end-user. It is built upon a validated software solution (MetX by Meteomatics) that will be adapted for CARMINE's specific needs. The visualization tool will offer:
 - A customizable multi-layer capability for interactive maps.
 - User-friendly tools to generate charts, animations, and tables.
 - Dedicated dashboards tailored to the specific climate challenges and adaptation options of each CSA.
- Security and data privacy components. This component ensures that access to data is managed securely, in full compliance with GDPR and other data privacy regulations. It will interface with the central [Keycloak](#) Identity Access Management service established in the DT Platform architecture to manage user authentication and authorization.

3.4. Interaction with the Digital Twin Platform

The synergy between the IDSS and the DT Platform is very important to the CARMINE project. The DT platform is responsible for hosting, processing, and managing the data streams from the eight individual Digital Twins. The interaction will be managed as follows:

- FDMS will expose the outputs of the DT simulations via standardized APIs.
- FDMS will facilitate the connection, allowing the IDSS to receive the required output data either on-demand or through an automated push mechanism when new data becomes available.
- The IDSS will consume these outputs (such as flood risk maps, heat health risk assessments, or wildfire scenarios from the various DTs) and integrate them into its own cross-sectoral dashboards for decision-makers.

This architecture ensures a clear separation of concerns: the DT Platform focuses on the simulation and management of digital replicas, while the IDSS focuses on the final integration and user-facing decision-support functionalities.

3.5. Core Visualization Component: The MetX Solution

The power of the DSS lies in its ability to translate complex data into clear, intuitive, and interactive tools for decision-makers. The technological core that enables this translation is [MetX](#), a validated, pre-existing software solution developed by project partner Meteomatics. Rather than building a visualization platform from scratch, the CARMINE project leverages this robust and proven solution, adapting it to meet the project's unique, cross-sectoral needs.

MetX is a web-based weather tool on top of the Weather API developed by Meteomatics. It visualizes all available meteorological and oceanographic API parameters, such as temperature, pressure, humidity, precipitation and cloud cover in tools like weather maps, plots or tables. In up to 9 views in one Tab, the following features are available in the MetX tool:

- Layers to showcase a wide range of weather parameters
- Real-time weather data calculation when zooming or changing views
- Points, lines, and areas of interest can be color-coded based on user-defined thresholds (e.g., red if temperature < 0 °C)
- Clear display of up to six maps on one screen offering the capability to compare different parameters
- Accurate data visualization even over large areas, including global view
- Various basemaps available, including topographic maps, border maps, aeronautical maps, and more
- Custom data import in various formats (e.g., CSV, GeoJSON)
- Data export available for each point on the map
- Save and share dashboards across users
- Display and export animations

Decision making in MetX:

MetX offers a highly customizable layerstack in which every parameter can be adjusted to the user needs.

The option to upload data in a CSV file or GeoJSON allows the tool to visualize dedicated user needs and support decision making by coloring different datasets by custom thresholds. For example, in Figure 2, all areas with a temperature above 10 °C are highlighted in red.

All the MetX visualizations can be downloaded and shared directly from the profile.

