# ECMWF Copernicus Procurement

## Invitation to Tender


## Copernicus Joint Services

# Extended Analysis-Ready and Cloud-Optimised (ARCO) capabilities within the Data Stores Services

## Volume II: Specification of Requirements

| |
|---|
| ITT Ref: CJS2_220b_bis |
| Issued by:<br>ECMWF<br>Administration Department<br>Procurement Section |
| Date: December 2025 |
| Version: Final |

Funded by the European Union

Implemented by ECMWF

# Contents

# 1 Introduction

## 1.1 Overview

This Invitation to Tender (ITT) has been prepared by the European Centre for Medium-Range Weather Forecasts (governed by its Convention and associated Protocol on Privileges and Immunities which came into force on 1 November 1975 and was amended on 6 June 2010) ("ECMWF") for the purposes of obtaining a proposal for the provision of extended Analysis-Ready and Cloud-Optimized (ARCO) capabilities that seamlessly integrates within the architectural components of an enhanced Data Stores Services (DSS).

ECMWF is both a research institute and a 24/7 operational service, producing and disseminating numerical weather predictions to its Member States. This data is fully available to the national meteorological services in the Member States. The Centre also offers a catalogue of forecast data that can be purchased by businesses worldwide and other commercial customers. The supercomputer facility (and associated data archive) at ECMWF is one of the largest of its type in Europe and Member States can use 25% of its capacity for their own purposes. The organisation was established in 1975 and now employs around 450 staff from more than 30 countries. ECMWF is based in three locations: Reading (United Kingdom), Bologna (Italy) and Bonn (Germany).

The European Copernicus Programme is an EU-wide flagship programme that aims to support policymakers, businesses and citizens with enhanced environmental information. The programme reached operational status in 2014 and entered its second phase in 2021. For further information, refer to Section 1 of the ITT Volume I.

The Data Stores Service (DSS) encompasses the technical infrastructure underpinning the data access platforms supporting the Copernicus Services implemented by ECMWF on behalf of the European Commission (EC) – i.e. the Copernicus Climate Change Service (C3S), the Copernicus Atmosphere Monitoring Service (CAMS) and contribution to the Copernicus Emergency Management Service (CEMS). The DSS exposes public user-facing interfaces, the Data Stores, for those Services - Climate Data Store (CDS) for C3S, the Atmosphere Data Store (ADS) for CAMS and the Early Warning Data Store (EWDS) for CEMS. DSS throughout its publicly available interfaces addresses a diverse community of users, including policymakers, businesses and scientists.

DSS has recently completed a modernization process inheriting the vision, architectural principles and functional scope of the former Climate and Atmosphere Data Store (CADS) platform whilst strongly relying on open-source, cloud-oriented and cutting-edge technologies. Modernization was cornerstone to overcome the obsolescence of former components and strength architectural synergies with the ECMWF Software Engine (ESEE).

Designed as a distributed system and open framework, the DSS encompasses different architectural layers (data, software, interfaces, cloud, external) with strong technical and functional interdependencies among them. The deployed platform (software layer) underpins interactive web portals and programmatic API-based interfaces (interfaces layer) providing access to a broad catalogue of heterogenous datasets (data layer), interactive applications and other digital assets. The use of Open-Source and Standards allow to serve users directly but also to federate with external platforms which are powered by the catalogued content and the functional range of capabilities offered by the DSS portfolio of data and services. The DSS, together with other services such as the European Weather Cloud (EWC), deploys and run on the ECMWF in-house Common Cloud Infrastructure (CCI) – (cloud layer) which is physically hosted at ECMWF Data Centre in Bologna.

ECMWF together with EUMETSAT, Mercator Ocean International and the European Environment Agency (EEA) also implements the Data and Information Access Service (DIAS) platform called WEkEO. This is a distributed cloud computing infrastructure used to process and make data generated by the Copernicus

Environmental Services (CMEMS, CAMS, C3S and CLMS) accessible to users. It also provides privileged access to derived products from these services and to all satellite data from the Copernicus Sentinels operated by EUMETSAT and ESA. Within the partnership, ECMWF is responsible for the procurement of the software to implement data access services, processing capabilities and tools.

DSS catalogued datasets are accessible in the WEkEO platform for discovery and download together with other Copernicus Data products. However, due to the nature of C3S and CAMS data (volume and dimensionality), there were certain barriers for the dynamic rendering of these products in the WEkEO viewer. To fill this gap, a contract was launched for the provisioning of DSS Serverless Data Access & Visualisation capabilities for C3S and CAMS datasets. Primarily focusing on addressing WEkEO requirements, the implemented DSS ARCO Data Lake (STAC+Zarr) supporting WMTS for WEkEO viewer has demonstrated as a valuable asset to, among others, power interactive user applications while providing efficient data access for certain typology of requirements such as timeseries or training ML models. In the current context, ARCO-based capabilities are called to play a pivotal role in the foreseen future of EU federated data platforms.

Driven by current experience and building on the existing setup, the aim of this ITT is to consolidate and expand the DSS ARCO-based capabilities, ensuring their seamless integration with the architectural components of the enhanced Data Stores Service (DSS). Outcomes of the contract subject to this ITT shall include a complementary set of ARCO capabilities that further improve data usability, optimise access for cloud environments, and support streamlined integration with modern processing frameworks and federated platforms, ultimately improving performance, interoperability, and the overall user experience for Copernicus users.

**Vision:** Building on the increasing demand and demonstrated value of the initial implementation of ARCO-based data and services in the DSS, the main objective of this ITT is to consolidate and broadly extend the portfolio of ARCO-based capabilities that seamlessly integrates across the various architectural layers and components of the enhanced Data Stores Service (DSS).

More detailed information about ECMWF-Copernicus services can be found in the following links:

| **Service** | **Link** |
|---|---|
| Copernicus Climate Change Service (C3S) | https://climate.copernicus.eu/ |
| Climate Data Store (CDS) | https://cds.climate.copernicus.eu/ |
| Copernicus Atmosphere Monitoring Service (CAMS) | https://atmosphere.copernicus.eu/ |
| Atmosphere Data Store (ADS) | https://ads.atmosphere.copernicus.eu/ |
| Emergency Management Service (CEMS) | https://emergency.copernicus.eu/ |
| Early Warning Data Store (EWDS) | https://ewds.climate.copernicus.eu/ |
| WEKEO (Copernicus DIAS platform) | https://wekeo.copernicus.eu/ |

## 1.2 Glossary

In this document the following definitions shall apply for the purpose of this ITT. Where there is a difference between the definitions below and the definitions in Volume I of the ITT documents (Instructions and Conditions), the definitions below shall take precedence.

| Name | Definitions |
|---|---|
| ADS | Atmosphere Data Store |
| API | Application Programming Interface |
| Application | An interactive web page that displays maps, graphs and/or textual information that are the results of computations performed on the data and products of the Data Stores. |
| ARCO | Analysis-Ready, Cloud Optimized |
| Broker | A middleware software component which will schedule and forward requests from the web portal to remote repositories. |
| CDS | Climate Data Store |
| CDM | Common Data Model |
| CIM | Content Integration Manager |
| CLMS | Copernicus Land Monitoring Service |
| Data | The raw data collected. |
| Data Repository | A generic name for a system that holds data and/or products. This can be a database, a collection of files, etc. |
| Data Supplier | An organisation that makes its data and products available through the Climate Data Store. |
| DSS | Data Stores Service encompassing all architectural layers (data, software and interfaces) |
| EQC | Evaluation & Quality Control |
| ESGF | Earth System Grid Federation |
| FAIR | Acronym for Findable, Accessible, Interoperable, Reusable |
| Metadata | Descriptive information about the data. |
| Product | A derived, valued added piece of information, usually generated from raw data. |
| Product Catalogue | A list of available products. |
| Tool | A software that performs computations on some input data or products and produces some output data or products. |
| URL | Uniform Resource Locator. |
| Users | The internal and external Users of the C3S infrastructure. |
| UX | User eXperience |
| Web portal | The web interface of the CDS that enables the users to view information, access/perform tasks on the Products Catalogue. |

| | |
|---|---|
| WEkEO | Data and Information Access Services (DIAS) Platform implemented by ECMWF in partnership with EUMETSAT, Mercator and EEA on behalf of the European Commission. |
| Workflow | A series of invocations of software tools whereby the output of the preceding tool becomes the input for the one which follows it until the required processing chain is completed. |

## 1.3 Background

### 1.3.1 DSS Architecture

The architecture of the current Data Stores Service (DSS) encompasses different layers which have strong technical and operational interdependencies between them. ARCO-related capabilities spread with components across all different layers.

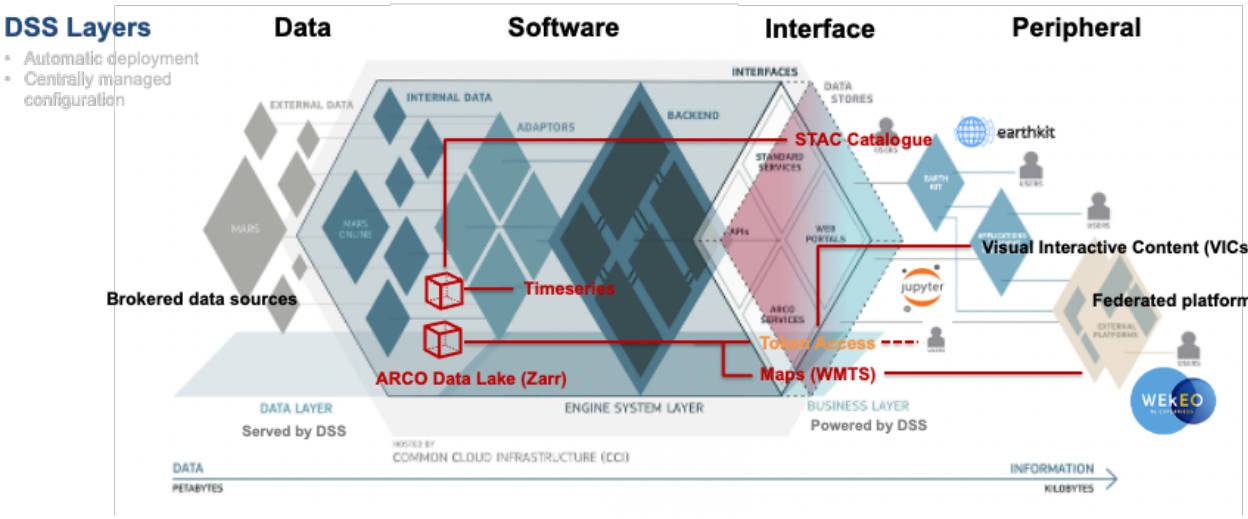


*Figure 1: Diagram of the different layers that integrate the Data Stores Service (DSS)*

- **Data layer:** Foundational pillar of the Service as it encompasses all the different data sources which are made available to users by the rest of layers and components. Depending on where data repositories are physically hosted in reference to DSS hosting cloud infrastructure (ECMWF Common Cloud Infrastructure - CCI), these are considered:
  - **External Data** repositories (hosted outside the CCI), ranging from federated and brokered datasets supported by third party organizations to the ECMWF MARS archive, where flagship datasets, such as Global Reanalysis, are primarily hosted.
  - **Internal Data** (hosted within CCI). The most representative is the on-line MARS where a subset of the most requested variables from the core ECMWF MARS archive are regularly uploaded. It also includes the Modernized Observations Repository, other small on-disk datasets, and the experimental ARCO Data Lake.

- **Software layer (core engine):** Set of Software components which deploy and run together to support the functional scope of DSS offering a seamless user journey for searching, discovering, sub-setting and retrieving data via interactive and programmatic interfaces. Software layer components integrate within the overarching ECMWF Software EnginE (ESEE). Functional scope of this layer covers the following:

  - Oversees the full life cycle of incoming data requests.

- Hosts the catalogues, including datasets metadata and configuration.
- Implements quality of service rules (QoS) to regulate user traffic and system resources.
- Dispatch requests to the Data Layer via dedicated adaptors.
- Exposes interactive and programmatic interfaces (APIs) highly configurable and compliant with standards.

- **Interface Layer (by-service instances):** Comprises operational deployments of the Software Layer, tailored by-service branding, configuration, and catalogued content. This layer includes both interactive web portals and programmatic access interfaces. Wherever feasible, interfaces are aligned with standards that support the adoption of FAIR principles (Findable, Accessible, Interoperable, Reusable). For example, catalogues are made publicly available via OGC API - Records, STAC, and OGC CSW.

