ECMWF Copernicus Procurement

Invitation to Tender

Copernicus Joint Services

Volume II

Serverless Data Access & Visualisation Platform for C3S/CAMS Datasets

ITT Ref: CJS2_220b
Issued by: ECMWF Administration Department Procurement Section
Date: 2 November 2023
Version: Final

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 an operational platform for Serverless Data Access & Visualisation of C3S/CAMS Datasets on the WEkEO platform (https://www.wekeo.eu/data).

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 Volume I.

The Climate and Atmosphere Data Store (CADS) Technical Infrastructure supports the implementation of the Climate Change (C3S) and Atmosphere Monitoring (CAMS) Services of Copernicus. Both services are implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF) on behalf of the European Commission (EC). CADS infrastructure exposes two user facing interfaces when sharing backend components: the Climate Data Store (CDS) and the Atmosphere Data Store (ADS). Each of those offers a dedicated catalogue with a wide variety of datasets from Climate and Atmospheric domains which are accessible via seamless Web-based and API-based search and retrieve facilities.

ECMWF together with EUMETSAT, Mercator Ocean International and the European Environment Agency (EEA) also implements a 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.

CDS and ADS 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 format), there is a limitation for the visualization of these products in the WEkEO viewer due to the lack of a suitable and responsive WMS/WMTS service in the CADS infrastructure. This project aims to address this infrastructure gap with the provision of extended CADS infrastructure components for Serverless Data Access & Visualisation which will allow the operationalization of such services and the visibility of data in the WEkEO viewer. Delivered components and interfaces are expected to be broadly use by CADS as part of the overarching adoption of FAIR principles (Findable, Accessible, Interoperable, Reusable) across the platform.

Useful links in the context of this document:

Reference	Link to site
ECMWF	https://www.ecmwf.int/en/about/what-we-do

Copernicus Climate Change Service (C3S)	https://climate.copernicus.eu
Copernicus Atmosphere Monitoring Service (CAMS)	https://atmosphere.copernicus.eu
Climate Data Store (CDS)	https://cds.climate.copernicus.eu
Atmosphere Data Store (ADS)	https://ads.atmosphere.copernicus.eu
WEkEO	https://www.wekeo.eu

> **Scope of activities subject of this ITT:** The service subject of this ITT aims to extend the current FAIR capabilities of the CADS Infrastructure implementing Serverless Data Access & Visualisation components that will foster the integration of CADS Data & Services with other external platforms such as WEkEO.

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.
CADS	An acronym for Copernicus Climate and Atmosphere Data Store (which incorporates both the ADS and CDS) used in this document as to refer to the overarching infrastructure where this tender will be integrated.
CDS	Climate Data Store
CDM	Common Data Model
CIM	Content Integration Manager
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
Developer	The type of CDS users adding tools to the CDS Toolbox framework
EQC	Evaluation & Quality Control
ESGF	Earth System Grid Federation
Expert	The type of CDS Toolbox users writing CDS Toolbox workflows to build applications.
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
SOA	Service-Oriented Architecture
Tool	A software that performs computations on some input data or products and produces some output data or products.
Toolbox	The set of available tools, workflows and applications within the CDS.
URL	Uniform Resource Locator.
Users	The internal and external Users of the C3S infrastructure.
UX	User eXperience
VLE	An acronym for Virtual Learning Environment.
Web portal	The web interface of the CDS that enables the users to view information, access/perform tasks on the Product Catalogue, Toolbox and CMS.
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   CADS Infrastructure

The current CADS infrastructure is designed as a distributed system and open framework, providing improved access to a wide range of datasets via a powerful service-oriented architecture. It offers seamless web-based and API-based search and retrieve facilities to access C3S and CAMS data, services and information.

The CDS/ADS catalogues contain a wide variety of datasets in the Climate and Atmosphere domains. These datasets are distributed and located at different data suppliers and are accessible via adaptors using a wide range of protocols.

CADS core infrastructure includes the following main architectural components:

- **Web Portal:** User entry point to the data and other catalogued resources. The portal provides discovery functionalities presenting detailed information about the available products and interactive forms for users to subset and retrieve data. These forms are based on a set of widgets, and dynamic constraints which allow users to construct and refine their selections. The Web portal component interacts with other components in the infrastructure by mean of interoperable APIs which are partially exposed to the external world.
- **Backend:** It is a container for different services that seat at the core of the infrastructure. This component manages the catalogued content and metadata, the user profiles and the processing services and adaptors for the dispatching and handling of incoming requests. It also exposes interoperable APIs which are consumed within the platform and exposed externally.
- **Compute:** Component that executes the tasks submitted by the backend, including external retrievals caches the results, and implements the Quality of Service (QoS) rules used to regulate user activity. This component covers the following services:
  - o **Broker:** Receives requests from the Backend, implements the QoS rules and handle the tasks to be submitted to the Scheduler.
  - o **Scheduler:** Dispatches tasks to the compute Workers based on the available resources of the system.
  - o **Workers:** Execute the incoming requests. They write final outputs in the cache or return errors.

- o **Adaptors:** These components guarantee interoperability between the broker and the various kinds of data repositories translating data requests into a format that is understood by the data sources.
  - o **Cache:** This holds the results of computations ready to be downloaded by the user; all data and computation results are held by the cache area.
- **Metrics:** Operational monitoring component that offers near real-time view of the system and long-term statistics and KPIs. It raises alerts based on predefined conditions and thresholds.
- **API:** The system provides a Programmatic Interface which allow users to submit and retrieve data. API endpoints are compliant with Standard OGC APIs and STAC.