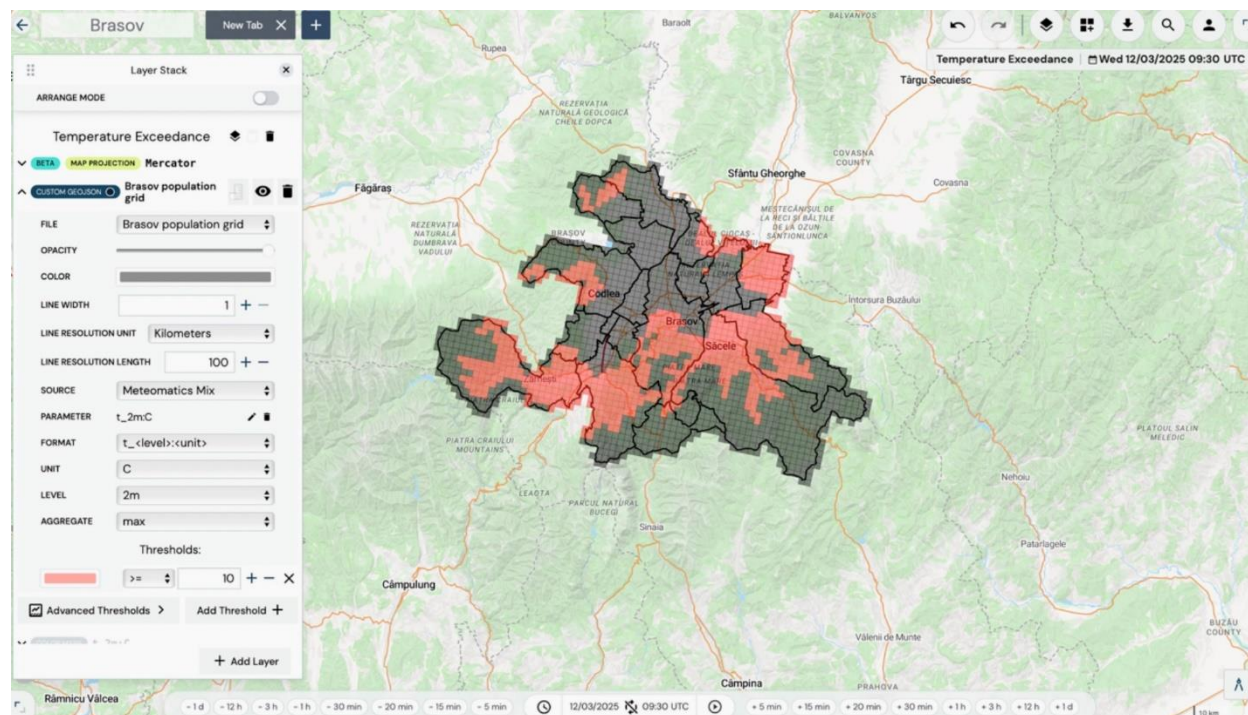


Figure 2: Example of CARMINE MetX graphical interface integrating data from Brasov CSA

3.6. Decision support engine

This component provides concrete recommendations and decision support based on data analysis and climate simulations. Decision-support algorithms help evaluate and prioritize response measures.

Functionalities:

- Optimization algorithms to propose mitigation and adaptation measures (e.g. based on adaptation tipping-point approach to develop adaptation measures - Kwadijk et al., 2010).
- Impact assessment on the economy, health, infrastructure, etc.
- Scenario planning that helps decision-makers protect communities.

Technologies:

- Expert systems: for automatically generating recommendations based on predefined rules.
- Integration technologies with other decision systems: via APIs and webhooks for automated decision-making.

3.7. Interoperability and collaboration layer

IDSS must be able to communicate with other platforms (e.g. local/regional platforms, C-A platform, etc).

Functionalities:

- Interfacing with other systems.
- Data sharing with other relevant institutions.
- Collaboration and communication solutions for decision-makers.

Technologies:

- Generic RESTful APIs for integration.
- OGC services for sharing geospatial data.
- Collaboration boards such as Slack, Mattermost, Microsoft Teams, Discourse, etc. for team communication.

Regarding the data sharing, in line with the project's open science policy, the data generated will be made available under open licenses, preferably Creative Commons Attribution (CC BY) or Creative Commons Public Domain Dedication (CC 0). An exception will be made for sensitive or restricted datasets provided to the IDSS by third-party authorities. This data will be used solely within the context of the service's analyses, and any potential for further sharing will be strictly governed by the original license and access conditions imposed by the data provider.

3.8. Security and data privacy

This component ensures that access to data is managed securely, in full compliance with GDPR and other data privacy regulations. It will interface with the central Keycloak Identity Access Management service established in the DT Platform architecture to manage user authentication and authorization, ensuring that sensitive data is protected.