Currently the DSS exposes the following web portal interfaces:

| Copernicus Service | Data Store Instance | Link |
|---|---|---|
| C3S | Climate Data Store (CDS) | https://cds.climate.copernicus.eu/ |
| CAMS | Atmosphere Data Store (ADS) | https://ads.atmosphere.copernicus.eu/ |
| CEMS | Early Warning Data Store (EWDS) | https://ewds.climate.copernicus.eu/ |

The deployment of operational clusters of the DSS Service (eg. Production, Test or Development) is done on top of Kubernetes as the container orchestration engine. The Kubernetes setup supports dynamic scaling of cluster components, allowing them to scale up or down as needed. Each DSS cluster consists of a shared backend that supports multiple environments and web portals which are defined at configuration level. Software code, catalogued content and configuration are hosted and managed in shared repositories on a version control platform, facilitating collaborative work between ECMWF and different partners and supporting automatic deployments.

Monitoring and metrics are cornerstone for DSS Operation. Infrastructure components across layers are tracked to guarantee its correct functioning, performance and reliability. Collected inputs are used among other purposes to raise alerts, troubleshoot issues, regulate system traffic, capacity planning and at higher level get operational statistics for reporting and insights to support informed decisions.

On the periphery of the core DSS layers, but closely interlinked with them, seats a series of components which are "powered by" and/or “complementary” to DSS. Of special relevance are the following:

- **Earthkit:** ECMWF open-source Python code repository offering a broad set of expert libraries optimised to work with ECMWF and DSS resources. Earthkit-data, the data handling component of earthkit, is a format-agnostic interface for geospatial data making it simple to read, inspect and slice from a wide range of input types such as grib, netcdf and ARCO complaint formats such as Zarr as an emerging capability.

| Service | |
|---|---|
| Earthkit’s documentation | https://earthkit.readthedocs.io/en/latest/ |
| Earthkit open code on Github | https://github.com/ecmwf/earthkit |

- **Visual Interactive Content (VICs):** this includes a broad set of user-oriented applications and training material that showcase data and services making combined use of the full range of resources and capabilities of the DSS Data Layer and earthkit. The underlaying DSS ARCO Data Lake leverages VICs functionality and responsiveness.

| Service | Link |
|---|---|
| Example application: ERA5 Explorer | https://era-explorer.climate.copernicus.eu/ |

- **External platforms:** Platforms and infrastructures which are "powered by" the catalogued content or any other available functional capability provided by DSS and earthkit. These platforms mostly interact with or consume data resources via the publicly available interfaces, embed technical components or integrate VICs as part of their portfolio. The WEkEO DIAS Platform, a partnership of ECMWF, EUMETSAT, Mercator Ocean and EEA, can be referred as the most representative example of an external platform "powered by" the Data Stores.

### 1.3.2 ECMWF Software EnginE (ESEE)

ECMWF has long been at the forefront of global weather prediction and atmospheric research, with a strong tradition of developing and maintaining high-performance computing systems and specialized software for atmospheric and climate modelling. In this context the ECMWF Software EnginE (ESEE) encompasses the broad family of software products and libraries designed, developed and managed by ECMWF. ESEE has evolved driven by community requirements, growing data volumes, increasing model complexity, and the emergence of new paradigms such as Machine Learning and high-throughput data workflows.

The DSS Software layer piggy-back on ECMWF's extensive in-house knowledge and experience running operational systems and developing expert software. DSS convergence with ESEE ensures operational compatibility, reusability of code, co-evolution and co-development of software libraries and the alignment of architectural guidelines and principles. All this lays the foundation for a more unified, functional, maintainable and scalable platform for serving Copernicus data to users. It also establishes the groundwork for the federation with other EU platforms such as WEkEO, DestinE or emerging Data Spaces.

As part of the ECMWF Software EnginE (ESEE), earthkit plays a key role in supporting DSS with libraries that integrates as building blocks across multiple system components to facilitate efficient data access, transformation, and integration workflows.

| | |
|---|---|
| **earthkit** | https://earthkit.readthedocs.io/en/latest/ |

Earthkit is designed to accelerate weather and climate science workflows by simplifying tasks such as data access, processing, analysis, and visualisation. It offers a modular suite of interoperable components built on trusted Python libraries such as NumPy, Pandas, and Matplotlib, while also integrating smoothly with ECMWF's operational software stack, including ecCodes, FDB, and other core infrastructure used within DSS.

Its high-level, user-friendly interface makes earthkit suitable for a wide range of users across ECMWF and the wider community. With scalability and operational readiness in mind, it helps bridge the gap between research and production, enabling the Data Stores Service to provide more flexible and efficient tools for accessing and using data.

The project is currently in BETA version, with several components available through PyPI and GitHub. As development continues, earthkit capabilities evolve to support emerging needs and a deeper integration with DSS, visual interactive content (VICs) and ARCO-based capabilities.

The Successful Tenderer will be required to use earthkit libraries and to actively contribute to its open-source code development for ARCO related capabilities in alignment with ECMWF software practices.

### 1.3.3 WEkEO

As part of the European Union's Copernicus Programme, ECMWF, together with EUMETSAT and Mercator Ocean International, have joined forces to implement a Data and Information Access Services (DIAS) Platform called WEkEO.

| WEkEO DIAS platform | https://www.wekeo.eu/ |
|---|---|

WEkEO is a distributed cloud-computing infrastructure used to process and make the data generated by the Copernicus Environmental Services (CMEMS, CAMS, C3S and CLMS) accessible to users. It also provides privileged access to derived products from these services and to all satellite data from the Copernicus Sentinels operated by EUMETSAT and ESA.

Within the partnership, ECMWF is responsible for the procurement of the software to implement Data Access Services, Processing and Tools. ECMWF developed the requirements for the software and awarded and managed a contract for the implementation of those requirements to a European provider of software engineering services. The services have been integrated with the delivery platform by the partner EUMETSAT and their suppliers. The service is now live on the Internet.

The delivery platform allows end-users to access a substantial portfolio of original Copernicus Programme and Sentinel satellite data using a harmonised data access mechanism and common data catalogue. Users can harness compute resources to process that data without the networking and storage costs associated with public cloud offerings. They have access to a wide range of tools and technologies in the areas of DevOps, including data access, artificial intelligence, machine learning, workflow, compilation, build and more. The portfolio of CDS, ADS and CEMS datasets is available through the WEkEO platform to allow bespoke processing of that data in unison with all other Copernicus data using a common data access mechanism.

WEkEO is a federated platform that regularly harvest Data Stores catalogues metadata via the programmatic standards interfaces, reuse configuration files to recreate data subsetting functionalities, forms and constraints, and access source data via APIs with dedicated Quality of Service rules. WEkEO also integrates earthkit as data access, manipulation and visualization tools.

The Successful Tenderer shall support the operational integration of the ARCO-based capabilities with WEkEO and other federated platforms, throughout the enhancement, implementation and maintenance of programmatic interfaces, standard services, operational monitoring and management capabilities

### 1.3.4 Status of DSS ARCO Capabilities

One of the most prominent user interfaces provided by the WEkEO platform is the Data Viewer (https://wekeo.copernicus.eu/). Due to the size, nature and archive format of C3S and CAMS datasets, the rendering and visualization of this data requires of dedicated architectural components to provide responsive and highly efficient WMS/WMTS services. For that purpose, a Serverless Data Access & Visualization platform was procured setting up the foundation for the ARCO capabilities on DSS.

Main output of this contract has been the implementation of an operational ARCO Data Lake supported by the following components provisioned as part of its operational capability:

- **Data Handling Layer:** Software components that handle the operational data streams and workflows for the extraction, transformation, and loading (ETL) of data from source repositories into the final ARCO format.
- **Data Repository:** A cloud-native data lake (S3 Object Store) that stores the curated ARCO datasets, using Zarr for data and STAC for metadata.
- **Standard Services for Visualization:** High-performance WMTS service that enable dynamic map tiles rendering of ARCO datasets in the WEkEO Data Viewer.
- **Monitoring & Reporting:** Operational dashboards providing real-time status, workflow metrics, and performance indicators (KPIs) about the different components.

- **Expert software:** Software libraries and service components that facilitate the management, accessibility and usability of the ARCO Data Lake.

The portfolio of the ARCO Data Lake is progressively being extended towards the aim of including a sub-extraction of the majority of the C3S and CAMS catalogued data. An up-to-date vision of the datasets currently available in the ARCO Data Lake can be obtained by checking in the WEkEO viewer the C3S and CAMS datasets which offer layers for visualization.

In addition to the DSS ARCO Data Lake and the supporting software infrastructure delivered by this initial contract, ECMWF is broadly working on ARCO related architectural components and software packages within the ECMWF Software EnginE (ESEE) and in contributions to other EU funded activities such as DestinE. On its current implementation, the DSS ARCO Data Lake is functionally and operationally detached from other ECMWF related activities and operational flows. One of the main objectives of this ITT is to strength these synergies and integrate the software components of the ARCO platform within the DSS and ESEE frameworks leveraging a more integrated and advanced Data Service.

The Successful Tenderer shall reuse, extend, and/or reengineer the existing ARCO-based components ensuring a seamless integration and/or transition to new or evolved components, while assuring the operational availability and the continuous update of the current ARCO Data Lake and all services relying on it.

Close collaboration between the Successful Tenderer, ECMWF and third parties will be required to define the architectural requirements of the evolving platform, ensuring full compatibility and alignment with the architectural principles of both the Data Stores Service (DSS) and the ECMWF Software EnginE (ESEE).

# 2 Contract Definition and Requirements

## 2.1 Vision and Objectives

**Vision:** To enhance the Data Stores Service (DSS) with advanced ARCO-based capabilities that integrate seamlessly across the architectural layers, software stacks, and operational methodologies of the existing DSS infrastructure and the ECMWF Software EnginE (ESEE). Delivered capabilities shall improve data accessibility and usability, enable optimised use of data in cloud environments, and support streamlined integration with modern, high-performance data processing frameworks and platforms. As extension of the current DSS offering, the enhanced ARCO-based capabilities shall improve the efficiency of resources and leverage the final user experience by providing an additional level of data abstraction in cloud-optimised and interoperable format.

The objectives of this contract can be summarised as:

- Deliver enhanced, ARCO-based capabilities and data services that improve data accessibility, usability and performance.
- Enable efficient, cloud-native data access which is optimised for large-scale, multidimensional, and heterogeneous datasets, including support for sparse data structures.
- Support further interoperability with modern data processing frameworks and federated platforms, ensuring consistency and alignment with other ECMWF and EU funded initiatives.
- Introduce an additional layer of data abstraction that allow users to access data through efficient, standardised, cloud-optimised formats (Zarr preferred but others to be considered), and discovery metadata (STAC).
- Ensure the operational management of the underlaying ARCO Data Lake, including workflows for data ingestion, monitoring and reporting tools, users access control, and actualization of catalogued metadata records.
- Provide an extended range of ARCO user-oriented tools, services and interfaces (e.g., interactive widgets in download forms, token management, standard APIs, QoS) that shall leverage user experience and streamline data access for different user profiles.
- Maintain close collaboration with ECMWF to ensure technical compatibility and architectural alignment with the DSS and ESEE (ECMWF Software EnginE) software stacks.

This ITT aims in summary to deliver operational, plug-in, and FAIR (Findable, Accessible, Interoperable, Reusable) ARCO-based capabilities that complement the evolving needs of the Data Stores Service (DSS) users from different domains and levels of expertise.

## 2.2 Statement of Requirements

### 2.2.1 General Requirements

The following sections detail the project requirements, summarise the architectural context and the functional scope of the envisaged ARCO-based capabilities to be considered in the context of this contract.

Enumerated requirements describe the baseline of how ECMWF envisages the ARCO-based capabilities to be implemented. They are by no means a constraint for the solutions to be proposed by the Tenderer but shall serve as a guide to define the scope of the proposal. In some instances, alternative solutions and approaches may exist to address the exposed requirements. ECMWF would like the Tenderer to submit details of such alternatives as part of their response to this ITT.

Being at the core of the technical implementation of the Copernicus Services, DSS has strong interlinkages with other Copernicus Services, EU platforms and funded projects. During the execution of this contract, interaction with different teams at ECMWF and external third parties will be required. ECMWF will facilitate interaction at all levels whenever this may be required.

The Tenderer is expected to propose a flexible work plan and implementation methodology that supports effective collaboration between different stakeholders and allows flexibility to adapt the scope and implementation roadmap in alignment with ECMWF and EU priorities and milestones.

Integrated within the DSS, all delivered components will be hosted, deployed and run, within the in-house Common Cloud Infrastructure (CCI) provided by ECMWF and physically located in Bologna (Italy). ECMWF will facilitate the computing and storage resources required for the development and operational implementation of the system and will provide support with technical requirements at Cloud Infrastructure level.