By means of these components CADS provides access to Data Repositories which are distributed and vary greatly in size, nature and formats. It should be noted that some of the repositories contain many petabytes of data and several billion products. Data repositories are expected to grow substantially in the coming future.

Currently there is an ongoing project for the modernization of the CADS Infrastructure. This project touches most of the current infrastructure components. New interfaces and endpoints will be released for users as result of this modernization project. At all effects CADS in this document refers to the modernized infrastructure.

The CADS infrastructure is hosted in on-premises Cloud Services (CADS Cloud) physically located at ECMWF Bologna Data Centre. Deployment of CADS instances perform on top of Kubernetes as the container orchestration engine. The underlaying Cloud Infrastructure high-level software specifications are:

- OpenStack (Yoga release) with all core functionality to which is added: Magnum, Designate and Manilla;
- Ceph (Quincy release) with all replication 3 storage.

Current CADS Cloud infrastructure will be decommissioned at the end of Q3 2024. Before that a Modernized Cloud Infrastructure will be provisioned where all current data and services will be migrated.

CADS Software and Hardware infrastructure modernization roadmap might partially overlap with the implementation of this contract. Tenders must be flexible to accommodate the potential impacts derived from actions in the CADS roadmap and to facilitate migration and transition between platforms.

ECMWF will provide the successful Tenderer with all the Cloud compute and storage resources required for the development and operational implementation of this contract.

1.3.2   WEkEO

As part of the European Union's Copernicus Programme, ECMWF, together with EUMETSAT, Mercator Ocean International and EEA, 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. The services have been integrated with the delivery platform by our partner EUMETSAT and their suppliers.

The WEkEO delivery platform uses OpenStack and Ceph. End-users are able to access a substantial portfolio of original Copernicus Programme and Sentinel satellite data using a harmonised data access mechanism and common data catalogue. They are also able to 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.

> **At the time of this ITT, WEkEO platform is in the process to be upgraded. Release roadmap might cause potential impacts on the activities envisaged in the context of this contract. The successful Tenderer should be flexible to accommodate the potential impacts derived from unavoidable actions of the overall WEkEO roadmap and to facilitate transition between platforms.**

1.3.3 Interaction between platforms

A broad selection of C3S and CAMS datasets are made available through the WEkEO DIAS platform to allow bespoke processing of that data in unison with all other DIAS data using a common data access mechanism.

CADS Standard Services and APIs are consumed by the WEkEO platform for the following purposes:

- Harvesting of CDS (C3S) and ADS (CAMS) catalogued metadata records to be discoverable in the WEkEO catalogue portfolio. This access is automatically done via Standard Metadata Services (CSW and STAC). Currently 119 C3S datasets and 13 CAMS datasets are available for discovery and are downloadable in WEkEO but these numbers are subject to change as both catalogues are in continuous evolution.
- Rendering of a replica of the CADS interactive forms for data sub-setting and request submission. Interactive forms also replicate the dynamic constraints to guide users on their selection process in a similar way that the CADS Web Portal interface does but more oriented to the final visualization as additional layers in the WEkEO viewer.
- Submission of Data requests via the WEkEO HDA API which is backend supported by CADS API for accessing CDS and ADS data. Incoming traffic from WEkEO platform is managed via ad-hoc quality service rules at broker level and supported by dedicated components and network connectivity.
- Monitoring and metrics that trigger alerts between platforms in the event of incidents affecting normal operation.

One of the most prominent user interfaces provided by the WEkEO platform is the Data Viewer (https://www.wekeo.eu/data). Due to the size, nature and format of C3S and CAMS datasets, rendering and visualization of this data requires dedicated architectural components which allow to provide responsive and highly efficient WMS/WMTS services.

Operational outputs from a previous proof-of-concept project, prototyping a scope of activities similar to the ones subject of this ITT, are available in WEkEO to illustrate the final aim:

- o ERA5 hourly data on single levels from 1940 to present.
- o CAMS Global atmospheric composition forecast.

> **CADS and WEkEO integration heavily relies on the implementation and adoption of Standard Data and Metadata Services. The Tenderer is strongly encouraged to describe how different Standards will be implemented at all levels on the scope of the proposed solution.**

2   Contract Definition and Requirements

2.1   Vision and Objectives

Vision: Set up the necessary infrastructure components that will provide operational Serverless ARCO Services and Visualization of C3S and CAMS datasets in the WEkEO viewer via standard WMS/WMTS data services.

The envisaged architecture is in the form of a dedicated ARCO Data Repository which integrates with the rest of the CADS infrastructure making use of its capabilities to access the raw data sources and perform operations close to the data that should optimize infrastructure resources and data transfers between main repositories and the Data Repository. After extraction from the source repositories, data is aimed to be transformed to the ARCO formats and stored in the destination Data Repository. An additional layer on top of the Data Repository exposes Standard Metadata and Data services that will be consumed by WEkEO and other external platforms and user communities.

The objectives of this contract can be then summarized with the provision of the following functionalities:

- Data handling layer supporting the implementation of configurable workflows that will translate data from the original data repositories to the ARCO Data Repository.
- An intermediate Data Repository hosting Cloud-Optimized (ARCO) versions of a selection of C3S and CAMS Datasets. Zarr format is envisaged for Data and Spatio-Temporal Asset Catalogue (STAC) for Metadata.
- WMTS Service supporting the visualisation of the datasets in the WEkEO Data Discovery Platform (viewer).
- Monitoring & reporting functionalities across the different components.

Provided infrastructure and implemented protocols should be extensible to allow the integration of existing and incoming datasets in the C3S and CAMS portfolio.

2.2   Statement of Requirements

2.2.1   General Requirements

The following sections detail the project requirements, summarise the envisaged architectural components and the functional scope of the implemented system.

