



TELESPAZIO

a LEONARDO and THALES company

<Template Service> Design Document

EOEPCA.SDD.xxx

TVUK System Team

Version 0.1, dd/mm/yyyy:

<Template Service> Design Document

| | |
|---|----|
| 1. Introduction | 2 |
| 1.1. Purpose and Scope | 2 |
| 1.2. Structure of the Document | 2 |
| 1.3. Reference Documents | 2 |
| 1.4. Terminology | 4 |
| 1.5. Glossary | 9 |
| 2. Overview | 11 |
| 3. Design | 12 |
| 3.1. Resource Catalogue Service | 12 |
| 3.1.1. Overview and Purpose | 12 |
| 3.1.2. Software Reuse and Dependencies | 12 |
| 3.1.3. Interfaces | 12 |
| 3.1.4. Data | 12 |
| 3.1.4.1. Configuration | 12 |
| 3.1.4.2. Data flow | 12 |
| 3.1.5. Applicable Resources | 13 |
| 4. Workflow | 14 |
| 4.1. Backend ingest and metadata processing | 14 |

| | |
|---|---|
| COMMENTS and ISSUES If you would like to raise comments or issues on this document, please do so by raising an Issue at the following URL https://github.com/EOEPCA/template-svce/issues . | PDF This document is available in PDF format here . |
| EUROPEAN SPACE AGENCY CONTRACT REPORT The work described in this report was done under ESA contract. Responsibility for the contents resides in the author or organisation that prepared it. | TELESPAZIO VEGA UK Ltd 350 Capability Green, Luton, Bedfordshire, LU1 3LU, United Kingdom. Tel: +44 (0)1582 399000 www.telespazio-vega.com |

AMENDMENT HISTORY

This document shall be amended by releasing a new edition of the document in its entirety. The Amendment Record Sheet below records the history and issue status of this document.

Table 1. Amendment Record Sheet

| ISSUE | DATE | REASON |
|-------|------------|---------------------------|
| 0.1 | dd/mm/yyyy | Initial in-progress draft |

Chapter 1. Introduction

1.1. Purpose and Scope

This document presents the <Template Service> Design for the Common Architecture.

1.2. Structure of the Document

Section 2 - Overview

Provides an over of the <Template Service> component, within the context of the wider Common Architecture design.

Section 3 - Design

Provides the design of the <Template Service> component.

1.3. Reference Documents

The following is a list of Reference Documents with a direct bearing on the content of this document.

| Reference | Document Details | Version |
|-----------------|---|--------------------------|
| [EOEPCA-UC] | EOEPCA - Use Case Analysis EOEPCA.TN.005 https://eoezca.github.io/use-case-analysis | Issue 1.0, 02/08/2019 |
| [EP-FM] | Exploitation Platform - Functional Model, ESA-EOPSDP-TN-17-050 | Issue 1.0, 30/11/2017 |
| [TEP-OA] | Thematic Exploitation Platform Open Architecture, EMSS-EOPS-TN-17-002 | Issue 1, 12/12/2017 |
| [WPS-T] | OGC Testbed-14: WPS-T Engineering Report, OGC 18-036r1, http://docs.opengeospatial.org/per/18-036r1.html | 18-036r1, 07/02/2019 |
| [WPS-REST-JSON] | OGC WPS 2.0 REST/JSON Binding Extension, Draft, OGC 18-062, https://raw.githubusercontent.com/opengeospatial/wps-rest-binding/develop/docs/18-062.pdf | 1.0-draft |
| [CWL] | Common Workflow Language Specifications, https://www.commonwl.org/v1.0/ | v1.0.2 |