### 2.2.2 Architectural Requirements

The architectural requirements of the delivered capabilities are expected to be fully aligned with those of ECMWF Software EnginE (ESEE), Data Stores Service (DSS) and the current DSS-ARCO implementation to guarantee seamless integration and full compatibility with the existing Software and Hardware assuring the operational continuity of the ARCO capabilities already in place.

The delivered architecture and system components shall be service-oriented, open, and fully aligned with the existing ECMWF and DSS architectural frameworks, software stacks, and operational methodologies, ensuring delivered outputs can be seamlessly integrated as plug-in extensions and/or be interfaced via programmatic standard interfaces. The design of delivered software components will be extensible, supporting scalability in terms of concurrent user access, data volumes, and processing capacity. All components will be designed to leverage the full potential of the underlying cloud infrastructure, ensuring elasticity, resource efficiency, and adaptability to future requirements, growth and demand.

The envisaged architectural components can be assimilable to the following functional levels:

- **Data Handling:** Architectural components that support the implementation and operational management of workflows responsible for extracting data from raw data sources, performing the necessary transformations, and loading the resulting datasets into the destination Data Repository in Cloud-Optimized (ARCO) formats. These components must ensure data integrity, enable automation of processes, and provide full visibility and control to system operators.
- **Data Repository:** A cloud-native Data Lake, in where data is loaded, stored and kept updated at a higher level of abstraction in Cloud-Optimized (ARCO) formats. Metadata describing available data in the Data Lake is indexed by using a Spatio-Temporal Asset Catalogue (STAC), enabling broader interoperability and efficient data discovery, filtering, and access. The repository must ensure scalability, data consistency, and support for both current and future dataset integrations.
- **Standard Services:** Primary interfaces through which ARCO Data Lake assets are made accessible to external platforms, applications, and users. This architectural layer is responsible for delivering data and metadata through open, standardised, and programmatic interfaces, ensuring interoperability, scalability, and seamless integration within federated data ecosystems.
- **User oriented tools and functionalities:** Interaction layer between the users and the ARCO Repository, providing functional tools and utilities that facilitate the discovery, configuration, and access of ARCO resources from different functional approaches. This layer will ensure that users can seamlessly interact with the underlying Data Lake without requiring deep understanding about the data formats or access protocols which are behind.

- **Operational Monitoring and Reporting:** Components at this level collect and make actionable real-time operational information for DSS administrators and operators. Outputs of this component will be used for operational monitoring and alerting purposes, generation of statistical analysis for reporting, tracking of KPIs and capacity planning.

The provided components are expected to maximize the capabilities of the underlying cloud infrastructure.

The proposed system architecture will address the following requirements:

| Ref. | Requirement | Remarks / Suggestions |
|---|---|---|
| 1. | Architectural alignment with ECMWF and DSS. | Architecture principles and recommendations, FAIR, coding standards, CD/CI pipelines. |
| 2. | Cloud-native architecture, Cloud technology oriented | Object Store, Kubernetes |
| 3. | Modular and Service Oriented design | |
| 4. | Scalable components | |
| 5. | Serverless (or hybrid) architecture | |
| 6. | Cloud-optimized (ARCO) formats | STAC, Zarr as baseline but others to be considered which are suitable for multidimensional /sparse data. |
| 7. | Cloud Platform agnostic | |
| 8. | Must run on a Linux/UNIX operating system | |
| 9. | Documentation is provided | Admin, user |
| 10. | Automatized deployment | Git, Ansible |
| 11. | Standards compliance | OGC or other de facto standards (eg. STAC). |
| 12. | Data Integrity & Traceability | Quality checks |
| 13. | Federation readiness | WEkEO, DestinE, Data Spaces |
| 14. | Seamless User Experience (UX) across DSS | Widgets, workflow editors |
| 15. | Real time observability - monitoring and statistics | Splunk, Opsview |

**The Tenderer must include in their response how their proposed architecture ensures seamless integration with DSS and the scalability of components.**

#### 2.2.2.1 FAIR Principles

DSS has a strong commitment to the adoption of FAIR principles (Findable, Accessible, Interoperable, Reusable) at all different levels. As one of the FAIR pillars, Interoperability is key for the evolution of the service and the ARCO-based capabilities.

- **Findability:** Datasets ingested into the ARCO Data Lake will be catalogued using a structured Spatio-Temporal Asset Catalog (STAC) framework, enriched with comprehensive metadata that adheres to

DSS configuration, templates and extensions. Metadata will be available via APIs**,** enabling users and external systems to efficiently navigate data assets (based on spatial, temporal, and thematic criteria) and automatically harvest catalogued records.

- **Accessibility:** Data and metadata are exposed through open, well-documented APIs that follow OGC standards (Features, Tiles) and RESTful practices. Data will be served in cloud-optimised formats such as Zarr, enabling efficient partial data access without the need for full dataset downloads. Access control will be based on DSS authentication mechanisms, supporting secure data access (token) and Quality of Service Rules (QoS).
- **Interoperability**: The implemented platform components expose open standards and protocols, ensuring internal (across DSS and ECMWF components) and external (public interfaces for federated platforms and users) interoperability. The use of harmonised data formats, metadata schemas, and programmatic APIs will allow ARCO Datasets to seamlessly integrate with analytical workflows (ML/AI/LLM), Visual Interactive Content (VICs such as viewers and applications), and federated platforms and ecosystems (WEkEO, DestinE, etc.).
- **Reusability**: Data Lake assets will be available for access and use subject to terms and conditions, licenses, citation, and usage guidelines which align with those of the raw data sources. The converted data will ensure scope alignment, consistency and versioning. Furthermore, ARCO datasets will be structured and formatted to facilitate direct use reducing the need for redundant preprocessing or post-processing.

Adoption of FAIR principles will allow DSS ARCO-based capabilities to enrich the service portfolio of data and services providing additional **robust, scalable, and user-centric data access capabilities** that maximises the long-term usability, accessibility, and scientific value of ECMWF and Copernicus data.

| Ref. | Requirement | Remarks / Suggestions |
|---|---|---|
| 16. | The Successful Tenderer must ensure the adoption of FAIR principles | Data, Metadata |
| 17. | Findability | STAC, APIs |
| 18. | Accessibility | Open, documented interfaces. Data formats (Zarr, COG). Authentication and authorization. |
| 19. | Interoperability | OGC, de-fact standards, programmatic interfaces. |
| 20. | Reusability | Licenses, Citation, Versioning |

#### 2.2.2.2 Open-Source Software

The Tenderer shall propose software solutions which are the most practical and cost effective for ECMWF and DSS intended use, scope of work and future evolution and maintenance. ECMWF expects the Tenderer to propose and be in condition to collaboratively contribute to open-source software, whenever this is available and suites to deliver a practical solution.

ARCO-based platform and components are preferred to be based on open-source software. Software tools, libraries, and development tools shall, whenever available and fit-for-purpose, be based on widely adopted open-source licenses (e.g., MIT, Apache 2.0). Similarly, the Tenderer shall rely on existing open-source projects and libraries (e.g., STAC, GDAL, Dask, Xarray, OpenLayers) to avoid vendor lock-in.

The delivered software components must facilitate the reuse, contribution, and sharing of code, adhering to established open-source development practices, including version control (e.g., Git), documentation, and issue tracking.

The Tenderer is invited to provide a tentative list of:

- Open-Source software proposed for use.
- Technologies per component.
- Programming languages that will be used.
- Any proprietary or non-Open-Source item that may be used.

If a proposed solution would require the modification of open-source software, ECMWF and the Successful Tenderer will first consult about the permissions or restrictions imposed by the relevant open-source licence and the risk that the modified software will become obsolete and unsupported over time.

Ownership of all rights in the software solutions, which the Successful Tenderer develops for ECMWF, shall be assigned, through ECMWF, to the European Commission (EC).

Technological choices will be based on existing open-source software whenever this is available, Python being the preferable programming language. Acceleration and integration with other languages is also relevant (e.g. Rust).

| Ref. | Requirement | Remarks / Suggestions |
|---|---|---|
| 21. | Open-Source Software | |
| 22. | Reuse of existing software and libraries | STAC, Xarray, Git, DSS |
| 23. | Python based development for ad-hoc components | Preferred language |
| 24. | Contribution and shared code | Version control, documentation, modularity |
| 25. | Modularity | Plug-in services |
| 26. | Maintainability and evolution | ECMWF, Community based |

#### 2.2.2.3 Supported Data

DSS Catalogued Datasets, split on distributed and heterogeneous sources, vary greatly in volume and formats. Different data sources are accessed via dedicated protocols and adaptors, facilitating users a simplified data access experience by using harmonized and standard interfaces.

DSS Broker dispatch incoming data requests to the distributed data repositories and manage traffic and workload via the Quality of Service rules (QoS). As part of the data access and retrieve process, DSS provides with additional functionalities for pre- and post-processing of data allowing among other things statistical aggregations, area sub-setting or format conversions before final delivery to optimize the use available resources.

The delivered platform and software components are expected to take full advantage of the capabilities offered by the DSS infrastructure when accessing raw data sources and be on itself flexible enough as to handle the wider diversity of datasets.

Each Dataset might present particularities on typology, scope, volume and format. Delivered solutions shall be flexible to handle different datasets in origin and support the implementation and configuration of ad-hoc data ingestion workflows which may include pre- and post- processing.

Some of the facts to be considered when evaluating the feasibility, scope and required effort for the integration and suitability of datasets to be handle under the ARCO-based capabilities are:

- Data format: Most common data formats available on CDS and ADS are WMO GRIB edition 1 and 2 and NetCDF conforming CF-1.6 or greater. Nevertheless, other formats may be also found.
- Number of variables: Datasets can cover a wide diversity of variables and sub-setting criteria.
- Data volume: Raw data volumes range between gigabytes to petabytes and several billion products.
- Data bridges: Best optimal ways and interfaces to access and extract data from original sources.
- Frequency of updates: Datasets are not static and are subject to regular incremental updates and releases of new versions.
- Grade of Popularity: User statistics provide insight about the preferences and profiles of users accessing datasets

| Ref. | Requirement | Remarks / Suggestions |
|---|---|---|
| 27. | Support ingestion for different archive data formats | GRIB, NetCDF |
| 28. | Heterogenous datasets supported | Sparse, irregular grids |
| 29. | Cloud-optimized ARCO formats | Zarr, COG |
| 30. | Metadata alignment with DSS catalogue | STAC, Temporal and Geographical updates, dimensions, updates, enriched metadata |
| 31. | Support for partial data access | Spatial sub-setting, time slicing, variables |

**Proposed solution will allow for technical adaptation and evolution to support the integration of a growing diversity of datasets.**

**The Tenderer is highly encouraged to explore the different DSS catalogues (CDS, ADS, EWDS) to evaluate the envisaged casuistic and potential burdens that the proposed solution might present for integration of the current data portfolio.**

### 2.2.3 System Requirements

**As part of the proposal, the Tenderer is expected to present the software architecture and functional capabilities of the envisaged ARCO-based capabilities describing the different components that would support their implementation and describe how they would interact among them and with the rest of DSS components. The Tenderer is expected to provide with use cases that illustrate how the proposed functionalities will improve the DSS data uptake and user experience when accessing data.**

**In those cases that proposed components are intended to build, extend or reengineer existing ones, the Tenderer is expected to provide arguments that support such decision, justify the scope of the proposed changes and detail the special requirements (e.g. GPUs) that may be needed to support such changes.**

**Interaction between different software components shall be ruled by clearly stated Interface Control Documents (ICD).**

**Project milestones will demonstrate functional integration with WEkEO viewer and potentially other platforms.**

**ECMWF will facilitate whenever possible the data access mechanisms to optimize the use of resources and avoid competition with DSS operational instances.**

#### 2.2.3.1 Data Handling

The purpose of the Data Handling component is to manage the end-to-end data workflows from the raw data sources into the final Data Repository (ARCO Data Lake). This set of components orchestrate the Extraction, Transformation and Loading (ETL) processes, ensuring that raw DSS datasets are ingested, processed, and delivered in the destination ARCO format. This set of components oversees the regular updates of data assuring integrity and consistency with the source.

Workflows might consist of different concatenated jobs such as data extraction, transformation and ingestion/creation of analysis-ready, cloud-optimised data formats in the destination ARCO Data Repository. Execution of workflows is orchestrated by a dedicated backend components and can be planned to be launched automatically or manually.

The Tenderer is expected to describe the different architectural components that will support the implementation of the required workflows, the scheduling and orchestration of jobs and the operational management and monitoring capabilities to handle their execution. This set of components must be flexible, modular and extensible enabling adaptation to different typologies of data.