3.9. Co-Development, validation and sustainability

The architecture of the IDSS is not static, it is designed to be co-developed and iteratively refined based on user feedback:

Co-development and calibration - early prototypes of the IDSS will be subjected to a critical evaluation process involving key stakeholders from all Target Groups. Interactive sessions, organized through the LL and the Stakeholder Community Hub (SC-HUB), will be used to gather feedback and ensure the service is legitimate and relevant to end-users.

Validation - a formal validation process will be applied to all IDSS components to ensure the system produces accurate and reliable results. User manuals, guides, and video tutorials will be developed to support this process and facilitate user adoption.

Sustainability and integration - the IDSS is designed for long-term sustainability beyond the project's lifetime. By building on the principles of the C-A platform and establishing operational links, the goal is for the CARMINE IDSS to be integrated as a new, valuable tool within the existing European climate services landscape.

Conclusions

The open architecture of the IDSS provides a robust and flexible framework for delivering on the core objectives of CARMINE. By leveraging the outputs of the DT Platform and integrating them with a wide range of climate, environmental and socioeconomic data, the IDSS will empower local and regional authorities with the tailored, science-based insights needed to build resilient and sustainable metropolitan communities. The co-development approach ensures that the final service will be a user-centric tool that effectively supports informed decision-making for climate adaptation.



**Funded by
the European Union**

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Climate, Infrastructure and Environment Executive Agency (CINEA). Neither the European Union nor the granting authority can be held responsible for them. UK participants in this project are co-funded by UK Research and Innovation (UKRI). The work of the Meteomatics partner (Switzerland) has received funding from the Swiss State Secretariat for Education (SERI).

 Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Swiss Confederation

Federal Department of Economic Affairs,
Education and Research (EAB)
State Secretariat for Education,
Research and Innovation (SERI)



**Co-funded by
UK Research
and Innovation**

Appendix 1

5.1. Introduction

This document outlines the methodology for building a comprehensive database for metropolitan regions for the purposes of the CARMINE project. It covers data sourcing, transformation, storage, publication and cataloging. The methodology is based on the experience gained in creating the database for CSA Brasov and it's provided to the project partners as a guide. The document is structured into several key phases.

5.2. Data acquisition

Data acquisition is the first step in building a CARMINE CSA database. It involves collecting spatial and non-spatial data from various sources, such as public authorities, open data portals and private entities. Key considerations include data licensing, access constraints and formats.

5.3. Sources of data

Public datasets

- Public governmental open data portals (in case of Romania <https://data.gov.ro>) for statistics, demographic, etc.
- INSPIRE portals (e.g. <https://inspire-geoportal.ec.europa.eu/srv/eng/catalog.search#/home>) for DEMs, orthophotos, general GIS layers.
- Institutional open data portals (examples in case of Romania: <https://www.roadapt.ro>, <http://statistici.insse.ro:8077/tempo-online>, [ANCPI](#), etc.) for socioeconomic and environmental data.
- [OpenStreetMap \(OSM\)](#) for transportation, points of interest, etc.
- [Overture Maps Foundation](#) for building footprints.
- [OpenTopography](#) for global and regional DEMs.
- [GeoNames](#) for global place names.
- [OpenAerialMap](#) for drone imagery.
- Copernicus satellite data for any purpose ([CDSE](#), [WEKEO](#)).
- Copernicus services for marine data ([CMEMS](#)), land data ([CLMS](#)), atmosphere ([CAMS](#)), climate change ([C3S](#)), emergency situations for natural or man-made disasters ([CEMS](#)).
- Copernicus thematic hubs for health information ([Health Hub](#)), European coastal zones ([Coastal Hub](#)), energy-related data ([Energy Hub](#)), Arctic region ([Arctic Hub](#)).
- [ESA Climate Change Initiative \(CCI\)](#) for climatic data derived from satellite imagery.
- Global environmental layers: [OpenLandMap](#), [ESA WorldCover](#), [Source Cooperative](#).
- [Eurostat population grids](#).
- [World Bank](#) or other international organizations for socioeconomic data.