| Reference | Document Details | Version |
|-------------------|--|-------------------------|
| [TB13-AP] | OGC Testbed-13, EP Application Package Engineering Report, OGC 17-023, http://docs.opengeospatial.org/per/17-023.html | 17-023, 30/01/2018 |
| [TB13-ADES] | OGC Testbed-13, Application Deployment and Execution Service Engineering Report, OGC 17-024, http://docs.opengeospatial.org/per/17-024.html | 17-024, 11/01/2018 |
| [TB14-AP] | OGC Testbed-14, Application Package Engineering Report, OGC 18-049r1, http://docs.opengeospatial.org/per/18-049r1.html | 18-049r1, 07/02/2019 |
| [TB14-ADES] | OGC Testbed-14, ADES & EMS Results and Best Practices Engineering Report, OGC 18-050r1, http://docs.opengeospatial.org/per/18-050r1.html | 18-050r1, 08/02/2019 |
| [OS-GEO-TIME] | OpenSearch GEO: OpenSearch Geo and Time Extensions, OGC 10-032r8, http://www.opengeospatial.org/standards/opensearchgeo | 10-032r8, 14/04/2014 |
| [OS-EO] | OpenSearch EO: OGC OpenSearch Extension for Earth Observation, OGC 13-026r9, http://docs.opengeospatial.org/is/13-026r8/13-026r8.html | 13-026r9, 16/12/2016 |
| [GEOJSON-LD] | OGC EO Dataset Metadata GeoJSON(-LD) Encoding Standard, OGC 17-003r1/17-084 | 17-003r1/17-084 |
| [GEOJSON-LD-RESP] | OGC OpenSearch-EO GeoJSON(-LD) Response Encoding Standard, OGC 17-047 | 17-047 |
| [PCI-DSS] | The Payment Card Industry Data Security Standard, https://www.pcisecuritystandards.org/document_library?category=pcidss&document=pci_dss | v3.2.1 |
| [CEOS-OS-BP] | CEOS OpenSearch Best Practise, http://ceos.org/ourwork/workinggroups/wgiss/access/opensearch/ | v1.2, 13/06/2017 |
| [OIDC] | OpenID Connect Core 1.0, https://openid.net/specs/openid-connect-core-1_0.html | v1.0, 08/11/2014 |

| Reference | Document Details | Version |
|------------|--|---------------------------|
| [OARec] | OGC API - Records - Part 1: Core, OGC 20-004, https://docs.ogc.org/DRAFTS/20-004.html | v1.0, 13/01/2020 |
| [OGC-CSW] | OGC Catalogue Services 3.0 Specification - HTTP Protocol Binding (Catalogue Services for the Web), OGC 12-176r7, http://docs.openeospatial.org/is/12-176r7/12-176r7.html | v3.0, 10/06/2016 |
| [OGC-WMS] | OGC Web Map Server Implementation Specification, OGC 06-042, http://portal.openeospatial.org/files/?artifact_id=14416 | v1.3.0, 05/03/2006 |
| [OGC-WMTS] | OGC Web Map Tile Service Implementation Standard, OGC 07-057r7, http://portal.openeospatial.org/files/?artifact_id=35326 | v1.0.0, 06/04/2010 |
| [OGC-WFS] | OGC Web Feature Service 2.0 Interface Standard – With Corrigendum, OGC 09-025r2, http://docs.openeospatial.org/is/09-025r2/09-025r2.html | v2.0.2, 10/07/2014 |
| [OGC-WCS] | OGC Web Coverage Service (WCS) 2.1 Interface Standard - Core, OGC 17-089r1, http://docs.openeospatial.org/is/17-089r1/17-089r1.html | v2.1, 16/08/2018 |
| [OGC-WCPS] | Web Coverage Processing Service (WCPS) Language Interface Standard, OGC 08-068r2, http://portal.openeospatial.org/files/?artifact_id=32319 | v1.0.0, 25/03/2009 |
| [AWS-S3] | Amazon Simple Storage Service REST API, https://docs.aws.amazon.com/AmazonS3/latest/API | API Version 2006-03-01 |

1.4. Terminology

The following terms are used in the Master System Design.

| Term | Meaning |
|-------|---|
| Admin | User with administrative capability on the EP |

| Term | Meaning |
|---|--|
| Algorithm | A self-contained set of operations to be performed, typically to achieve a desired data manipulation. The algorithm must be implemented (codified) for deployment and execution on the platform. |
| Analysis Result | The <i>Products</i> produced as output of an <i>Interactive Application</i> analysis session. |
| Analytics | A set of activities aimed to discover, interpret and communicate meaningful patterns within the data. Analytics considered here are performed manually (or in a semi-automatic way) on-line with the aid of <i>Interactive Applications</i> . |
| Application Artefact | The 'software' component that provides the execution unit of the <i>Application Package</i> . |
| Application Deployment and Execution Service (ADES) | WPS-T (REST/JSON) service that incorporates the Docker execution engine, and is responsible for the execution of the processing service (as a WPS request) within the 'target' Exploitation Platform. |
| Application Descriptor | A file that provides the metadata part of the <i>Application Package</i> . Provides all the metadata required to accommodate the processor within the WPS service and make it available for execution. |
| Application Package | A platform independent and self-contained representation of a software item, providing executable, metadata and dependencies such that it can be deployed to and executed within an Exploitation Platform. Comprises the <i>Application Descriptor</i> and the <i>Application Artefact</i> . |
| Bulk Processing | Execution of a <i>Processing Service</i> on large amounts of data specified by AOI and TOI. |
| Catalogue Service for the Web (CSW) | OGC standard that provides a simple HTTP interface for searching collections of descriptive metadata for data, services and relation information objects. |
| Code | The codification of an algorithm performed with a given programming language - compiled to Software or directly executed (interpreted) within the platform. |
| Compute Platform | The Platform on which execution occurs (this may differ from the Host or Home platform where federated processing is happening) |
| Consumer | User accessing existing services/products within the EP. Consumers may be scientific/research or commercial, and may or may not be experts of the domain |
| Data Access Library | An abstraction of the interface to the data layer of the resource tier. The library provides bindings for common languages (including python, Javascript) and presents a common object model to the code. |