Given the heterogeneity and complexity of raw data sources, the number and complexity of workflows is expected to grow. The proposed architecture shall provide optimal ways to manage these workflows in the long term. ECMWF aims to have full access rights and control for the management and operation of the implemented workflows. Detailed documentation is then required to be delivered both for the administration of the system and for the operation of implemented workflows.

Workflows will rely on ECMWF Software EnginE (e.g. earthkit), DSS components (e.g. API, adaptors) to access the data. ECMWF will work together with the Successful Tenderer on defining the most optimal ways for accessing, retrieving and processing raw data sources minimizing the concurrence and competition with DSS operational traffic.

Workflows are expected to support parallelization and optimize the consumption of resources across all different steps (jobs), from data access to computing and ingestion.

Management and Monitoring tools in the Data Handling component will allow operators of the system to:

- Create and modify workflows and jobs.
- Tune configuration.
- Schedule executions and manage queues.
- Monitor status and outputs.
- Access logs and debug issues.
- Get statistics and KPIs.

| Ref. | Requirement | Remarks / Suggestions |
|---|---|---|
| 32. | System allows for the implementation of ad-hoc workflows per dataset. | |
| 33. | Workflows can be scheduled, adapted and configured. | By the operational administrators |
| 34. | System guarantees the integrity and consistency of data. | |

| | | |
|---|---|---|
| 35. | Workflows can be monitored. | By the operational administrators |
| 36. | Documentation is provided for the administration and operation of workflows. | |

#### 2.2.3.2 Data Repository

The ARCO Data Lake serves as an additional Data Repository within the DSS Data Layer, providing an abstraction data layer in a cloud-native format that simplifies and optimises certain data access requirements. Designed to store and deliver complete datasets, partial subsets, derived and ad-hoc products in ARCO formats, it enables scalable, high-performance access for analytics, modelling, and visualisation use cases. Supporting large-scale, multidimensional, and heterogeneous datasets, the Data Lake leverages formats like Zarr for arrays and Cloud Optimized GeoTIFF (COG) for raster data, allowing users to efficiently retrieve spatial, temporal, or variable subsets without downloading unnecessary portions of data.

To accommodate diverse usage patterns, the Data Lake shall implement dimension-based chunking strategies, optimised for both time-focused and spatial-focused analyses. Metadata will be structured using the Spatio-Temporal Asset Catalog (STAC), ensuring efficient discovery and interoperability through standard APIs. The architecture must support horizontal scalability, maintain data consistency and integrity, and allow seamless integration of incoming datasets. Standard interfaces will ensure low-latency access, responsive visualisation, and close-to-source processing via internal platforms as the DSS JupyterHub.

Data access protocols shall allow to define ingestion and consumption mechanisms, supporting user profiles with differentiated access privileges and usage policies, governed through an overarching access control layer.

The ARCO Data Lake will facilitate serverless or hybrid deployments, elastic scalability for high-resolution data, and granular usage monitoring.

| Ref. | Requirement | Remarks / Suggestions |
|---|---|---|
| 37. | Cloud-native, scalable architecture | Object Storage |
| 38. | Allow partial data access | Subsetting, chunking |
| 39. | ARCO Data is expected to be optimized for multiple consumption profiles | |
| 40. | Data Access protocols for ingestion and consumption are described | S3, HTTP |
| 41. | Expose compliant metadata | STAC |
| 42. | Optimized data formats | Zarr, COG, others to be considered. |
| 43. | Assure data integrity, consistency and versioning | |
| 44. | Implement access control and policies | Token access, QoS |

#### *2.2.3.3 Standard Services*

The Standard Services layer of the ARCO-based capabilities serves as the primary interface through which data and metadata of assets available in the ARCO Data Lake are made accessible to external platforms, applications, and users. This layer is responsible for delivering data through open, standardised, and programmatic interfaces**,** ensuring interoperability, scalability, and seamless integration with federated

platforms and data ecosystems. Datasets shall be exposed for external consumption using standard services and protocols, including among others the SpatioTemporal Asset Catalog (STAC) API for metadata discovery, OGC APIs (Features and Tiles) for spatial data access, and the Web Map Tile Service (WMTS) for efficient, tile-based visualisation. These standard services support primarily the integration with platforms such as the WEkEO Viewer**.**

A core requirement of this layer is the implementation of a high-performance WMTS service**.** This WMTS service must support multiple resolutions, offer tile sets in various map projections, and render tiles in PNG format with transparency**,** ensuring flexibility and responsiveness for user interactions. The WEkEO platform will interact with this service both at backend level, periodically harvesting information about viewable layers, and at frontend level, where user actions such as adding layers, panning, or zooming in the map viewer will trigger direct tile requests from the WMTS service.

The Standard Services layer must ensure high availability and performance, supporting concurrent access at scale, while adhering to open standards that guarantee internal and external interoperability. It must also be extensible, allowing for the seamless integration of new services and data formats as the DSS evolves, and enforce secure and controlled access aligned with the DSS Quality of Service (QoS) rules (permission, limits and priorities).

By exposing data through these cloud-native, standard-based services, the ARCO platform aims to enhance the usability, reach, and operational efficiency of DSS data portfolio, fostering wider adoption across diverse user communities and application domains.

The Tenderer is invited to propose the use of emerging data formats, standard services, or technologies that can further enhance the platform with cloud-native analytics and on-demand processing capabilities.

| Ref. | Requirement | Remarks / Suggestions |
|---|---|---|
| 45. | Datasets are made available via responsive WMS/WMTS that cover the full extend (variables and dimensions) | |
| 46. | WMTS integrates with WEkEO viewer maximizing user capabilities | Temporal and geographical navigation, zooming. |

**Implemented service must be compliant with OGC specifications for WMTS services as described here: https://www.ogc.org/standard/wmts/**

**Integration of WMTS with WEkEO Viewer is at the core of the requirements of the provided infrastructure.**

#### 2.2.3.4 User tools and functionalities

The **user tools and functionalities layer** represents the interaction point between users and the ARCO-based capabilities within the DSS, providing intuitive interfaces and utilities that enable users to configure and manage their access to assets in the ARCO Data Lake. This layer shall also offer users a comprehensive set of tools and functionalities that facilitate the entire data access journey across the DSS platform, from discovery and exploration to configuration, access, manipulation, and visualisation. This architectural layer abstracts

the complexity of underlying data formats, access protocols, and infrastructure, enabling users to seamlessly interact with the underlaying ARCO Data Lake without requiring advanced technical expertise.

Core components of this layer may include:

- Token access management interfaces for users to request, manage, and monitor their access credentials in line with defined permissions and data access policies.
- Interactive widgets, embedded within the DSS Web Portal, widgets allow users to explore ARCO Data Lake holdings, apply spatial and temporal filters, preview, use and retrieve data.
- Earthkit extensions, enabling users to interact programmatically with ARCO data by leveraging the earthkit suite of tools to facilitate efficient data access, transformation, and analysis directly within their own workflows and computational environments applications.
- Additionally, the layer may support user-driven workflows configuration, enabling authorised users to define and execute constrained ad-hoc data transformation workflows that generate ephemeral data products or virtual data cubes.

This layer must align with DSS architectural design, principles and user experience, maintaining consistency across the platform, while delivering responsive, scalable interfaces that address different user needs. Full integration with backend access control mechanisms is required, ensuring that all user interactions are governed by data access policies and Quality of Service (QoS) rules. The design must also ensure extensibility to allow for future enhancements, including new user-facing functionalities as data services and user requirements evolve.

In addition to core tools and functionalities, the Tenderer is welcome to propose further capabilities that may range from advanced catalogued records exploration**,** new standards to be explored or potential amendments that might allow a better uptake of the DSS data in highly-demanding frameworks, or plug-ins for software packages (e.g., GIS tools) and federated platforms such as WEkEO, DestinE, and emerging Data Spaces**.** By bridging the gap between the ARCO Data and potential user communities, this layer will play a critical role in ensuring accessible, efficient, and user-centric interaction with the DSS data and services.

| Ref. | Requirement | Remarks / Suggestions |
|---|---|---|
| 47. | User interfaces integrate with DSS components | UX, widgets, APIs |
| 48. | Tokenize access to Data Lake | |
| 49. | Earthkit extensions for ARCO Data | |
| 50. | Access control and Quality of Service (QoS) | Permissions, priorities, limits |
| 51. | Scalable and low latency interfaces | |
| 52. | Interfaces based on OGC or other de facto standards | OGC APIs (eg. EDR), STAC |

#### 2.2.3.5 Monitoring and Metrics

The Monitoring and Metrics component is a cornerstone layer of the system that shall ensure the operational observability, reliability, and continuous tuning and improvement of the ARCO-based capabilities within the DSS. This component shall provide accurate and up-to-date information on the operational status and performance of all system components, enabling proactive management, troubleshooting, capacity planning, and user activity tracking. It will also play a vital role in fulfilling internal and external reporting obligations, including Key Performance Indicators (KPI) in compliance with European Commission requirements.

The Monitoring and Metrics component shall integrate seamlessly with the broader DSS Observability ecosystem, aligning with existing operational tools, methodologies, and best practices implemented by ECMWF and DSS. The system shall rely on tools which are currently in use such as Grafana, Prometheus, Elastic, Splunk or Opsview for system health monitoring and alerting, ensuring a consistent and centralised operational approach.

The proposed solutions are expected to encompass the following capabilities:

- Log system activities**,** capturing detailed information for diagnostics, troubleshooting, and supporting capacity planning through integration with DSS existing log management solutions.
- Supervise data integration workflows**,** ensuring the correct execution, data integrity, and consistency between ingested datasets in the ARCO Data Lake and their respective source repositories.
- Monitor system health and raise alerts**,** integrating automated checks and probes into the existing monitoring framework to provide real-time visibility about the status of all ARCO components.
- Track system usage and activity metrics**,** including user access patterns, workflow executions, and data retrieval statistics, with the ability to translate these into ECMWF/DSS defined KPIs.
- Ability to collect usage statistics to full fill reporting obligations towards the European Commission and third parties.
- Ensure that collected monitoring data can be utilised to enhance dynamic system configuration and user-oriented functionalities**,** providing feedback loops for automated adjustments and improvements.

The Monitoring and Metrics Component must be designed to support operational continuity and provide insights for service improvements through detailed observability across the ARCO-based components. It will be essential for guaranteeing service availability, identifying bottlenecks, optimising resource usage, and maintaining high-quality service delivery within the DSS operational framework.

This component will cover the following baseline functionality:

| Ref. | Requirement | Remarks / Suggestions |
|---|---|---|
| 53. | Ability to create health checks of the system components to the Operators. | To be integrated with Prometheus, Opsview. |
| 54. | Ability to track execution of data integration workflows. | |
| 55. | Ability to generate system and application logs. | To be integrated with Elastic, Splunk. |
| 56. | Ability to conduct audits. | |
| 57. | Ability to manage faults and record errors. | Self-explanatory and human readable messages. |
| 58. | Ability to minimize downtime during upgrades. | |
| 59. | Ability to provide system metrics based on KPIs. | |

# 3 Contract Structure and Implementation

## 3.1 Contractual Approach

As stated in previous sections, the final objective of the Framework Agreement and associated Service Contracts subject to this ITT is to enhance the Data Stores Service (DSS) with advanced ARCO-based capabilities that seamlessly integrates across the architectural layers, software stack, and operational methodologies of the existing DSS infrastructure. Delivered capabilities will improve data accessibility and usability, enable optimised use of data in cloud environments, and support streamlined integration with modern, high-performance data processing frameworks and platforms. As an extension of the current DSS offering, the enhanced ARCO functionality shall improve resource efficiency and leverage user experience by introducing an additional level of data abstraction in a cloud-optimised and interoperable format.

The future Framework Agreement and associated Service Contracts subject to this ITT will be complementary to the current DSS ARCO set up, aiming to consolidate and extend the operational and functional capabilities of the current offering and set up the basis for further evolution towards broader federation of Cloud-based services. Among the core baseline objectives of the future Framework Agreement and associated Service Contracts subject to this ITT, an objective is to continue supporting the visualization of C3S and CAMS data products on the WEkEO Viewer via standard WMS/WMTS data services.

The proposed approach for the Framework Agreement and associated Service Contracts subject to this ITT is driven by the following objectives:

- ECMWF's vision and architectural principles are aligned with those of the Tender.
- Knowhow, experience and current set up from the current ECMWF and DSS ARCO implementations are taken onboard in the form of functional requirements and technical specifications.
- Requirements and expectations of main stakeholders are properly addressed and handled as baseline drivers for requirements.
- Functional integration and interfaces with external platforms such as WEkEO are properly established.
- Project roadmap is aligned with those of ECMWF, DSS and WEkEO.