All enumerated requirements describe a baseline of how ECMWF envisages the potential solution to be implemented. They are by no means a constraint for the solution to be proposed by the Tenderer but shall serve as a guide to define the scope of the proposal.

Being at the core of the implementation of the Copernicus Services, CADS has strong interdependencies with other EU platforms and funded projects. During the execution of this contract, interaction with third parties will be required with special mention to WEkEO stakeholders. ECMWF will facilitate interaction at all levels whenever this may be required.

The Tenderer is expected to propose a flexible workplan and methodologies which will facilitate interaction with third parties and coordination with the roadmaps of CADS and WEKEO infrastructures.

All delivered components will be hosted and run in the In-house Cloud Infrastructure physically located in ECMWF premises in Bologna (Italy). Delivered components must be compatible and optimized to deploy and run in the underlaying CADS Cloud infrastructure. ECMWF will facilitate the computing and storage resources

required for the development and operational implementation of the system and will support technical requirements at Cloud level and required data bridges.

2.2.2 Architectural Requirements

The architectural requirements of the proposed system are expected to be aligned with those of CADS to guarantee full compatibility with the current Software and Hardware infrastructures.

The architecture of the system will be Service Oriented, Open and Extendable in terms of the number of concurrent accesses, handled data volumes and processing capacity. Architecture will then include components which are scalable and are designed to take full advantage of the underlying Cloud Infrastructure.

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

- **Data Handling:** Architectural components in charge of extracting data from raw data sources, perform transformations and populate the destination Data Repository in Cloud-Optimized (ARCO) formats. This component must guarantee the integrity of data, allow automatization, and provide full visibility to operators of the system.
- **Data Repository:** This component can be also described as a Cloud-native Data Lake where data is stored in cloud-optimized (ARCO) formats and Metadata described by a Spatio-Temporal Asset Catalogue (STAC).
- **Standard Services:** Data Repository holdings are exposed via Standard Services which are ready to be consumed programmatically by external platforms and users (WMS/WMTS).
- **Operational Monitoring and Reporting:** Real time information for operators presenting the overall status and activity of the system and the executed workflows. These will be used for statistics, 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:

Reference	Requirement	Remarks
1.	Compatibility with CADS Software and Hardware Infrastructure	
2.	Cloud technology oriented	Object Store, Kubernetes
3.	Scalable components	
4.	Serverless (or hybrid) architecture	
5.	Cloud-optimized (ARCO) formats	Zarr, STAC
6.	Cloud Platform agnostic	
7.	Must run on a Linux/UNIX operating system	
8.	Administrator documentation is provided	
9.	Deployment is fully automated	
10.	Compliance with OGC and de-facto Standards	Zarr, WMS, WMTS, STAC
11.	Activity is registered in logs which can be parsed by standard monitoring tools	Splunk, Opsview

The Tenderer should include in their response how their proposed architecture ensures scalability.

2.2.2.1 Interoperability

CADS has a strong compromise with the implementation of FAIR principles at all levels. As one of the FAIR pillars, Interoperability is key for the evolution of the infrastructure as it facilitates the integration and communication at different levels, from data sources to external platforms and services.

The interoperability shall be achieved:

- Internally, between the dedicated infrastructure components and CADS core components; and
- Externally, with external information systems and platforms such as the INSPIRE Geoportal, the WMO Information System (WIS), the Global Earth Observation System of Systems (GEOSS) or the WEkEO platform.

To achieve interoperability, CADS and by extension the proposed system implements agreed international or de-facto standards (see Appendix 5.1).

Reference	Requirement	Remarks
12.	The successful Tenderer must ensure the adoption of FAIR principles aligned with CADS objectives	

2.2.2.2 Open-Source Software

The Tenderer shall propose software solutions which are the most practical and cost effective for ECMWF's intended use and future evolution and maintenance.

ECMWF expects the Tenderer to propose open-source software, whenever this is available and can provide or contribute to a practical solution.

The Tenderer must provide a list of:

- Open-Source software proposed for use.
- Components to be developed.
- 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.

Reference	Requirement	Remarks
13.	Open-Source Software	Whenever available
14.	Python based development for ad-hoc components	Preferred language

2.2.2.3 Supported Datasets

C3S and CAMS Catalogued Datasets vary greatly in volume and formats. Datasets are stored in distributed repositories which are accessed by dedicated protocols and adaptors.

CADS Broker component is in charge for dispatching incoming requests to the distributed data repositories and manage traffic and workload of the system via the implemented quality of service rules (QoS). As part of the data access and retrieve process, CADS system provides with additional functionalities to implement pre-

and post-processing close to the data. This allows among other things to convert final formats and optimize data transfers. All this complexity is seamless and transparent for CDS and ADS users as access is harmonized by the provided interfaces.

The delivered system under this contract is expected to take full advantage of the capabilities offered by the CADS infrastructure when accessing data sources and be on itself flexible to integrate the wider diversity of datasets.

Each Dataset might present particularities on typology, scope, volume and format. System should be flexible to handle different datasets and support the implementation and configuration of ad-hoc data integration workflows.

Some of the facts to be considered when evaluating the feasibility, scope and required effort for the integration of each dataset in the delivered infrastructure 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.

Reference	Requirement	Remarks
15.	System supports different incoming data formats	GRIB, NetCDF

Proposed architecture 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 CDS and ADS catalogues to evaluate the envisaged casuistic and potential burdens that the proposed solution might present for integration of current datasets portfolio.

2.2.3   System Requirements

As part of the proposal, the Tenderer is expected to present the software architecture describing the different components, their envisaged functionalities and how they will interact among them and with CADS.

