

Objectives

- Design and develop an EO Software Platform that both supports & controls access to the EO Data Warehouse;
- Design and Implement easy to use customisable and aggregatable services that support recurrent data /needs adhering to EO data standards;
- Provide a brokerage service by which to locate/rank appropriate data source(s)/providers;
- Provide support for satellite image cropping/trimming through bounding box definition for specific area;
- Develop services for the conflation/fusion of disparate data streams within the spatiotemporal domains;
- Provide services for manipulation of satellite images to clean/cluster features based on pixel clustering and pre-processing;
- Provide data analytics to examine satellite images to clean/cluster features based on pixel clustering utilising pre-processing and machine learning techniques;
- Design and Develop an EO data source discovery service

Description of work (where appropriate, broken down into tasks), lead partner and role of participants

This work package will identify *generic, recurring and much needed* EO data services. The identification of such will *in part* derive from stakeholder engagement with the sectoral areas in WP7. Prioritisation of those services to develop will be based upon perceived sectoral need. The first wave of such services will seek to demonstrate the power deriving from the effective leveraging of EO data and thereafter to encourage the extension and consolidation of their usage.

These services will be constructed as a suite of microservices adopting the emerging de facto industry standard of Microservices Architecture (MSA). In the first instance these will be largely stand alone services however in time process chains together will be developed whereby individual microservices can be aggregated to deliver end-to-end solutions to specific problems.

EO data embraces a myriad of potential data sources and streams including but not limited to: satellite, drone, lidar, ground segment and citizen derived. The CAMEO EO platform will make provision for such diverse and rich datasets through functionality for storage and access; cataloguing and searching; data processing and analysis; visualization; secure authentication and authorization (dealt with in the security WP4).

Task 2.1: Design & Delivery of CAMEO EO Software Platform

The CAMEO EO software platform will provide the scaffolding within which the overall CAMEO functionality will be delivered. It will involve a number of functionalities:

1. Design and delivery of a *data brokerage service* that will source data from a range of data service providers according to user specified queries including: price, quality, latency and spatio-temporal adjacency.
2. To develop the EO platform to support data warehouse access and provide cloud resources for the CAMEO platform to run on. Oracle will specifically support this activity by providing a dedicated resource throughout the project duration to provide cloud architectural support to developers. Oracle will also ensure the adoption of industry best practises and that what is built within the CAMEO platform is built secure and scalable.
3. To integrate both the data quality filter developed within **WP3** and data security provisions from **WP4**;
4. To support on-line access to training resources developed in **WP5** through an off the shelf Virtual Learning Environment.

Task 2.2: Core EO Microservices

Task 2.2 will develop a catalogue of EO microservices which will be developed which support:

- Imagery services: for cropping, window query, image enhancement (antialiasing; contrast enhancement), filtering (eg. sharpening/smoothing), pixel clustering, feature extraction, segmentation, affine transformations; preprocessing functionality including orthorectification, projection, data conversion (including image compression); spectral analysis; data export into a variety of different formats (SHP, XMLMetadata, GeoJSON);
- LiDAR services: downsampling, filtering, window query, Digital Terrain Model (DTM) generation, spatio-temporal data integration (e.g. airborne and terrestrial LiDAR datasets, etc.).

Task 2.3 Advanced EO Microservices

A number of advanced services will be developed to supplement the initial CAMEO core microservices.

These will include:

- *Intelligent EO data services* eg. image classification, longitudinal change detection, data analytics (application of ML techniques) to facilitate identification and visualisation of data attribute correlations, data attribute spatio-temporal clustering at pixel level using eg Moore I [Ref], Ref, Ref]. UCD involvement in the ESA AIREO project²⁰ (AI-Ready Earth Observation Training Datasets) concerned with the design specifications and best practices for AI ready (Machine Learning) training datasets for EO data including generating, structuring, describing and curating training datasets. This experience will enable CAMEO to overcome demonstrable bottlenecks to the application of ML to EO. Version 1.0 of the AIREO specifications and best-practices will be available summer 2021 and will frame dataset curation and the effective application of ML to EO data. Dell will contribute to the development of intelligent data query mechanisms using ontologies.
- *Advanced data conflation services*, facilitating spatio-temporal data joins, spatio-temporal integration of different data sources (eg LiDAR and imagery, ground segment, UAV, citizen derived). Dell EMC will leverage its experience of image processing, IoT data factorisation and sensor fusion, using its knowledge and expertise in ML algorithms, intelligent image processing and semantic lifting.

Task 2.4: Design of Supports for EO Process Chaining

In order to effectively leverage the functionality offered through the CAMEO microservices catalogue (core plus advanced) services a *modular process chaining facility* will be designed to both inform and enable service selection akin to *plug and play* enabling service composition where service interfaces are clean and adhere to interface standards (Restful APIs) facilitating service interoperability and data exchange. UCD will draw upon its experience to deliver orchestration of different services into a usable framework and pipeline. It is anticipated that a multi-agent systems approach will be used to introduce multiple agents collaborating within the process chain.

Task 2.5: User Experience (UX) Design

User adoption and sustained user usage is an imperative for the CAMEO user interface. FAIR data principles (Findable, Accessible, Interoperability and Reusability) and ease of use for non-EO experts will be central to CAMEO. Ease of access and interaction with the data warehouse and associated resources/services will be ensured through *user centred design* and iterative interface releases and usability testing. This task will provide a suite of visualisation options appropriate for particular data view requirements as established through stakeholder engagement in **WP7**. It is intended to accommodate the Oracle analytics engine in rapidly supporting data visualisation options. Oracle has extensive UX experience from interface delivery across diverse user sectors and this know-how will be harnessed within this task. **Treemetrics, TWM and Icon** will assist with usability testing while UCD will contribute to design and development of the icon based UI along with NLP processing for intuitive search.

CAMEO will adhere to guidelines of the OGC EO Exploitation Platform DWG²¹. The key activities of such an

²⁰ <https://eo4society.esa.int/projects/aireo/>

²¹ <https://www.ogc.org/projects/groups/eoexplatform>