To achieve the above-mentioned objectives, it is proposed to implement within the Framework Agreement and associated Service Contracts subject to this ITT an approach with different phases and/or areas of work:

- **Preliminary analysis phase (under SC1):** During this phase, the Successful Tenderer will interact with ECMWF and other parties to define the functional scope, architecture, and technologies to be used and to define the development and implementation strategy, roadmap and work plan. Also during this phase, the Successful Tenderer will be requested to define and execute a plan for knowledge transfer with the current contractor in charge of DSS ARCO activities to understand the operational functioning and technical set up of the current ARCO implementation.
- **Transitional maintenance, adaptation and operational uptake (under SC2):** Upon completion of the handover of the operational ARCO-based capabilities performed in close collaboration with the contractor formerly in charge, the Successful Tenderer will gradually start the required actions for the migration, adaptation and takeover of operational processes and workflows ruling the assets in DSS-ARCO Data Lake and their final integration via standard WMTS services in the WEkEO viewer. This phase will temporally overlap with the former ARCO contract. Tenderer will be expected to take full ownership over the operational ARCO capabilities at the end of this overlapped period with the support of the former contractor.
- **Development and implementation phase (under SC2):** After agreement and approval of the functional scope, architectural design and technical framework coming out from the initial analysis

phase (SC1), the Successful Tenderer will start the gradual implementation of agreed requirements. This phase might imply the following activities:

- Configuration and tuning of the underlaying cloud infrastructure to support the deployment of the dedicated ARCO components and allow for its efficient and correct functioning.
- Development and implementation of software components.
- Grant of access rights, privileges, and permissions for those required components not directly managed by the Successful Tenderer.
- Integration, configuration and adaptation of DSS to the newly developed components.
- Validation and testing of delivered components and final outputs.

Some of these activities may involve different internal and external stakeholders. ECMWF will facilitate communication and coordination between different actors at all levels. ECMWF will also support the Successful Tenderer to implement those technical requirements required by the system which may lay out of the scope of the contract and depend on ECMWF or other external contractors.

- **Data Integration (under SC2):** The purpose of this phase is to complete the adaptation of remaining data workflows to the new framework and to uptake the implementation of new ones to continue increasing the ARCO Data Lake portfolio of available data assets.

  Workflows will rely for their implementation and operational functioning on the combined capabilities of ECMWF-ESEE, DSS, existing ARCO suites and those delivered in the context of the contract during the development and implementation phase. ECMWF will work together with the Successful Tenderer on a case-by-case to define the most optimal ways for accessing, retrieving, processing and storing data minimizing the concurrence and competition with DSS operational traffic and assuring alignment between raw and ARCO data.

  This area of work might imply the following activities:

  - Identify target datasets to be integrated.
  - Define the scope (subset) of data assets to be integrated.
  - Evaluate the optimal configuration of ARCO formats.
  - Implement ad-hoc jobs and workflows.
  - Define data integrity checks.
  - Schedule the execution of workflows according to raw data updates.
  - Provide detailed documentation for operators and system administrators.
  - Adapt the monitoring tools to track the workflows executions.
  - Validate final outputs of the workflows:
    - Showcase serverless access to the ARCO data and metadata.
    - Prototype the Standard services on top of the ARCO Data Lake.
- **Maintenance & Support (under SC2):** This area of work will cover the outputs of previous phases for both Software components and Data Lake assets. This area might imply the following activities:
  - Perform corrective maintenance and support to guarantee operational continuity.
  - Generate incident reports and root cause diagnosis.
  - Perform preventive maintenance actions to avoid failures or service degradation.
  - Support integration and software updates.
  - Evolutive maintenance of implemented workflows.

- Advice on the reusability of existing components.

**ECMWF proposes to apply the Agile methodology as reported in Annex 4 of the Volume V Framework Agreement template part of this ITT. Nevertheless, any other suggestion from the Tenderer in terms of methodology to apply for each of the different phases of the project, based on their knowledge and experience with projects of a similar nature, is welcomed. The methodologies proposed by the Tenderer must ensure that the Deliverables and Milestones are fit for purpose, aligned with the project vision and remain within project budget and schedule. The Tenderer must provide in any case examples of how they have applied this (these) methodologie(s) in similar projects they have previously worked on.**

## 3.2 Implementation Schedule

ECMWF intends to award a **Framework Agreement** for a total period of 24 months, which shall be implemented via **two Service Contracts**:

- The **Service Contract No 1 (SC1)** shall be composed of **WP0 and WP1**. The SC1 is expected to commence in Q2 2026, ideally 01/06/2026, and last four (4) months, therefore until 30/09/2026.
- Since the SC1 shall allow ECMWF and the Successful Tenderer to agree on the exact content of the SC2, a period of one (1) month is expected between the end date of SC1 and the start date of SC2 to allow the signature of the SC2 and, if necessary, to perform all necessary amendments of the SC2 as originally agreed (e.g. Pricing Tables and List of Deliverables and Milestones parts of the Framework Agreement Annex 2, see also Clause 2.1.2.1 (i) of the ITT Volume V) and formalise them thanks to an amendment agreement to the Framework Agreement and the Service Contract(s).
- The **Service Contract No 2 (SC2)** shall be composed of **WP0, WP2, WP3 and WP4**. The SC2 is expected to commence in Q4 2026, ideally on 01/11/2026, and last nineteen (19) months, therefore until 30/06/2028.
- **Important**: For the sake of clarity, the Parties will first sign only the Framework Agreement and its SC1. The signature of the SC2 will be fully depending on the outcomes of the implementation of the SC1 and their formalisation, if necessary, through an amendment agreement of the contractual documents originally agreed between ECMWF and the Successful Tenderer.

  If ECMWF cannot sign the SC2 with Successful Tenderer, a restricted new ITT will be launched (CJS2_220b_bis2). This new ITT will be open only to the Tenderers who were shortlisted in this ITT CJS2_220b_bis (those having achieved a technical score above 60%). The Successful Tenderer of CJS2_220b_bis will not be able to participate to the ITT CJS2_220b_bis2.

The Tenderer shall provide a detailed implementation plan of proposed activities for the total period.

## 3.3 Work Packages

The following sections describe the different Work Packages (WP) subject of this ITT.

### 3.3.1 WP1: Preliminary Analysis of the ARCO-based capabilities to be implemented

This work package aims to carry out a detailed functional and technical analysis of the ARCO-based capabilities to be implemented prior to the development phase. This phase is expected to be run jointly between ECMWF staff and the Successful Tenderer's team in a series of iterations (i.e. workshops and meetings) to be agreed. For certain components and actions, interaction with third parties will be required.

The final objectives of this phase are the following:

- Define the architecture principles and functional scope of the components to be implemented.
- Produce user stories/use cases, to define the structure and functional scope of the proposed components.
- Perform a gap analysis of DSS-ARCO capabilities currently in place and identify areas for further improvement and development.
- Define the software components and technologies to be implemented.
- Define the development strategy and implementation roadmap.
- Define the ICDs that will drive the integration between infrastructure components.
- Define the work methodology and set up the required tools.
- Identify key stakeholders, roles and responsibilities.
- Contextualize requirements to be aligned with the ECMWF landscape of ARCO related initiatives.
- Define communication flows and channels.
- Set up the working tools and environments.
- Define roadmap and core objectives for the next project milestones.
- Collect and review existing documentation.

Dedicated objectives for the transitional maintenance of the current DSS-ARCO infrastructure:

- Knowledge transfer and handover of the current system components.
- Run the required knowledge sharing and handover sessions jointly with the former ARCO contractor.

**Important**: All Deliverables as result of this phase must be approved by ECMWF, unless otherwise agreed, as conditional requirement before triggering the development phase (WP2), migration (WP3) and maintenance phase (WP4) as it will be agreed between the parties under the frame of the Service Contract No 1.

A close and iterative collaboration between ECMWF, the Successful Tenderer (the contractor) and designated third parties is envisaged during this entire phase through a series of events as described in the following table. Duration and number of participants is estimative just for budget allocation purposes:

| Scope | Participants | Duration | Objectives | Location |
|---|---|---|---|---|
| Initial workshop | ECMWF, contractor | To be agreed<br>Initial estimation 1 week with up-to 3 persons attending | • Introduce the project teams<br>• Define the requirements and expectations of the project<br>• Perform a preliminary exchange of knowledge and ideas<br>• Define a first draft of architectural components (WPs)<br>• Plan further actions | ECMWF premises |
| Initial hand-over workshop | ECMWF, contractor, former ARCO contractor (CJS2_220b) | To be agreed (potentially as part of the Initial workshop) | • Introduce the current ARCO set-up.<br>• Knowledge transfer of the technical components and operational actions.<br>• Review of existing documentation.<br>• Hands-on sessions. | ECMWF premises |

| | | | | |
|---|---|---|---|---|
| Work meetings | ECMWF, contractor, third parties | To be agreed<br>Initial estimation up-to 6 hours of remote meetings involving up-to 4 participants. | • Follow up the business and technical analysis.<br>• Technical deep-dives and clarification sessions.<br>• Progress of work and status of deliverables.<br>• Plan further actions | Virtual |
| Presentation of results | ECMWF, contractor | To be agreed.<br>Initial estimation 3 days with up-to 3 persons attending. | • Present the outputs of the work as described in the objectives described above<br>• Plan further action | ECMWF premises |

Final scope and agenda for this series of events shall be agreed at a later stage during the contract negotiation taking account of the Tenderer's proposal.

Expected deliverables for this phase / WP1 are, as a minimum:

| **Nature** | **Title** |
|---|---|
| Report | Business Analysis and technical architecture |
| Report | Interface Control Documents (ICD) |
| Report | Work plan and implementation roadmap (including concerning the migration and adaptation activities subject to WP3 ) |
| Report | Best and final offer about Service Contract No 2 |

Reference material for this phase / WP1 will be:

- Solutions proposed by the contractor as response to this ITT.
- Current DSS ARCO-based infrastructure and available documentation.
- Inputs and requirements collected from key stakeholders.

### 3.3.2 WP2: Development and implementation of an incremental set of ARCO-based components

The final approach and scope of this work package will be driven by the outputs and agreements produced by the preliminary analysis phase (WP1). The scope of work under this WP2 covers the development and implementation of an incremental set of ARCO-based components that gradually integrates in the operational ecosystem of the Data Stores Service (DSS).

It may be the case that delivered components rely on third party actions or developments to be released into operations. In those cases, strong collaboration and interaction are expected between ECMWF, the Successful Tenderer and third parties to align the requirements and coordinate development, deployment, test and release roadmap. The work methodology proposed by the Tenderer shall provide with the flexibility to manage these coordinated actions and adapt the roadmap for certain software deliverables to those of

third parties whenever dependencies are identified. ECMWF will facilitate the coordination and communication activities between the different parties involved.

The final objectives of this work package are the following:

- Implement an iterative/continuous integration, delivery and deployment approach (CI/CD).
- Conduct continuous iteration, consultation and revision of requirements with key stakeholders.
- Produce early prototypes to demonstrate functionality (whenever required).
- Produce the envisaged software components.
- Conduct component, unit, system, and integration testing.
- Deliver the components and perform acceptance tests.
- Deliver all required documentation.

Internal organization of this work package can be split in different sub-packages if this serves the project to shorten the roadmap, deliver individual components or make a more efficient use of resources.

Given the developmental nature of this phase, ECMWF anticipates that the adoption of a CI/CD (Continuous Integration and Continuous Development) approach will be beneficial to ensure its successful delivery. This would typically involve a design framework comprising of:

- Clarification of requirements
- Analysis of solutions.
- Implementation plan
- Development
- Validation and test
- Deployment (release)

The development strategy and implementation roadmap will define a detailed schedule of the activities to be carried out within the contract including activity planning and prioritization of requirements, analysis, deep dives, reviews, deliverables and milestones, validation and deployment.

Taking account of the proposed Agile approach to be implemented, it is expected that iterative development phases will deliver incremental working versions of the building blocks of the envisaged components which will combine software deliverables, documentation, test, acceptance criteria and implementation plan to be agreed between the Successful Tenderer and ECMWF and, in some cases, with third parties as well.

The Kanban approach is recommended but ECMWF welcomes suggestions from the Tenderer on what agile methodology they propose to apply based on their knowledge and experience with software projects of a similar nature. To this regard, the Tenderer must provide examples of how they have applied such approach in similar contracts they have previously worked on. The development methodology proposed by the Tenderer must ensure that the Deliverables are fit for purpose and remains within the contract cost and schedule.

The Successful Tenderer will be required to discuss and iterate on the design/architecture with ECMWF and third parties throughout the lifetime of the development phase (WP2). Due to the dynamic nature of the Copernicus program and Services, incoming requirements are expected during the project.