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

Project milestones will demonstrate functional integration with WEkEO Viewer.

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

2.2.3.1   Data Handling

The purpose of the Data Handling component is to manage and orchestrate the data workflows that will rule transfers of data from the raw data sources to the destination Data Repository. Workflows might consist of different concatenated jobs such as those for data extraction, transformation and ingestion/creation of analysis-ready, cloud-optimised data formats in the destination ARCO Data Repository.

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 of their execution.

Number and complexity of workflows are expected to growth with the evolution of the system and the growth of the catalogues portfolio. The provided architecture must propose optimal ways to implement and 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 CADS architecture components and interfaces 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 CDS/ADS operational traffic.

Processing is expected to use parallelization and optimize the consumption of resources across all different steps in the workflow, 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.

Reference	Requirement	Remarks
16.	System allows for the implementation of ad-hoc workflows per dataset	
17.	Workflows can be scheduled, adapted and configured.	
18.	System guarantees the integrity and consistency of data.	
19.	Workflows can be monitored	
20.	Documentation is provided for the administration and operation of workflows	

2.2.3.2   Data Repository

The core of the service is envisaged as a Data Lake, an object store hosting the ARCO assets versions of CAMS/C3S data and metadata:

- **ARCO data**: Data Cubes generated from CDS/ADS data, in OGC Zarr format; and
- **ARCO metadata**, in Spatio-Temporal Asset Catalogue (STAC) format, always in sync with the data itself.

The idea of a Data Lake aims to provide high-performance, direct data access to co-located data consumers through CADS and WEkEO, as well as to potentially remote data consumers from the Internet at large. The Data Lake is expected to be the foundation for the implementation of Standards Services (WMS/WMTS) that will expose the hosted Data and Metadata resources.

ARCO Data is expected to be optimised for multiple consumption profiles and produced in different dimensional chunks at least for Time and Geospatial:

- **Time chunked dimension**: for users interested in analysing large geographic domain over a short time span. Structured in small chunk size in the *time* dimension but large in geographical.

- **Geospatial chunked dimension**: for users interested in analysing large time series over a relatively small geographical area. Structured in small chunk size in the *geographical* dimension but large in time.
- **Other dimensions**: Depending on the casuistic of each dataset, additional chunking dimensions might be considered.

The data redundancy of the Data Lake is expected to require large storage consumptions which are assumed by ECMWF. Nevertheless, to optimize resources, ECMWF will agree with the successful Tenderer the suitability and scope of each Dataset prior to start the ingestion work based on the different criteria described above. Data redundancy is however expected to bring some benefits:

- **Resilience against data loss**: Data hosting structure split in buckets and the inherent redundancy of the object store, guarantee that no data is lost in case of catastrophic events.
- **High performance for all access types**: combination of chunking schemes and high-speed connection to federated and internal networks, including HPC infrastructures.

Data Access Protocols:

The Tenderer is expected to describe the data access protocols and policies that will be implemented by the Data Lake, the different typologies of users, assigned privileges and the access limitations to data and services. Described protocols and policies will address both:

- Data ingestion (e.g. S3 access in read-write); and
- Data consumption (e.g. http accesses).

File Formats:

The format supported by the Data Lake is envisaged to be ARCO Data Storage format (Zarr). It will support compression and chunking schemas for different dimensions (time, geographical, other). Encoding schemas are expected to consider usage and storage.

Metadata is expected to be provided as Spatio-Temporal Asset Catalogue (STAC) to increase discoverability and interoperability with downstream cataloguing services.

Data Lake is aimed to support the implementation of Standard interfaces for users and downstream services. Direct access for consumers to data and metadata is intended to bring:

- High Performance availability and latency.
- Responsive Visualization of large datasets.
- Data-proximate processing.
- Web-based visualization and exploration at full resolution.

From the Data Provider perspective, the implementation of an intermediate Data Lake is aimed to bring:

- Serverless/hybrid architecture and use of object store to optimize required resources.
- Scalability to accommodate high-resolution products and traffic spikes.
- Leverage the use of WEkEO Jupyter Hub enabling data proximate processing.
- Granular usage monitoring at data chunk level.

Reference	Requirement	Remarks
21.	Data Lake aims to provide high-performance	
22.	ARCO Data is expected to be optimised for multiple consumption profiles	
23.	Data Access protocols for ingestion and consumption are described	S3, HTTP
24.	Final Data formats optimize final usage and storage	

2.2.3.3 Standard Services

The envisaged Data Lake described above will host the data in a format and structure which will facilitate the implementation of performing Data and Metadata Standard services.

This Services layer is expected to bring together a set of standards (Zarr, STAC, and WMTS) to improve CADS interoperability with different platforms and facilitate the visualisation of C3S and CAMS datasets on CADS and the WEkEO viewer.

A core requirement of this contract is the implementation of a WMTS Service on top of the Data Lake to render on-demand Map Tiles. This service will allow visualisation of the supported CAMS/C3S datasets in the WEkEO Data Discovery Platform (Viewer).

The key features provided by the WMTS services is aimed to address:

- Resolution: capacity to render in different resolution tiles.
- Projections: offer tile sets in different projections.
- Formats: Render tiles in PNG format with transparency.

WEkEO's Data Discovery Platform will interface with the provided WMTS Service in two different ways:

- Backend: Periodically WEkEO backend will query and harvest information about viewable layers from the WMTS service using the standard capabilities of the service.
- Frontend: User actions in the map browser (add layer, pan, zoom) will trigger the request of map-tiles directly from the WMTS service.

Operational outputs from a previous proof-of-concept project, prototyping WMTS integration with WEkEO viewer, are currently available for the following datasets:

- ERA5 hourly data on single levels from 1940 to present.
- CAMS Global atmospheric composition forecast.

