**DTIF3 Contract Reference: DT 2020 0209**

Title/Acronym: **C**reating an **A**rchitecture for **M**anipulating **E**arth **O**bservation data (CAMEO)

WP3

(Data Quality Assurance)

**Project Workplan, Deliverables:**

# GANTT Chart: Timing of Work Packages and their components

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | M1 | M2 | M3 | M4 | M5 | M6 | M7 | M8 | M9 | M10 | M11 | M12 | M13 | M14 | M15 | M16 | M17 | M18 |
| WP0 | D0.1 | D0.2 |  |  |  |  |  |  |  |  |  | D0.3 |  |  |  |  |  |  |
| WP1 |  |  |  |  |  |  |  | D1.1 |  |  | D1.2 | D1.3.1 |  |  |  |  |  | D1.3.2 |
| WP2 |  |  |  |  |  |  |  | D2.1 |  | D2.2 | D2.3 |  |  |  |  |  | D2.5.1 | D2.2 |
| WP3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| WP4 |  |  |  |  |  |  |  |  | D4.1 |  |  | D4.2 |  |  |  |  |  |  |
| WP5 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | D5.1 |
| WP6 |  |  | D6.1, D6.2 |  |  |  |  |  |  |  |  | D6.3 |  |  |  |  |  | D6.4 |
| WP7 |  |  |  |  | D7.1 |  |  | D7.2 |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | M19 | M20 | M21 | M22 | M23 | M24 | M25 | M26 | M27 | M28 | M29 | M30 | M31 | M32 | M33 | M34 | M35 | M36 |
| WP0 |  |  |  |  |  | D0.4 |  |  |  |  |  |  |  |  |  |  |  | D0.5 |
| WP1 |  |  |  |  |  |  |  |  | D1.3.3 |  |  |  |  |  |  |  |  |  |
| WP2 | D2.4 | D2.3 |  |  |  |  |  | D2.2 |  |  |  |  |  | D2.5.2 |  |  | D2.3 |  |
| WP3 | D3.1 |  |  |  |  | D3.2 |  |  |  |  |  |  |  |  |  |  |  |  |
| WP4 |  | D4.3 |  |  |  |  |  | D4.4 |  |  |  |  |  |  |  |  |  |  |
| WP5 |  | D5.3, D5.4 |  |  | D5.2 |  | D5.3 |  |  |  |  | D5.3 |  | D5.3 |  |  |  |  |
| WP6 |  |  | D6.6.1 |  |  |  |  |  |  |  |  | D6.6.2 |  |  |  | D6.5 | D6.6.3 |  |
| WP7 |  |  |  | D7.3.1 |  | D7.4.1 |  |  |  | D7.3.2 |  | D7.4.2 | D7.3.3 |  | D7.4.3 |  |  | D7.5 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Work package number** | WP3 | | **Start Date** | |  |  | M6 |  |  |
| **Work package title** | Data Quality Assurance | | | |  |  | |  |  |
| **Participant number** | **1** | 2 | | 3 | 4 | 5 | | 6 | 7 |
| **Short name of participant** | **UCD** | VC | | ES | ICO  N | TM | | TW  M | Dell  Technologies |
| **Person/months** | **65 (-**  **5)\*** | 4 | | 4 | 3 | 16 | | 8 | 6 |

**Objectives**

●Design and formulation of mechanisms for the adjudication of data quality;

●Use of discovery services to identify temporally and geographically adjacent data sources;

●Provision of services for ground truthing data with relevant (location and temporal adjacency) and known high quality data sets;

●Design and implementation of trusted mechanisms to filter ‘poor quality’ data and ensure non-admittance to the data warehouse

**Description of work**

Poor quality data will invariably lead to poor decisions. It is imperative therefore to seek to ensure that the CAMEO data warehouse is only populated with quality data or at the very least data for which the indicative quality of data is known.

Adjudication of data quality and mechanisms for doing so need to be incorporated throughout the entirety of the big data model including data collection, data pre-processing, data processing and analytics, and data use. This work package will involve 4 subtasks.

**Deliverables (brief description and month of delivery)**

D3.1 Design of Data Quality Adjudication Framework (M19)

D3.2 Design and Implementation of Data Quality Filter (M24)

**Milestones**

MS3.1 Delivery of Data Quality Filter (M24)

**Deliverables for D3.1**

1. **Study of CAMEO framework**
2. **Identification of various earth observatory sources**
3. **Categorization of earth observatory data types**
4. **Identifying need of Data Quality and its evaluation matrix**
5. **Implementation of EO raster data Quality matrix**
6. **Predictive analysis of usability for a SME**

**Deliverables for D3.2**

1. **Study of Vector data quality metrics**
2. **Classification of various vector EO data and identification of Vector data quality metrics**
3. **Implementation of EO vector data Quality matrix**
4. **Data filters for SME raster and vector data**
5. **Predictive analysis of usability of vecdtor data for a SME**

**September**

Week 1:

* Introduction to CAMEO
* Basic employment formalities

Week 2:

* Introduction to CAMEO architecture
* Introduction to GIS and Its data

Week 3:

* Introduction to Work Package 3 &2

Week 4:

* Working of spartial data quality and its types

**October**

Week 1:

* Data-security course enrolled
* Research ethics course

Week 2

* Data-security course Completed
* Collection of GIS raster data from various sources
* ARCGIS workshop completed
* Report in WP3 deliverables and outcome

Week 3

* Report in WP3 deliverables and outcome

November

* EO conference attended
* ARGIS workshop attended
* Ethics workshop
* Open Access Publishing Options workshop library
* Postdoc Career Orientation
* Discover Datawrapper To Show Your Data On A Map
* Intellectual Property And Your Research

December

* ML and cloud classification Landsat & sentinel
* Data-security course completed

January

* New architecture for WP3
* Implementation of SSIM, MS-SSIM
* Defining precision, completeness

February:

* Defining new architecture of WP3
* Defining precision, completes, accuracy and consistency
* Review article submitted

March:

* PPT for the demonstration
* Implementing completeness
* Implementing precision

Targets:

* Implement: precision, completes, accuracy (raster)
* Implement: PSNR, SSIM, MS-SSIM, MSE for sample data
* And test it with sample data/AOI/POI
* Finding sample timeseries data from CAMEO
* Find appropriate SDQ for time series data