### 3.3.3 WP3: Migration and adaptation of existing ARCO-related data workflows and functionalities

Depending on the scope, technologies and design of the newly delivered components, the Successful Tenderer shall support the adaptation, migration, or reengineering of existing ARCO-related data workflows and functionalities under this WP3**.** Triggering of this work package will be subject to the functional scope of the final delivered components and the potential reusability of existing components. These WP3 activities will be agreed, planned and executed in agreement with ECMWF and can be triggered in parallel of the implementation phase (WP2).

**Important**: Whatever actions are to be taken, these shall not create service disruption or functional regression on existing services and always be carried out assuring operational continuity as a priority with none or minor disruption to ongoing operational services.

Whenever necessary to facilitate the migration and adaptation activities, the Tenderer will be required to organise dedicated knowledge transfer, planning and deployment sessions with the contractor responsible for the current ARCO implementation while benefiting from ECMWF's support to this regard. These sessions will serve to gather detailed information on functional dependencies, understand operational behaviour, and clarify technical requirements relevant for the adaptation and migration process. Overall plan for the migration and adaptation phase shall be based on the agreed plan on SC1 (cf. deliverable "Work plan and implementation roadmap" part of WP1 activities).

Expected activities may include:

- Assessment of existing ARCO workflows and their alignment with the new components.
- Refactoring or redeployment of workflows to run under the new technological environment.
- Migration or reingestion of data and metadata assets to the enhanced ARCO Data Lake.
- Ensuring uninterrupted operational switch and integration with existing platforms such as the catalogued timeseries entries or the WEkEO Data Viewer**.**
- Integration with existing automation, monitoring and statistical tools.
- Implementation of test suites and check procedures to ensure quality, traceability, and consistency.

All such activities must be designed to complement the newly delivered components and ensure continuity, stability, and improved performance across the DSS ARCO-based capabilities.

### 3.3.4 WP4: Operational maintenance and support

This work package covers the operational maintenance and support of the components delivered under this contract, as well as any migrated or adapted workflow or operational functionality. The aim is to ensure continuous availability, reliability, and performance of the deployed components in alignment with the rest of DSS operational processes. The Tenderer shall define, together with ECMWF and during the preliminary analysis phase / WP1, the work methodology, communication flows and operational Service Level Agreements (SLAs) that shall apply from the identification to the final resolution of operational bugs and issues.

The operational maintenance phase / WP4 will include regular interaction with ECMWF to review the status of components, perform wash-ups, and define mitigation or corrective actions to be implemented. The Successful Tenderer will also contribute to maintaining accurate and up-to-date system.

Tasks under this work package include**:**

- Monitor the operational status and performance of delivered and migrated components.
- Implement corrective actions for identified bug, issues and failures in a timely manner in accordance with the agreed SLAs.
- Run wash-up sessions, provide retrospective reports and facilitate recipes or workarounds for operators.
- Handle the implementation and deployment of system updates, bug fixes, and minor enhancements upon the final solution of identified issues.
- Support the operational continuity of services provided to external platforms (e.g. WEkEO viewer).
- Supervise access control and security mechanisms in line with defined policies;
- Supporting capacity planning through the analysis of monitoring metrics and usage patterns.
- Document operational changes, configurations, and known issues for future reference.

The operational maintenance must ensure high availability, reliability, and security of services, contributing to the long-term sustainability and evolution of the DSS ARCO-based capabilities.

### 3.3.5 WP0: Management and coordination

The WP0 "Management and coordination" shall ensure the correct overall implementation of the Framework Agreement and associated Service Contracts during their whole Term and shall be further described and completed, if necessary, by the Tenderer in its technical proposal. The WP0 tasks are the following ones:

**Task 0.1: Service management and coordination**

| Ref. | Requirement |
|---|---|
| **[R01.01]** | To appoint a Service Manager with experience of managing projects and contracts of this type and size. This person will be the main point of contact for administrative matters. Amongst others this person shall ensure the implementation of all contractual obligations as described in the ITT Volume V and all necessary **day-to-day service management and coordination** including the planning and monitoring of all Work Packages activities and corresponding resources. This person shall be involved in the activities of this ITT at a minimum level of 10% of their total working time, unless otherwise agreed. |
| **[R01.02]** | To appoint a Technical Lead[1] with more than 5 years of experience in technical activities related to this ITT that shall ensure that all technical requirements are met. This person will be the point of contact on technical matters and shall be involved in the activities of this ITT at a minimum level of 10% of their total working time, unless otherwise agreed. |
| **[R01.03]** | To ensure **quality assurance and control activities**, including in what concerns the Sub-contractors' activities if any. The final quality check of all deliverables (contents, use of ECMWF's templates for deliverables and reports, format, deliverables/milestones numbering and naming, typing errors, etc.) shall be made by the Successful Tenderer. The Tenderer shall provide a list of its quality assurance processes and management systems and if applicable, any quality related accreditations or certifications it holds. |
| **[R01.04]** | To enforce a proactive **risk management**, including in what concerns the Sub-contractors' activities if any. The technical proposal shall therefore include a risk register which shall be updated regularly by the Successful Tenderer. Any update (related to new risks, likelihood or impact) shall be reported during the PRM as well as be part of the QIR and AIR. |
| **[R01.05]** | To ensure proactive and dynamic **internal and external communication**, towards and between all parties involved in the contract. |

[1] The role of Service Manager and Technical Lead can be performed by the same person (in this case to be considered as the Service Manager only) if duly justified. In that case a second contact person needs to be appointed as alternative contact.

| | |
|---|---|
| **[R01.06]** | To **manage the personal data** in accordance with the ITT Volume V Clause 2.8 and its Annex 6 “Personal Data Protection”. |
| **[R01.07]** | To **manage the Sub-contractor(s)**, if any, in accordance with the ITT and annexes, including dispute resolution (the Successful Tenderer will be responsible for settling disagreements, although advice/approval from ECMWF may be sought on the subject). The Tenderer shall also describe in its Tender how the ITT Volume V, in particular Clause 2.9 “Sub-contracting”, has been flowed down to all its Sub-contractors. |
| **[R01.08]** | The Successful Tenderer shall **support ECMWF in its general communication activities** for the C3S/CAMS services, where they are related to the activities described in this ITT. |

⇨ The Tenderer shall outline and justify the proposed overall management methodology for this contract in its technical proposal.

**Task 0.2: Reporting activities for the benefit of the European Commission**

| Ref. | Requirement |
|---|---|
| **[R02.01]** | To ensure the timely and quality delivery of the Deliverables as stated in ITT Volume IIIA and in accordance with ITT Volume V, especially in its Clause 2.3 "Reporting and Planning", necessary to ECMWF's **reporting towards the European Commission**. |

**Task 0.3: Meeting management and coordination**

| Ref. | Requirement | | |
|---|---|---|---|
| **[R03.01]** | To organise and/or attend the following meetings: | | |
| **Meeting** | **Objective** | **Date/Frequency, format and location, duration** | **Expected attendance (<u>as a minimum</u>)** |
| Kick-Off Meeting (KOM) | (All) To introduce all involved parties, remind the requirements, report on the first efforts performed and next steps to implement. | • Once (max. T0+1)<br>• Teleconference<br>• 2 to 3 hours | • ECMWF<br>• Prime contractor's Service Manager and Technical Lead<br>• Sub-contractors' representatives |
| Progress Review Meeting (PRM) | (Contractor) To report on the progress accomplished.<br>(All) To take actions and/or make decisions for the sake of the contract implementation. | • Quarterly<br>• Teleconference<br>• 1 hour | • ECMWF<br>• Prime contractor |
| C3S General Assembly (GA) | (All) To present C3S activities and service provision and perform side-meetings if need be. | • Annually<br>• On site attendance<br>• Europe<br>• Ca. 2-3 days | • ECMWF<br>• Prime contractor |
| CAMS General Assembly (GA) | (All) To present CAMS activities and service provision and perform side-meetings if need be. | • Annually<br>• On site attendance<br>• Europe<br>• Ca. 2-3 days | • ECMWF<br>• Prime contractor |
| Ad-hoc meetings | (All) To solve any issue faced during implementation and/or take actions and/or make decisions for the sake of the contract implementation. | • On ECMWF and contractor's request<br>• Teleconference | • ECMWF<br>• Prime contractor<br>• On-request attendants |
| <u>Note</u>: If considered necessary, the Tenderer can propose additional project meetings, whose added value must nevertheless be precisely substantiated. | | | |

The Tenderer shall therefore incorporate the following Deliverables and Milestones in the relevant spreadsheet part of the ITT Volume IIIA. All Deliverables and Milestones shall be numbered as indicated below.

| ***Deliverable / Milestone ID*** | ***Deliverable / Milestone title*** | ***Due date*** | ***Main contributor*** | ***Nature*** |
|---|---|---|---|---|
| D0.1-QIR-YYYYQQ | Quarterly Implementation Report YYYY QQ<br>*YYYY QQ being the previous quarter (i.e. either Q1, Q2 or Q3)* | Quarterly on 15/04, 15/07 and 15/10 | Tenderer | Report |
| D0.2-AIR1-YYYY | Annual Implementation Report Part 1 YYYY<br>*including both:*<br>• *the Quarterly Implementation Report YYYY Q4 and* | Annually on 15/01 | Tenderer | Report / Other |

| | | | | |
|---|---|---|---|---|
| | • *the requested preliminary financial information YYYY*<br>*YYYY being the Year n-1* | | | |
| D0.3-AIR2-YYYY | Annual Implementation Report Part 2 YYYY<br>*YYYY being the Year n-1* | Annually on 28/02 | Tenderer | Report |
| D0.4-AIP-YYYY | Annual Implementation Plan YYYY<br>*YYYY being the Year n+1* | Annually on 30/09 | Tenderer | Report |
| D0.5-FS-YYYY | Copy of Prime Contractor's general financial statements and audit report YYYY<br>*YYYY being the Year n-1* | Annually, not later than on 15/12 | Tenderer | Other |
| D0.6 | Final Implementation Report | End date of the contract, only once all other activities have been duly performed | Tenderer | Report |
| D0.7[2] | Evidence of or a declaration about the Contractor's insurance | Max 30 days after start of contract | Tenderer | Other |
| M0.1 | Kick-Off Meeting | Max 30 days after start of contract | Tenderer | Presentation and MoM |
| M0.2-PRMxx | Progress Review Meeting No xx<br>*x being the number of the PRM* | ~ every 3 months | Tenderer | Presentation and MoM |
| M0.3-C3S-YYYY | C3S General Assembly YYYY<br>*YYYY being the Year of the GA edition* | Annually, not later than on 15/12 | Tenderer | Attendance |
| M0.3-CAMS-YYYY | CAMS General Assembly YYYY<br>*YYYY being the Year of the GA edition* | Annually, not later than on 15/12 | Tenderer | Attendance |

**⚠ Important note ⚠**

Given the proposed structure of the Agreement, i.e. two Service Contracts under a Framework Agreement (see also section 3.2 above), some specific aspects must be considered by the Tenderer in its offer regarding the WP0 activities:

- Each Service Contract shall include the WP0 activities and shall include in its respective List of Deliverables and Milestones only the Deliverables and Milestones that are due during the term of the concerned Service Contract.
- Only one KOM (cf. Milestone M0.1) is required and shall take place under the frame of the SC1 only.
- Only one FIR (cf. Deliverable D0.6) is required to be delivered under the Framework Agreement, therefore a FIR shall first be added to the List of Deliverables and Milestones of both Service Contracts No 1 and No 2. In case the parties agree to sign the Service Contract No 2, then the FIR part of the list of Deliverables and Milestones of SC1 will be cancelled and its price deducted.
- The Milestones M0.3-C3S-YYYY and M0.3-CAMS-YYYY corresponding to the C3S and CAMS General Assemblies shall only be considered under the frame of the SC2.
- The evidence of or a declaration about the Contractor's insurance (cf. Deliverable D0.6) shall be considered under each SC since the Overall Price may evolve depending on the Parties' agreement concerning the SC2 (see also Clause 2.1.9.2 in the ITT Volume V).

[2] This Deliverable is not necessary under SC1 provided it has been provided before the Effective Date of the Agreement, see also Clause 2.1.9.4 in the ITT Volume V.

## 3.4 Project Team

### 3.4.1 Tenderer

The Tenderer shall demonstrate the availability of expertise as required for the implementation of the services in line with the components and requirements stated in this ITT Volume II.

The Tenderer shall demonstrate for itself, and for any proposed sub-contractor that they have participated in national or international research and/or private sector software integration projects in the last 5 years for the activities subject of this ITT. ECMWF may ask for evidence of performance in the form of certificates issued or countersigned by the competent authority.

The Successful Tenderer will appoint a **Technical Lead**, responsible for the delivery of the system, to oversee the technical progress of this project, as well as a **Service Manager**, that will represent their interests in day-to-day discussions and meetings and will oversee all contract management activities.