Reference	Requirement	Remarks
25.	Datasets are made available via responsive WMS/WMTS that cover the full extend (variables and dimensions)	
26.	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 Monitoring and Metrics

Operationalization of the delivered infrastructure will require dedicated monitoring and metrics components.

The proposed system should have the ability to:

- Log system activities, for diagnostics and troubleshooting, to be used by analysts and as support of capacity planning (ECMWF currently uses Splunk for log indexing and analysis).
- Supervise the correct execution of data integration workflows assuring the integrity and consistency of the ingested data in the Data Lake and the original data sources.

- Monitor and raise alerts. System will need to integrate a set of checks into a monitoring tool, showing the health of the system (ECMWF currently uses Opsview).
- Measure system activity in terms of number of users, workflow executions, data accesses and to be able to convert these inputs into ECMWF Key Performance Indicators (KPI).
- Generate statistics about the use and access to the system to address European Commission and internal reporting obligations.

Collected information will be used to guarantee the quality and availability of the service, fulfil reporting obligations and to enable continuous improvements of the infrastructure (capacity planning).

Monitoring outputs will be suitable for further usage as inputs for user-oriented functionalities and dynamic configuration.

This component will cover the following baseline functionality:

Reference	Requirement	Remarks
27.	Ability to create health checks of the system components to the Operators.	To be integrated with Opsview.
28.	Ability to track execution of data integration workflows	
29.	Ability to generate system and application logs.	To be integrated with Splunk.
30.	Ability to conduct audits.	
31.	Ability to manage faults and record errors.	Self-explanatory and human readable messages.
32.	Ability to minimize downtime during upgrades.	
33.	Ability to provide system metrics based on KPIs.	

3 Contract Structure and Implementation

3.1 Contract Approach

As stated in previous chapters, the final objective of this contract is to set up the required infrastructure components that will allow CADS to operationalise the integration and visualisation of C3S and CAMS data products on the WEkEO Viewer via standard WMS/WMTS data services. The proposed approach for this contract is driven by the following objectives:

- ECMWF's vision and architectural principles are aligned with those of the Tender.
- Knowhow and experience from previous tests with ARCO Services on top of C3S and CAMS data sets are taken onboard in the form of functional requirements and technical specifications.
- Requirements and expectations of main stakeholders are properly addressed.
- Functional integration and interfaces with WEkEO platform are properly established.
- Project roadmap is aligned with those of CADS and WEkEO.

To achieve the above-mentioned objectives an approach to the contract by different phases or areas of work is proposed:

- **Implementation of the Software Infrastructure:** After agreement and approval of the final architecture, the successful Tenderer is expected to implement the technical infrastructure that will support the operationalization of the system. Infrastructure must be understood in this context as all the software components that will intervene in the data handling, data hosting and standard service provision. This area might imply the following activities:

- Configuration and tuning of the underlaying cloud infrastructure to support the deployment of the dedicated system components and allow for its correct functioning.
- Development and implementation of software components provided by the contract.
- Grant of access rights, privileges, and permissions for those required components not directly managed by the successful Tenderer.
- Validation and testing of network connectivity, internal and external.
- Configuration and tuning of dedicated CADS components that will be consumed by the system.
- 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:** The purpose of this phase is the implementation of ad-hoc data integration workflows that will support the scheduled transfers of data from the raw data sources to the Data Lake where those will be hosted in the ARCO formats. Workflows might require the development or configuration of ad-hoc jobs per dataset in order to assure final compatibility.

  Workflows will rely for their implementation and operational functioning on the combined capabilities of CADS components and those delivered by this contract. ECMWF will work together with the successful Tenderer on defining the most optimal ways for accessing, retrieving, and processing raw data minimizing the concurrence and competition with CDS/ADS operational traffic. This area of work might imply the following activities:

    - Identify the datasets to be integrated.
    - Define the scope (subset) of data 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 WMTS service on top of the ARCO Data.

- **Maintenance & Support:** This area of work will cover the outputs of previous phases both Software Infrastructure and Data Integration. 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 the Agile methodology reported in Annex 4 of the Volume V Framework Agreement template part of this ITT but welcomes suggestions from the Tenderer on what methodology they propose to apply for the different phases of the project based on their knowledge and experience with projects of a similar nature. The Tenderer must provide examples of how they have applied this approach in similar projects they have previously worked on.

The methodologies proposed by the Tenderer must ensure that final deliverables and milestones are fit for purpose, aligned with the project vision and remain within project budget and schedule.

3.2 Implementation Schedule

ECMWF intends to award a single Framework Agreement for a period of 18 months, with the right to extend for a further 12 months at ECMWF's discretion, which shall be implemented via a single Service Contract expected to commence in Q2 2024. The Tenderer shall provide a detailed implementation plan of proposed activities for the initial 18-months contract period.

3.3 Work Packages

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

General note about the Deliverables and Milestones IDs: 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.

3.3.1 Work Package 1: Software Infrastructure

This Work Package 1 (WP1) aims at implementing the technical infrastructure that will support the operationalization of the system. Infrastructure must be understood in this context as all the software components that will intervene in the data handling, data hosting and standard service provision.

The final objectives of this work package are the following:

- Implement an iterative/continuous improvement development approach.
- Produce a prototype to demonstrate structure and functionality (proof of concept).
- Produce the components of the pre-operational phase.
- Conduct component, unit, system, and integration testing.
- Deliver the system and perform acceptance tests at ECMWF.
- Deliver all required documentation.