Governmental authorities

Requests for data can be made to specific departments, such as:

- Urban planning departments (for land use, zoning, etc.).
- Environmental agencies (for climate, water and biodiversity data).
- Transportation authorities (for road networks, public transit data).
- Utilities (for energy, water and telecom infrastructure).

Data access may require formal data requests, often under European/national open data policies (e.g. [EC Open data and the reuse of public-sector information](#)).

Private sector

- Real estate firms, telecommunications companies and other industries may have useful datasets available through partnerships or licensing agreements.
- Satellite or drone imagery, LiDAR data and other geospatial information can be sourced from vendors like [Planet](#).

5.4. Licensing and access constraints

When acquiring data, it is important to identify:

- Licensing constraints. Public domain (e.g., [CC0](#)), open licenses (e.g., [CC-BY](#), [CC-BY-SA](#)), or proprietary licenses (requiring permissions for use, distribution, etc.).
- Access constraints. Some data may have restrictions based on sensitivity (e.g., critical infrastructure), requiring secure storage or limited access.

5.5. Data formats

Different formats should be anticipated:

- Vector data: [Shapefiles](#), [GeoJSON](#), [KML](#), [GML](#), [DXF](#).
- Raster data: [GeoTIFF](#), [COG \(Cloud-Optimized GeoTIFF\)](#), [NetCDF](#), [GRIB1/GRIB2](#), [TIFF/PNG/JPEG](#) + [world file](#).
- Tabular data: [CSV](#), Excel, databases, scanned documents.

5.6. Data transformation (ETL - Extract, Transform, Load)

After acquiring the data, it must be processed for integration into a homogenized CSA database. This phase involves data cleaning, standardization and optimization.

Extraction

- Initial loading. Raw data from various sources is extracted and loaded into a working environment (GIS software, ETL tools like [GDAL](#) or Python-based scripts).
- Format conversion. Convert different formats into compatible formats. For example:
 - Shapefiles or KML may be converted to GeoPackage or GeoParquet.
 - Raster data (TIFF, JPEG) may be converted to COG for efficient cloud usage.
 - Tabular data (CSV, Excel) may need to be geocoded or joined with spatial datasets.

Transformation

Data often requires cleaning, enrichment, or re-projection before it can be integrated into the IT systems like database, CARMINEDTs, CARMINE IDSS:

- Re-projection. Convert datasets to a consistent coordinate reference system (e.g., [EPSG:4326](#) or [EPSG:3857](#)).
- Cleaning and validation. Remove duplicates, validate geometries and ensure data completeness. Tools like [QGIS](#), [ArcGIS](#) or Python libraries (e.g., [geopandas](#), [shapely](#)) are useful here.
- Data enrichment. Join demographic or economic data from tabular formats to spatial boundaries (e.g., joining census data to municipality boundaries).
- Standardization. Harmonize attribute names, units of measurement and classification systems across datasets.
- Optimization for storage. Simplify geometries (if appropriate) and compress data for optimal storage and access. This is important for web services and cloud environments.

Load

- Database loading. Load the processed data into a spatial database (e.g., [PostGIS](#), [SpatialLite](#), [DuckDB](#)) or a data lake.
- Cloud storage. Store optimized datasets in cloud-friendly formats such as:
 - [GeoParquet](#) - a columnar data format optimized for fast access and analytics.
 - [FlatGeobuf](#) - an efficient binary format for streaming large amounts of vector data.
 - [COG](#) - for raster datasets like imagery or DEMs.
 - [Zarr](#) - for multi-dimensional arrays (e.g., climate data) in cloud environments.