| Term | Meaning |
|------------------------------------|--|
| Development | The act of building new products/services/applications to be exposed within the platform and made available for users to conduct exploitation activities. Development may be performed inside or outside of the platform. If performed outside, an integration activity will be required to accommodate the developed service so that it is exposed within the platform. |
| Discovery | User finds products/services of interest to them based upon search criteria. |
| Execution | The act to start a <i>Processing Service</i> or an <i>Interactive Application</i> . |
| Execution Management Service (EMS) | The EMS is responsible for the orchestration of workflows, including the possibility of steps running on other (remote) platforms, and the on-demand deployment of processors to local/remote ADES as required. |
| Expert | User developing and integrating added-value to the EP (Scientific Researcher or Service Developer) |
| Exploitation Tier | The Exploitation Tier represents the end-users who exploit the services of the platform to perform analysis, or using high-level applications built-in on top of the platform's services |
| External Application | An application or script that is developed and executed outside of the Exploitation Platform, but is able to use the data/services of the EP via a programmatic interface (API). |
| Guest | An unregistered User or an unauthenticated Consumer with limited access to the EP's services |
| Home Platform | The Platform on which a User is based or from which an action was initiated by a User |
| Host Platform | The Platform through which a Resource has been published |
| Identity Provider (IdP) | The source for validating user identity in a federated identity system, (user authentication as a service). |
| Interactive Application | A stand-alone application provided within the exploitation platform for on-line hosted processing. Provides an interactive interface through which the user is able to conduct their analysis of the data, producing <i>Analysis Results</i> as output. Interactive Applications include at least the following types: console application, web application (rich browser interface), remote desktop to a hosted VM. |
| Interactive Console Application | A simple <i>Interactive Application</i> for analysis in which a console interface to a platform-hosted terminal is provided to the user. The console interface can be provided through the user's browser session or through a remote SSH connection. |

| Term | Meaning |
|------------------------------|--|
| Interactive Remote Desktop | An Interactive Application for analysis provided as a remote desktop session to an OS-session (or directly to a 'native' application) on the exploitation platform. The user will have access to a number of applications within the hosted OS. The remote desktop session is provided through the user's web browser. |
| Interactive Web Application | An Interactive Application for analysis provided as a rich user interface through the user's web browser. |
| Key-Value Pair | A key-value pair (KVP) is an abstract data type that includes a group of key identifiers and a set of associated values. Key-value pairs are frequently used in lookup tables, hash tables and configuration files. |
| Kubernetes (K8s) | Container orchestration system for automating application deployment, scaling and management. |
| Login Service | An encapsulation of Authenticated Login provision within the Exploitation Platform context. The Login Service is an OpenID Connect Provider that is used purely for authentication. It acts as a Relying Party in flows with external IdPs to obtain access to the user's identity. |
| EO Network of Resources | The coordinated collection of European EO resources (platforms, data sources, etc.). |
| Object Store | A computer data storage architecture that manages data as objects. Each object typically includes the data itself, a variable amount of metadata, and a globally unique identifier. |
| OGC API - Records (OARec) | Draft OGC API standard that provides the capability to create, modify, and query metadata on the Web. The draft specification enables the discovery of geospatial resources by standardizing the way collections of descriptive information about the resources (metadata) are exposed. OARec is the successor to OGC CSW specification. |
| On-demand Processing Service | A <i>Processing Service</i> whose execution is initiated directly by the user on an ad-hoc basis. |
| Platform (EP) | An on-line collection of products, services and tools for exploitation of EO data |
| Platform Tier | The Platform Tier represents the Exploitation Platform and the services it offers to end-users |
| Processing | A set of pre-defined activities that interact to achieve a result. For the exploitation platform, comprises on-line processing to derive data products from input data, conducted by a hosted processing service execution. |
| Processing Result | The <i>Products</i> produced as output of a <i>Processing Service</i> execution. |
| Processing Service | A non-interactive data processing that has a well-defined set of input data types, input parameterisation, producing <i>Processing Results</i> with a well-defined output data type. |