This work package may require some coordinated actions with ECMWF and other contractors such as:

- Configuration and tuning of the underlaying Cloud Infrastructure to support the deployment of the dedicated system components and allow for its correct functioning.
- Development and implementation of software components provided by the contract aligned with the rest of the CADS Infrastructure.
- Grant of access rights, privileges, and permissions for those required components not directly managed by the successful Tenderer.
- Configuration and tuning of dedicated CADS components that will be consumed by the system.
- Validation and testing.

ECMWF will facilitate communication and coordination between different actors at all levels. ECMWF will also support the Tenderer implementing those technical requirements required by the system which may lay out of the scope of the contract and depends on ECMWF or other external contractors.

WP1 Deliverables and Milestones				
Deliverable / Milestone ID	*Responsible*	*Nature*	*Deliverable / Milestone title*	*Due date*
D1.y.z	Tenderer	Report		
D1.y.z	Tenderer	Report		
M1.y.z	Tenderer	Milestone		

3.3.2   Work Package 2: Dataset integration

This Work Package 2 (WP2) aims at the progressive ingestion of C3S and CAMS datasets in the Data Lake and its final operational integration for visualisation in the WEkEO Viewer. This work package is proposed to follow an incremental approach based on Agile Methodology where datasets will be sequentially submitted, evaluated, approved, and finally integrated.

The integration of a dataset implies all the necessary workflows, jobs and system configurations that will drive the operational implementation of the complete data life cycle to handle data flows from the original Data Sources to the Data Lake and towards the final consumption via the Standard WMTS Service by the WEkEO Viewer.

This work package is envisaged to be based on an iterative process between ECMWF and the successful Tenderer, consisting of the following steps:

- Define the backlog of datasets to be integrated.
- Evaluate the feasibility for registered datasets to be converted to ARCO formats.
- Estimate and plan the work for the implementation of ad-hoc jobs and workflows.
- Once estimated and approved, develop and implement the workflows.
- Schedule the workflows and populate the Data Lake.
- Test and validate the outputs.
- Configure the workflows for automatic updates.
- Update documentation and provide guidelines for operating the workflows.
- Adapt the monitoring tools to track the workflows executions.
- Integrate with the Operational WEkEO platform.

WP2 Deliverables and Milestones				
Deliverable / Milestone ID	*Responsible*	*Nature*	*Deliverable / Milestone title*	*Due date*
D2.y.z	Tenderer	Report		
D2.y.z	Tenderer	Report		
M2.y.z	Tenderer	Milestone		

3.3.3　Work Package 3: Operational maintenance

As part of the work subject of this ITT, the successful Tenderer is expected to take over the maintenance of the delivered software infrastructure and implemented workflows through this Work Package 3 (WP3).

The scope of this work package covers the following objectives:

- Perform corrective maintenance and support to guarantee operational continuity.
- Generate incident reports and root cause diagnosis.
- Perform preventive maintenance actions to avoid failures or performance degradation.
- Provide support for the operational management of software and workflows.
- Advice on the reusability of existing components.

WP3 Deliverables and Milestones				
Deliverable / Milestone ID	*Responsible*	*Nature*	*Deliverable / Milestone title*	*Due date*
D3.y.z	Tenderer	Report		
D3.y.z	Tenderer	Report		
M3.y.z	Tenderer	Milestone		

3.3.4　Work Package 0: Management and coordination

The following management aspects shall be briefly described in the Tender (this is not an exhaustive list):

- Contractual obligations as described in the Volume V Framework Agreement Clause 2.3 on reporting and planning.
- Meetings (classified as tasks and listed in a separate table as part of the Proposal):
  - ○ ECMWF and the successful Tenderer will organise a Kick-Off Meeting during the first month of implementation of the contract.
  - ○ ECMWF will host monthly teleconference meetings to discuss this Copernicus Joint Services (CJS) service provision and other topics. The Project Manager appointed by the successful Tenderer will represent the successful Tenderer in such meetings.
  - ○ ECMWF and the successful Tenderer will organise quarterly Progress Review Meetings, linked to Payment Milestones, unless otherwise agreed.
  - ○ ECMWF will organise annual C3S and CAMS General Assemblies. The successful Tenderer is required to attend these meetings with team members covering the different topics that are part of this ITT.
  - ○ The Tenderer can propose additional project meetings as part of their response.
- Quality assurance and control: the final quality check of the deliverables should be made by the prime contractor (contents, use of ECMWF reporting templates for deliverables and reports (Microsoft Word and Excel), format, deliverables/milestones numbering and naming, typing errors, etc.). All reports in this project shall be in English. Unless otherwise specified, the specific contract Deliverables shall be made available to ECMWF in electronic format.
- Communication management (ECMWF, stakeholders, internal communication).
- Resources planning and tracking using the appropriate tools.
- Implementation of checks, controls and risk management tools for both the prime contractor and sub-contractors.
- Sub-contractor management, including conflict resolution, e.g., the prime contractor is responsible for settling disagreements, although advice/approval from ECMWF may be sought on the subject.

- A list of sub-contractors, if any, describing their contribution and key personnel shall be provided, as well as back-up names for all key positions in the contract. The Tenderer shall describe how the Volume V Framework Agreement, in particular Clause 2.9, has been flowed down to all their sub-contractors.
- Management of personal data and how this meets the requirements of Clause 2.8 and Annex 6 of the Volume V Framework Agreement.

The Tenderer shall complete the relevant table in Volume IIIA as part of its Tender, which shall include the deliverables and milestones for this work package already indicated in the tables below. Volume IIIA will be used by the Tenderer to describe the complete list of Deliverables, Milestones and corresponding schedules for each Work Package. All Deliverables and Milestones shall be numbered as indicated. All document deliverables shall be periodically updated and versioned as described in the tables.