The Tenderer must outline the project team in its Tender. The outline must contain the following:

- Relevant experience of key staff and management personnel.
- Names of Technical Lead and Service Manager, and main technical contact and number of work hours dedicated to the project team for the duration of the project.
- The Tenderer must state if there are any sub-contractors and define their roles.
- An indication of how many staff members will be part of the project team and at what level.

The project team assigned to this contract is expected to demonstrate:

- solid knowledge in:
  - Web technologies.
  - Service Oriented Architectures (SOA).
  - Open source.
  - Containerization and orchestration technologies.
  - Geospatial data handling within a Linux and Python environment.
  - Advanced Python.
  - Data structures and algorithms.
  - NetCDF and GRIB formats.
  - OGC standards (CSW, WMS, WMTS, Zarr).
  - Cloud technologies.
  - UML and Business architectures.
- experience on running large scale systems preferably related with the domain (Meteorological, Climate, etc.)
- adequate understanding of the current Data Stores Service (DSS) and WEkEO DIAS Platform.

The Successful Tenderer's project team is expected to work very closely with ECMWF's team for the whole duration of the contract.

### 3.4.2 ECMWF

ECMWF will appoint a **Product Owner** and a **Technical Lead** to oversee the development and deliverables. The Product Owner will be ECMWF's technical point of contact for the Successful Tenderer. The Product Owner will:

- Monitor the Successful Tenderer's work execution.

- Review the Successful Tenderer's specifications and architectures to ensure that they are "fit for purpose".
- Be the focal point to provide the Successful Tenderer with ECMWF's inputs required at each stage.
- Facilitate interaction with third Parties.
- Validate and prioritise the requirements list.
- Agree time boxing priorities.
- Be the focal point to support the incremental testing of each iterative phase.
- Sign off key milestones and deliverables.

Furthermore, ECMWF will appoint a **Contract Management Officer** to oversee the implementation of the contract, the approval of Deliverables and achievement of Milestones. The Contract Management Officer will be ECMWF's contractual point of contact for the Successful Tenderer and will represent ECMWF's interests in day-to-day discussions and meetings.

ECMWF's team will attend project meetings as deemed necessary for the monitoring of the Successful Tenderer's activities and will be granted unrestricted access by prior agreement to the Successful Tenderer's facilities where the work is being carried out.

## 3.5 Deliverables and Milestones

### 3.5.1 General guidance about Deliverables and Milestones

The Tenderer shall use the ITT Volume IIIA "Pricing and deliverables" spreadsheet "Deliverables List" to provide the list of Deliverables and Milestones for each WP while considering the following:

- A **Deliverable** is a substantial, tangible or intangible good or service produced as a result of a project (see also the deliverable definition in this ITT Volume V Clause 1.2 and Clause 3.2). In other words, a Deliverable is a verifiable outcome produced in response to the specific objectives of the contract and is subject to approval by both ECMWF's Technical Officer (TO) and Contract Management Officer (CMO) before being considered as contractually approved.
- A **Milestone** is designed as a marker of demonstrable progress in service development and/or quality of service delivery during the contract implementation (see also the milestone definition in this ITT Volume V Clause 1.2). The Milestones shall not duplicate deliverables.

The objectives of each Work Package are outlined in section 3.3 above. As part of the Tender, the Tenderer is expected to propose a list of Deliverables and Milestones for each Work Package which shall be consistent with the set objectives. The Deliverables can be in the form of documents or reports, data sets or databases, services, user support, etc.

The following shall apply to the Deliverables and Milestones:

- All Deliverables and Milestones must be consistent with the activities and objectives described in this ITT and shall allow the Successful Tenderer to comply with the technical requirements listed in this ITT Volume II;
- All Deliverables shall be produced in English unless otherwise expressly agreed between ECMWF and the Successful Tenderer;
- The quality of reports shall be equivalent to the standard of peer-reviewed publications and practice;
- Unless otherwise specified in the contract, or requested by ECMWF during the contract implementation, the final version of each Deliverable shall be made available to ECMWF without any comments and tracked changes, in electronic format (Microsoft Word/Microsoft Excel/HTML or compatible, PDF in case of signed version, while all other formats – if any – must be agreed during the contract negotiation or implementation) via the Copernicus Deliverables Repository portal – OpenText Core (OTC) – and any

other expressly requested platform. The Tenderer is invited to propose the use of any other platforms, servers or portal depending on the type of deliverables considered (e.g. datasets).

⇨ ECMWF will provide all necessary templates and guidance at the beginning of the contract.

The following shall be considered by the Tenderer when completing the ITT Volume IIIA "Pricing and deliverables" spreadsheet "Deliverables List" to attach to its Tender:

- All Deliverables and Milestones already listed in this ITT Volume II shall keep their ID. Additional Deliverables and Milestones, if any, shall abide by the following referencing system:
    - *Deliverables and Milestones shall respectively be numbered as per the following format Dx.y.z (for Deliverables) and Mx.y.z (for Milestones), where x is the WP number, y is the task number and z is the Deliverable or Milestone number in this task.*
    - *Deliverables delivered annually should be numbered Dx.y.z-yyyy, where yyyy is the year the Deliverable refers to (e.g. Dx.y.z-2024). Deliverables that will be delivered quarterly should be numbered Dx.y.z-yyyyQa, where Qa is the quarter of the year the Deliverable refers to (e.g. Dx.y.z-2024Q1). The same numbering format shall be applied for Milestones. Continuous deliverables at higher frequency can be labelled in the same way as quarterly deliverables.*
- Each Deliverable shall have an associated resource allocation and price (cf. column I "Nb of PM allocated" and column J "Estimated price"), while the only resource types to be considered is "payroll", "travel" and "professional fees" (the total of these allocated resources and prices shall therefore amount to the total price associated with "payroll", "travel" and "professional fees" in ITT Volume IIIA spreadsheet "Costs and Prices").
- Milestones should not have any associated resource allocation and price unless there are Deliverables associated to them such as presentations, MoMs, etc.
- Each Deliverable and Milestone shall have a realistic and therefore achievable due date.

As part of its Tender, the Tenderer shall complete the relevant table in the ITT Volume IIIA template which shall therefore include the Deliverables and Milestones already indicated in this ITT Volume II, unless otherwise substantiated.

⇨ For the sake of clarity, the Tenderer can add to the list of Deliverables and Milestones any Deliverables and Milestones that are deemed necessary to duly cover the entire scope of activities to be performed and comply with all associated requirements.

***[CONCERNING THE OPTIONAL DELIVERABLES, if any]*** Please note that the optional Deliverable(s) shall neither be delivered by the Successful Tenderer nor subject to payment unless expressly requested and activated by ECMWF in due time. In case of optional Deliverables, please:

- clearly mention "[optional]" next to the reference of each optional Deliverable in the list of Deliverables and Milestones, and
- add for each optional Deliverable or set of optional Deliverables an associated Milestone *"ECMWF's decision concerning the activation of optional [DelivRef]"* with ECMWF as main contributor with as a due date the deadline for ECMWF to activate the said optional Deliverable or set of optional Deliverables.

***[CONCERNING THE CONTINUOUS ACTIVITIES, if any]*** Please add to the list of Deliverables and Milestone all necessary Deliverables concerning the continuous activities (e.g. user support, maintenance) that will allow their formal approval by ECMWF. The period of activities covered and reported within the concerned Deliverable (technical note) must be clearly indicated in the list of Deliverables and Milestones.

IMPORTANT NOTE:

⇨ The Tenderer shall detail in its offer **any specific dependency of any of the activities, Deliverables and/or Milestones on the provision of external input/data, by ECMWF or third party, to the Successful Tenderer**, including especially the origin of the necessary input. This shall be duly accounted for in the list of Deliverables and Milestones (cf. ITT Volume IIIA "Pricing and deliverables" spreadsheet "Deliverables List" column "Comment/Means of verification") as well as in the risk register where relevant mitigation shall be envisaged (cf. ITT Volume IIIB Section 5.6).

### 3.5.2 Data and IPR

It is a condition of EU funding for Copernicus that ownership of any datasets/software developed with Copernicus funding passes from the suppliers to the European Union via ECMWF. Ownership will pass from the date of creation of the datasets/software. Suppliers will be granted a non-exclusive licence to use the datasets/software which they have provided to Copernicus for any purpose.

All software and products used by the Successful Tenderer to produce the Copernicus datasets/software will remain the property of the Successful Tenderer, except for those components which are acquired or created specifically for Copernicus purposes, with Copernicus funding, and which are separable and useable in isolation from the rest of the successful Tenderers' production system. The identity and ownership of such exceptional components will be passed to the European Union annually. The Successful Tenderer will be granted a non-exclusive licence to use them for any purpose.

For more details please refer to ITT Volume V, Clause 3.

## 3.6 Key Performance Indicators

The Successful Tenderer shall report to ECMWF on a set of SMART (Specific, Measurable, Actionable, Realistic and Time bound) KPIs suitable for monitoring the service performance. The tables below provide the minimum set of KPIs to be incorporated by the Tenderer in its Tender and shall serve as template to be used by the Tenderer to describe the KPIs, relevant for this ITT, together with performance targets, delivery schedules and explanations if needed.:

| WP(s) | KPI # | KPI title | Performance target and unit of measure | Frequency of delivery / reporting | Explanations / Comments |
|---|---|---|---|---|---|
| ***Data and service quality*** | | | | | |
| **…** | **KPI.D1** | … | … | … | … |
| **…** | **…** | … | … | … | … |
| ***User support*** | | | | | |
| **…** | **KPI.U1** | … | … | … | … |
| **…** | **…** | … | … | … | … |
| ***Contract management*** | | | | | |
| **all** | **KPI.C1** | Timely delivery for review of Deliverables and completion of Milestones | 100% of Deliverables submitted for review and Milestones passed on time during previous quarter | Quarterly via WP0 QIR and PRM | The due dates associated to each Deliverable and Milestone in the contractual list of Deliverables and Milestones shall be considered the deadlines (inclusive) for the Deliverables to be submitted for review and Milestones to be passed. |

The Tenderer may propose in its Tender additional KPIs that are assumed suitable for the activities subject of this ITT and for monitoring the following aspects:

- Code quality (performance, output, etc.);
- Service delivery;
- Contract management;
- User support;
- Contract management.

However, the Tenderer shall limit those additional KPIs to the sole KPIs whose reporting and analysis by both Parties may help to optimize the performance of the contract in case of deviation observed per comparison with the performance targets.

All KPIs shall be labelled and numbered as indicated in the tables above and additional ones, if any, shall abide by a similar coding.

During the contract implementation, all KPIs shall be duly reported by the contractor in the WP0 QIR and PRM, in accordance with their frequency of delivery.

## 3.7 Communication

Copernicus is a single brand, owned and overseen by the European Commission and entrusted to the entities delivering the six Copernicus services. ECMWF and its third-party suppliers must consequentially adhere to the Copernicus brand guidance set by the European Commission for all communication material. All communication activities considered by the Successful Tenderer and its Sub-contractors must therefore be agreed with the ECMWF Copernicus Communication team in advance. This includes, without being limited, communication planning, branding and visual style, media outreach, website and social media activity, externally facing written and graphic content and events.

If needed, the Successful Tenderer can be requested to support ECMWF in its communication activities for the C3S and/or CAMS services, where they are related to the activities described in this ITT. Examples are contributions to the C3S and CAMS website news items, C3S and CAMS brochures and flyers.

## 3.8 Price and Payment Plan

### 3.8.1 Type of price

It is envisaged to use pre-agreed (fixed) price under the frame of the Framework Agreement and its two Service Contracts. This might be subject to discussion during the negotiation phase, if need be.

### 3.8.2 Price breakdown

The price breakdown between the two Service Contracts is envisaged as follows:

| **Service Contract** | **Price Service Contract (Maximum Budget) in Euro** |
|---|---|
| SC1 | 100.000,00 |
| SC2 | 600.000,00 |
| Total | 700.000,00 |

### 3.8.3 Substantiation of the price

The Tenderer is invited to provide a short explanation about the allocation of the resources between the different WPs/activities/Deliverables (e.g. bottom-up approach), focusing on the items that require the most significant efforts. In case of pre-agreed (fixed) price, any indication about difficulty in quoting precisely some activities shall be shared as soon as possible with ECMWF.

Additional exchanges and requests for explanations concerning the efforts and corresponding prices will take place during the negotiation phase if the Tenderer is selected as the Successful Tenderer.

Additional guidance about the prices: WP0 management and coordination is expected to amount to approx. 7-10% of the total effort (person months) allocated to the entire Framework Agreement.