5.7. Publication and access

Once the data is processed and loaded, it should be made available for CARMINE consortium partners/applications and stakeholders through modern geospatial services and/or cloud-based solutions.

OGC-Compliant Services

- [WMS \(Web Map Service\)/API - Maps](#). Allows users to visualize spatial data as map images over the web.
- [WFS \(Web Feature Service\)/API - Features](#). Serves vector data (e.g., points, lines, polygons) and allows querying and downloading.
- [WCS \(Web Coverage Service\)/API - Coverages](#). Serves raster data like satellite imagery and DEMs.
- [WMTS \(Web Map Tile Service\)/OGC - Tiles](#). Optimized for serving pre-rendered tiles for fast map access (useful for base maps).

Cloud optimized formats

- [GeoParquet](#) and [FlatGeobuf](#). These formats enable efficient querying and storage of large vector datasets, optimized for big data processing in cloud environments ([AWS S3](#), [Google Cloud](#), etc.).
- [PMTiles](#). A single-file archive format for pyramids of tiled data.
- [Zarr](#). Ideal for handling large, multi-dimensional climate or environmental datasets.
- [COG](#). Allows partial access to raster files stored in cloud environments, improving performance when accessing or visualizing large geospatial images.

5.8. Data catalog and metadata

A data catalog is essential for managing, discovering and accessing the various datasets in the CSA database.

Cataloging

- Data inventory. A systematic inventory of datasets including source, date of acquisition, coverage and format.
- Metadata standards. Adopt international standards such as [ISO 19115](#) or [Dublin Core](#) for metadata to describe the datasets, including:
 - Title, description, keywords.
 - Spatial extent (bounding box).
 - Data quality (accuracy, completeness, timeliness).
 - Access restrictions, licensing (e.g., public, private, restricted).

- Contact information for the data provider.
- Cataloging tools. Tools like [GeoNetwork](#), [pycsw](#), [CKAN](#) or [Datahub](#) can be used to create searchable catalogs for all datasets.

5.9. Maintenance and updates

Regular updates are important to keep the database relevant and accurate. CSA database administrators should take in consideration:

- Versioning. Implement a version control system for tracking changes in datasets over time.
- Automated updates. For datasets like traffic, public transit, or air quality, implement ETL pipelines that automate the regular updating of data.
- Data quality assurance. Regularly validate and check data integrity, especially when updating datasets.

References

Anghilea A, Smith G, Meissl S, Achtsnit S (2024) EarthCODE – ESA’s Earth Science Collaborative Open Development Environment. <https://doi.org/10.5194/isprs-archives-XLVIII-4-2024-19-2024>

Kwadijk, J.C.J., Haasnoot, M., Mulder, J.P.M., Hoogvliet, M.M.C., Jeuken, A.B.M., van der Krogt, R.A.A., van Oostrom, N.G.C., Schelfhout, H.A., van Velzen, E.H., van Waveren, H. and de Wit, M.J.M. (2010), Using adaptation tipping points to prepare for climate change and sea level rise: a case study in the Netherlands. *WIREs Clim Chg*, 1: 729-740. <https://doi.org/10.1002/wcc.64>

Palutikof, J.P., Street, R.B. & Gardiner, E.P. Decision support platforms for climate change adaptation: an overview and introduction. *Climatic Change* **153**, 459–476 (2019). <https://doi.org/10.1007/s10584-019-02445-2>

Wenkel K-O, Berg M, Mirschel W, Wieland R, Nendel C, Köstner B (2013) LandCaRe DSS – An interactive decision support system for climate change impact assessment and the analysis of potential agricultural land use adaptation strategies. *Journal of Environmental Management*. 127, S168-S183. <https://doi.org/10.1016/j.jenvman.2013.02.051>

<https://eoepca.org/eoepcaplus/>

<https://tool.adapt2clima.eu/en/home/>

<https://guide.cloudnativegeo.org/>