WP0 Deliverables and Milestones				
Deliverable / Milestone ID	*Responsible*	*Nature*	*Deliverable / Milestone title*	*Due date*
D0.y.z-YYYYQQ	Tenderer	Report	Quarterly Implementation Report QQ YYYY *QQ YYYY being the previous quarter*	Quarterly on 15/04, 15/07 and 15/10
D0.y.z-YYYY	Tenderer	Report / Other	Annual Implementation Report YYYY – Part 1 *YYYY being the Year n-1* including both: • the Quarterly Implementation Report Q4 YYYY and • the requested preliminary financial information YYYY	Annually on 15/01
D0.y.z-YYYY	Tenderer	Report	Annual Implementation Report YYYY – Part 2 *YYYY being the Year n-1*	Annually on 28/02
D0.y.z-YYYY	Tenderer	Report	Annual Implementation Plan for year YYYY *YYYY being the Year n+1*	Annually on 31/09
D0.y.z-YYYY	Tenderer	Other	Copy of Prime Contractor's general financial statements and audit report YYYY *YYYY being the Year n-1*	Annually
D0.y.z	Tenderer	Report	Final Report	60 days after end of contract
M0.y.z-KOM	Tenderer	Presentation and MoM	Kick-Off Meeting and associated MoM	30 days after start of contract
M0.y.z-PRMx	Tenderer	Presentation and MoM	Progress Review Meetings / Payment Milestone SC1-PMx and associated MoM	~ Every 6 months
M0.y.z-C3SGAxxxx	Tenderer	Attendance	C3S General Assembly YYYY	Annually
M0.y.z-CAMSGAxxxx	Tenderer	Attendance	CAMS General Assembly YYYY	Annually

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 document.

The Tenderer shall demonstrate for itself, and for any proposed sub-contractors that they have participated in national or international research and/or private sector software integration projects in the last 5 years for the activities for which the Tender is proposed. 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 Project Manager, responsible for the delivery of the system, to oversee the progress of this project, as well as a Service Manager, that will represent their interests in day-to-day discussions and meetings.

The Tenderer must outline the project team. The outline must contain the following:

- Relevant experience of key staff and management personnel.
- Names of Project Manager 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:

- Have 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.
- Demonstrated experience on running large scale systems preferably related with the domain (Meteorological, Climate, etc.)
- Have an adequate understanding of the current CADS Infrastructure and WEkEO DIAS Platform.

The 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 the 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 the 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

A deliverable is a substantial, tangible or intangible good or service produced as a result of a contract. In other words, a deliverable is an outcome produced in response to the specific objectives of the contract and is subject to acceptance by both ECMWF's Technical Officer (TO) and Contract Management Officer (CMO).

Milestones should be designed as markers of demonstrable progress in service development and/or quality of service delivery. They should 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 (except for Work Package 0), which are consistent with the set objectives. The deliverables can be in the form of documents or reports, data sets or databases, services, and user support. Requirements for each type are described in the following sub-sections.

3.5.1   Documents and Reports

All deliverables, i.e. reports, documentation and presentations for the contract shall be consistent with the technical requirements specified in this ITT and produced in English.

The quality of reports and deliverables shall be equivalent to the standard of peer-reviewed publications and practice.

When defining deliverables, please consolidate their numbers against a specific deadline where possible.

Unless otherwise specified in the specific contract, deliverables shall be made available to ECMWF in electronic format (PDF/Microsoft Word/Microsoft Excel or HTML) via the Copernicus Deliverables Repository portal. The details will be agreed at the negotiation stage.

A high-level project management plan must be delivered as part of this ITT.

The following documents shall be delivered as part of the contract (this is not an exhaustive list):

- An updated project management plan including milestones.
- Progress reports.
- Work package documentation.
- Detailed description of the test cases used for the internal validation of the software and workflows including test results:
  - o Functional tests.
  - o Performance tests.
  - o Availability tests.
  - o Consistency tests.
- Interface descriptions of all modules.
- Software quality assurance plan.
- Sign offs.
- Training plans.
- Documentation (system, software, source code).
- Release notes.
- User guide, data supplier guide, administration guide, installation guide.

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.

3.6 Tasks to be performed

The successful Tenderer is required to:

- Implement the set of activities and corresponding Work Packages (WP) described in this ITT.
- Deliver related WP deliverables and achieve milestones following the CADS processes and requirements when applicable.
- Deliver all required documentation (in English).
- Carry out training of ECMWF personnel if required (transfer of knowledge).

3.7 Key Performance Indicators

As part of the Tender, the Tenderer shall specify a proposed set of Key Performance Indicators (KPIs) suitable for monitoring various aspect of the service performance. Therefore, the KPIs shall be designed to quantify different aspects of quality of service against the requirements described in this document.

Contractors shall report to ECMWF on a set of SMART (Specific, Measurable, Actionable, Realistic and Time bound) KPIs suitable for monitoring various aspect of service performance, including (but not limited to):

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

The KPIs will be reported in the Quarterly and Annual Implementation Reports (cf. WP0 deliverables). At the end of each year, a Service Readiness Review shall take place that will include assessment of performance against the set of KPIs.

The table below provides the 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. All KPIs shall be labelled and numbered as indicated. All KPIs shall be periodically updated as described in the tables. Tenderers shall provide preliminary versions of the completed tables as part of their Tender.

KPI #	KPI Title	Performance Target and Unit of Measure	Frequency of Delivery	Explanations / Comments
KPI_1	xxx	…	xxx	xxx
KPI_x	xxx	…	xxx	xxx

These initial specifications shall be refined together with ECMWF during the negotiation of the contract.

3.8   Communication

The successful Tenderer shall 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.