### 3.8.4 Travel Costs

The travel costs shall be presented in accordance with the following provisions:

Travel costs should, in principle, be based on the European Commission's calculator [Table 3: Unit cost per distance band for air or combined air/rail travel, Commission Decision C(2024)5405], and consider a daily subsistence allowance not to exceed €300.

Travel costs must reflect **estimated actual costs and must not include any profit margin**.

If the proposed travel costs deviate from these reference values, the deviation shall be clearly indicated and duly justified.

Tenderers are requested to provide a summary table as shown below as part of their Tender.

| **Type of cost** | **Route/Destination** | **Estimated number of missions** | **Estimated unit price [€]** |
|---|---|---|---|
| Travel/Subsistence | | | |
| Travel/Subsistence | | | |

ECMWF will reserve the right to re-claim any declared unspent or unaccounted budget for "Travel", as it will be described in the Annex 2 Pricing Tables of the Framework Agreement.

### 3.8.5 Payment Plan

The Tenderer shall propose a draft Payment Plan in its Tender thanks to the ITT Volume IIIA "Pricing and deliverables" spreadsheet "Payment Plan preparation":

- It is foreseen to assess the Services and Deliverables on a (Payment) Milestone basis in accordance with Clause 4.5.3.4 of the ITT Volume V. Therefore, the Payment Milestones must relate to the Deliverables and Milestones delivered during the period subject of the corresponding Payment Milestone (e.g. the payment covering the period January-June must relate to the Deliverables and Milestones whose due dates are part of the same period).
- Given the total duration of the Framework Agreement and each Service Contract, it is recommended:
    - For the SC1, to have a single payment at the end of the SC1 term, and
    - For the SC2, to abide by a circa 6-month frequency between each Payment Milestone anticipated date of completion.
    - Any other frequency can be proposed by the Tenderer but shall be duly substantiated.
- The due dates of the Progress Review Meetings shall be adjusted to ensure that each Payment Milestone has an associated PRM.
- In case of request of a payment at contract signature, please note that this must be duly and precisely substantiated by the Tenderer (e.g. in terms of necessary investment prior to implementation or during first weeks/months for ensuring the initial set up of the project). It is necessary to relate this payment to activities and prices subject to other Payment Milestones.

# 4 Tender Format and Content

General guidelines for the Tender are described in the ITT Volume IIIB. This section describes specific requirements to prepare the proposal for this particular Tender, along with guidelines for minimum content expected to be included in the proposal, additional to the content described in the general guidelines of the ITT Volume IIIB. This is not an exhaustive description and additional information may be necessary depending on the Tenderer's response.

## 4.1 Page Limits

As a guideline, it is expected that individual sections of the Tenderer's response do not exceed the page limits listed below. These are advisory limits and should be followed wherever possible, to avoid excessive or wordy responses.

*Table 1 Page limits per section*

| *Section* | *Page Limit* |
|---|---|
| Executive Summary | 2 |
| Track Record | 2 (for general) and 2 per entity |
| Quality of Resources Applied | 2 (excl. Table 1 in the ITT Volume IIIB and CVs with a maximum length of 2 pages each) |
| Technical Solution Proposed | 30 for the technical solution and Work Packages (Table 2 in the ITT Volume IIIB, the section on reference, publications, patents and any pre-existing IPR are excluded from the page limit and have no page limit) |
| Management and Implementation | 10 (excl. Table 4 and Table 5 in the ITT Volume IIIB) + 2 per each Work Package template (Table 3 in the ITT Volume IIIB) |
| Pricing table | No limitation |

## 4.2 Specific additional instructions for the Tenderer's response

The following is a guide to the minimum content expected to be included in each section, additional to the content described in the general guidelines of the ITT Volume IIIB. This is not an exhaustive description and additional information may be necessary depending on the Tenderer's response.

### 4.2.1 Executive Summary

The Tenderer shall provide an executive summary of the proposal, describing the objectives, team and service level.

### 4.2.2 Track Record

The Tenderer shall demonstrate the availability of expertise as required for the implementation of the services in line with the work package descriptions.

The Tenderer shall demonstrate for itself, and for any proposed sub-contractors that they have participated in national or international research and private sector software development projects in the last 5 years for the activities for which this ITT is proposed. ECMWF may ask for evidence of performance in the form of certificates issued or countersigned by the competent authority.

The Tenderer shall in particular demonstrate its experience in:

- Kubernetes
- UX, Front-end development

- Python
- Open source
- Data Base administration
- Geospatial standards (OGC, ISO, INSPIRE)
- AGILE development methodology
- Implementation of large operational systems

### 4.2.3 Quality of Resources to be deployed

The Tenderer shall propose a team that meets at least the following requirements:

- A Service Manager with more than 5 years of experience in managing activities related to such a large-scale contract, with experience in the appropriate delivery methodology proposed in section 3 above.
- A Technical Lead with more than 5 years of experience on performing activities related to the various aspects of this ITT.

The CV, proven track records and certifications of key individuals are required, including a brief and clear description of the role these individuals will play in the contract.

### 4.2.4 Technical Solution Proposed

The Tenderer shall give a short background to the proposed solution to demonstrate its understanding of that solution and of the C3S/CAMS context.

This section shall also especially include:

- information on any other third-party suppliers that are used as part of the technical solution, and
- a statement/matrix of compliance for each requirement formulated throughout this document, describing how the proposed solution maps to the requirements.

**Important**: Any dependencies on input data, whose origin must be specified, shall be detailed in this section of the technical proposal. They must also be accounted for in the risk register table.

### 4.2.5 Management and Implementation Plan

The Tenderer shall provide a detailed implementation plan of proposed activities for the duration of the ITT Volume V Framework Agreement (cf. section 3.2 above). As part of the general project management description, the Tenderer shall consider the elements listed in section 3.3.5 above.

# 5 Appendices

## 5.1 Standards, Protocols and APIs

| | | |
|---|---|---|
| **"AAA" / "Triple A" Accessibility** | title for compliancy with Priority 1, 2 and 3 of the Web Content Accessibility Guidelines 1.0 (WCAG 1.0) | http://www.w3.org/TR/WCAG10/ |
| **CF** | Climate and Forecast metadata conventions | http://cfconventions.org/ |
| **CSV** | Comma Separated Value | |
| **DataCite** | Digital citations to find, access and reuse data | https://www.datacite.org/ |
| **DOI** | Digital Object Identifier system | http://www.doi.org/ |
| **ECMWF-ODB** | ECMWF's Observations Database | |
| **FTP** | File Transfer Protocol | http://www.w3.org/Protocols/rfc959/ |
| **GeoJSON** | a format for encoding a variety of geographic data structures | http://geojson.org/ |
| **GeoTIFF** | file standard which allows geo-referencing information to be embedded within a Tagged Image File Format (TIFF) file (inc. Animations) | http://trac.osgeo.org/geotiff/ |
| **GRIB 1** | General Regularly-distributed Information in Binary form Version 1 | https://community.wmo.int/site/knowledge-hub/programmes-and-initiatives/wmo-information-system-wis/grib-edition-1 |
| **GRIB 2** | General Regularly-distributed Information in Binary form Version 2 | https://www.ecmwf.int/en/newsletter/175/computing/migration-grib1-grib2-preparing-ecmwf-model-output-future |
| **HDF** | Hierarchical Data Format | https://www.hdfgroup.org/ |
| **HTTP** | Hypertext Transfer Protocol | http://www.w3.org/Protocols/ |
| **INSPIRE** | Infrastructure for Spatial Information in the European Community | http://inspire.ec.europa.eu/ |
| **ISO19115** | Defines the schema required for describing geographic information and services by means of metadata. | http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=53798 |
| **ISO/AWI 19119** | Identifies and defines the architecture patterns for service interfaces used for geographic information | https://www.iso.org/standard/91679.html |
| **ISO/TS 19139-1:2019** | Geographic information — XML schema implementation | http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=32557 |
| **JSON** | JavaScript Object Notation | http://json.org/ |
| **MARS** | Meteorological Archival and Retrieval System | https://www.ecmwf.int/en/forecasts/access-forecasts/access-archive-datasets |
| **NetCDF** | Network Common Data Form | http://www.unidata.ucar.edu/software/netcdf/ |
| **OAI-PMH** | Open Archives Initiative Protocol for Metadata Harvesting | https://www.openarchives.org/pmh/ |
| **OAUTH** | open protocol to allow secure authorization in a simple and standard method | http://oauth.net/2/ |
| **WIS** | WMO Information System | http://www.wmo.int/pages/prog/www/WIS/ |
| **OGC - SOS** | OGC Sensor Observation Service | http://www.opengeospatial.org/standards/sos |
| **OGC - WCPS** | OGC Web Coverage Processing Service | http://www.opengeospatial.org/standards/wcps |
| **OGC - WCS** | OGC Web Coverage Service | http://www.opengeospatial.org/standards/wcs |
| **OGC - WFS** | OGC Web Feature Service | http://www.opengeospatial.org/standards/wfs |
| **OGC - WMS** | OGC Web Map Service | http://www.opengeospatial.org/standards/wms |
| **OGC - WMTS** | OGC Web Map Tiles Service | https://www.ogc.org/standards/wmts/ |

| | | |
|---|---|---|
| **OGC - WPS** | OGC Web Processing Service | http://www.opengeospatial.org/standards/wps |
| **OGC-WCTS** | OGC Web Coverage Tile Service | http://www.opengeospatial.org/ |
| **OGC-WMTS** | OGC Web Map Tiling Service | http://www.opengeospatial.org/ |
| **OpenDAP** | Open Source Project for a Network Data Access Protocol | http://www.opendap.org/ |
| **OpenLayers** | Open Source JavaScript library for displaying map data in web browsers | http://openlayers.org/ |
| **PNG** | Portable Network Graphics file (inc Animations) | http://www.libpng.org/pub/png/ |
| **Rasdaman** | enables Web-based geo data offerings and Big Data Analytics on multi-dimensional raster ("array") data of unlimited size | http://www.rasdaman.com/ |
| **REST** | Representational State Transfer | |
| **SensorML** | OGC standard encoding for describing sensors and measurement processes | http://www.ogcnetwork.net/SensorML |
| **SFTP** | Secure File Transfer Protocol | http://www.w3.org/Protocols/rfc959/3_DataTransfer.html |
| **STAC** | SpatioTemporal Asset Catalogs<br><br>The STAC specification is a common language to describe geospatial information | https://stacspec.org/en |
| **THREDDS Data Server** | Thematic Real time Environmental Distributed Data Services | https://www.unidata.ucar.edu/software/thredds/current/tds/ |
| **TimeseriesML** | OGC encoding standard for the representation of time series observations data | https://portal.opengeospatial.org/files/60856 |
| **UV-CDAT** | Ultrascale Visualization Climate Data Analysis Tools | http://uvcdat.llnl.gov/ |
| **WaterML** | OGC standard encoding for the representation of water observations data | http://www.opengeospatial.org/standards/waterml |
| **Zarr** | Specifications and software for storage of large N-dimensional typed arrays. | https://zarr.dev/ |

## 5.2 Software

| | | |
|---|---|---|
| **Cartopy** | Python package for advanced map generation with a simple matplotlib interface | http://scitools.org.uk/cartopy/index.html |
| **CDO** | Climate Data Operators | https://code.zmaw.de/projects/cdo |
| **ecCodes** | Package developed by ECMWF which provides an application programming interface and a set of tools for decoding and encoding messages | |
| **earthkit** | **A** new open-source Python project led by ECMWF | https://earthkit.ecmwf.int |
| **GDAL** | translator library for raster and vector geospatial data formats | http://www.gdal.org/ |
| **GI-Axe** | Brokering framework | http://essi-lab.eu/do/view/GIaxe/WebHome |
| **GI-Cat** | Broker catalogue service | http://essi-lab.eu/do/view/GIcat |
| **IRIS** | Python package for analysing and visualising meteorological and oceanographic data sets | http://scitools.org.uk/iris/index.html |
| **Leaflet** | Open Source JavaScript library used to build web mapping applications | http://leafletjs.com/ |
| **Matplotlib** | a python 2D plotting library | http://matplotlib.org/ |
| **NCO** | netCDF Operators | http://nco.sourceforge.net/ |
| **NumPy** | NumPy is the fundamental package for scientific computing with Python | http://www.numpy.org/ |
| **OpenLayers** | Open Source JavaScript library for displaying map data in web browsers | http://openlayers.org/ |
| **Rasdaman** | enables Web-based geo data offerings and Big Data Analytics on multi-dimensional raster ("array") data of unlimited size | http://www.rasdaman.com/ |
| **Rust** | | https://rust-lang.org |
| **Scipy** | Python-based ecosystem of open-source software for mathematics, science, and engineering. | http://www.scipy.org/ |