All communication activity must be agreed with the ECMWF Copernicus Communication team in advance. This includes, but not exhaustively, communication planning, branding and visual style, media outreach, website and social media activity, externally facing written and graphic content and events.

Agreed activity would also need to be evaluated and reported on, once complete, so that success measures and KPIs can be provided to the European Commission.

4   Tender Format and Content

General guidelines for the Tender are described in 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 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.

Section	Page Limit
Executive Summary	2
Track Record	2 (for general) and 2 per entity
Quality of Resources Applied	2 (excl. Table 1 in 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 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 Volume IIIB) + 2 per each Work Package template (Table 3 in Volume IIIB)
Pricing table	No limitation

Table 1: Page limits

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 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 an ITT of this size, with experience in the appropriate delivery methodology proposed in section 3 above.
- A technical project team with more than 5 years of experience on performing activities related to the various aspects of this ITT.

The CVs, proven track record and certification of key individuals is required, including a brief 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 understanding of that solution and of the C3S/CAMS context. This section shall also include information on any other third-party suppliers that are used as part of the technical solution, and a statement of compliance for each requirement formulated throughout this document, describing how the proposed solution maps to the requirements.

4.2.5   Management and Implementation Plan

The Tenderer shall provide a detailed implementation plan of proposed activities for the duration of the Volume V Framework Agreement (cf. section 3.2 above). Deliverables and Milestones should be consistent with requirements specified in this document.

The Tenderer is requested to include management and implementation activities within the dedicated Work Package 0 (WP0, cf. section 3.3.4 above). The number of milestones is not restricted, but they should be designed as markers of demonstrable progress in service development and/or quality of service delivery. Adjustments to the proposed implementation plan can be made on an annual basis depending on needs for service evolution, changed user requirements, or other requirements as agreed between the European Commission and ECMWF.

As part of the general project management description, the Tenderer shall consider the elements listed in section 3.3.4 above.

4.2.6 Other

Please note that in Volume IIIA Pricing and deliverables (cf. Excel sheet "Deliverables List"):

- • Volume IIIA will be used by the Tenderer to describe the complete list of deliverables, milestones and schedules for each WP. All deliverables and milestones shall be numbered as indicated. All document deliverables shall be periodically updated and versioned as described in the tables.

- • Each deliverable shall have an associated resource allocation (person-months in column "Nb of PM allocated", and financial budget - resource type to be considered is payroll only – in column "Estimated price"). Therefore, the total of these allocated resources shall amount to the requested budget associated with payroll.

- • Milestones should not have any associated budget unless there are deliverables associated to them such as presentations, MoMs, etc.

- • The Tenderer shall ensure that the proposed due dates of deliverables and milestones are realistic and achievable.

- • A payment plan shall be proposed and provided as part of the Tender in Volume IIIA.

Any dependencies on input data (whose origin must be specified) shall be detailed and also accounted for in the risk register table.

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/
GEMINI	UK Discovery Metadata Standard	http://www.agi.org.uk/join-us/agi-groups/standards-committee/uk-gemini
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://www.wmo.int/pages/prog/www/WMOCodes/Guides/GRIB/Introduction_GRIB1-GRIB2.pdf
GRIB 2	General Regularly-distributed Information in Binary form Version 2	https://www.wmo.int/pages/prog/www/WMOCodes/Guides/GRIB/Introduction_GRIB1-GRIB2.pdf
GridFTP	high-performance, secure, reliable data transfer protocol optimized for high-bandwidth wide-area networks	http://toolkit.globus.org/toolkit/docs/latest-stable/gridftp/
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
ISO19119	Identifies and defines the architecture patterns for service interfaces used for geographic information	http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=39890
ISO19139	defines Geographic MetaData XML (gmd) encoding, an XML Schema implementation derived from ISO 19115	http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=32557
JSON	JavaScript Object Notation	http://json.org/
KML	Keyhole Markup Language	http://www.opengeospatial.org/standards/kml/
LAS	Live Access Server is a web server to provide flexible access to scientific data	http://ferret.pmel.noaa.gov/LAS
MARS	Meteorological Archival and Retrieval System	http://www.ecmwf.int/en/what-mars
ncBrowse	Java application to visualise netCDF files	http://www.epic.noaa.gov/java/ncBrowse/
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/
OGC - CSWWIS	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 - 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
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

5.2   Software

Apache Open Climate Workbench	software that performs climate model evaluation using model outputs from a variety of different sources	https://climate.apache.org/
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	
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
GrADS	Grid Analysis and Display System	http://iges.org/grads/
IRIS	Python package for analysing and visualising meteorological and oceanographic data sets	http://scitools.org.uk/iris/index.html
java-netcdf	java netcdf library	https://www.unidata.ucar.edu/netcdf-java
Leaflet	Open Source JavaScript library used to build web mapping applications	http://leafletjs.com/
Magics	ECMWF's Meteorological plotting software	https://software.ecmwf.int/wiki/display/MAGP/Magics
Matplotlib	a python 2D plotting library	http://matplotlib.org/
Metview	ECMWF's Meteorological workstation application	https://software.ecmwf.int/wiki/display/METV/Metview
MIR	computer display server for the Linux operating system	http://unity.ubuntu.com/mir/
ncBrowse	Java application to visualise netCDF files	http://www.epic.noaa.gov/java/ncBrowse/
NCO	netCDF Operators	http://nco.sourceforge.net/
NumPy	NumPy is the fundamental package for scientific computing with Python	http://www.numpy.org/
ECMWF_odb_api	API to the ODB	
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/
Scipy	Python-based ecosystem of open-source software for mathematics, science, and engineering.	http://www.scipy.org/