| Term | Meaning |
|--|--|
| Products | EO data (commercial and non-commercial) and Value-added products and made available through the EP. <i>It is assumed that the Hosting Environment for the EP makes available an existing supply of EO Data</i> |
| Resource | A entity, such as a Product, Processing Service or Interactive Application, which is of interest to a user, is indexed in a catalogue and can be returned as a single meaningful search result |
| Resource Tier | The Resource Tier represents the hosting infrastructure and provides the EO data, storage and compute upon which the exploitation platform is deployed |
| Reusable Research Object | An encapsulation of some research/analysis that describes all aspects required to reproduce the analysis, including data used, processing performed etc. |
| Scientific Researcher | Expert user with the objective to perform scientific research. Having minimal IT knowledge with no desire to acquire it, they want the effort for the translation of their algorithm into a service/product to be minimised by the platform. |
| Service Developer | Expert user with the objective to provide a performing, stable and reliable service/product. Having deeper IT knowledge or a willingness to acquire it, they require deeper access to the platform IT functionalities for optimisation of their algorithm. |
| Software | The compilation of code into a binary program to be executed within the platform on-line computing environment. |
| Systematic Processing Service | A <i>Processing Service</i> whose execution is initiated automatically (on behalf of a user), either according to a schedule (routine) or triggered by an event (e.g. arrival of new data). |
| Terms & Conditions (T&Cs) | The obligations that the user agrees to abide by in regard of usage of products/services of the platform. T&Cs are set by the provider of each product/service. |
| Transactional Web Processing Service (WPS-T) | Transactional extension to WPS that allows adhoc deployment / undeployment of user-provided processors. |
| User | An individual using the EP, of any type (Admin/Consumer/Expert/Guest) |
| Value-added products | Products generated from processing services of the EP (or external processing) and made available through the EP. This includes products uploaded to the EP by users and published for collaborative consumption |
| Visualisation | To obtain a visual representation of any data/products held within the platform - presented to the user within their web browser session. |
| Web Coverage Service (WCS) | OGC standard that provides an open specification for sharing raster datasets on the web. |

| Term | Meaning |
|--|--|
| Web Coverage Processing Service (WCPS) | OGC standard that defines a protocol-independent language for the extraction, processing, and analysis of multi-dimensional coverages representing sensor, image, or statistics data. |
| Web Feature Service (WFS) | OGC standard that makes geographic feature data (vector geospatial datasets) available on the web. |
| Web Map Service (WMS) | OGC standard that provides a simple HTTP interface for requesting geo-registered map images from one or more distributed geospatial databases. |
| Web Map Tile Service (WMTS) | OGC standard that provides a simple HTTP interface for requesting map tiles of spatially referenced data using the images with predefined content, extent, and resolution. |
| Web Processing Services (WPS) | OGC standard that defines how a client can request the execution of a process, and how the output from the process is handled. |
| Workspace | A user-scoped 'container' in the EP, in which each user maintains their own links to resources (products and services) that have been collected by a user during their usage of the EP. The workspace acts as the hub for a user's exploitation activities within the EP |

1.5. Glossary

The following acronyms and abbreviations have been used in this report.

| Term | Definition |
|-------------|---|
| AAI | Authentication & Authorization Infrastructure |
| ABAC | Attribute Based Access Control |
| ADES | Application Deployment and Execution Service |
| ALFA | Abbreviated Language For Authorization |
| AOI | Area of Interest |
| API | Application Programming Interface |
| CMS | Content Management System |
| CQL | Common Query Language |
| CSW | Catalogue Service for the Web |
| CWL | Common Workflow Language |
| DAL | Data Access Library |
| EMS | Execution Management Service |
| EO | Earth Observation |
| EP | Exploitation Platform |
| FUSE | Filesystem in Userspace |

| Term | Definition |
|-------------|---|
| GeoXACML | Geo-specific extension to the XACML Policy Language |
| IAM | Identity and Access Management |
| IdP | Identity Provider |
| JSON | JavaScript Object Notation |
| K8s | Kubernetes |
| KVP | Key-value Pair |
| M2M | Machine-to-machine |
| OARec | OGC API - Records |
| OGC | Open Geospatial Consortium |
| PDE | Processor Development Environment |
| PDP | Policy Decision Point |
| PEP | Policy Enforcement Point |
| PIP | Policy Information Point |
| RBAC | Role Based Access Control |
| REST | Representational State Transfer |
| SSH | Secure Shell |
| TOI | Time of Interest |
| UMA | User-Managed Access |
| VNC | Virtual Network Computing |
| WCS | Web Coverage Service |
| WCPS | Web Coverage Processing Service |
| WFS | Web Feature Service |
| WMS | Web Map Service |
| WMTS | Web Map Tile Service |
| WPS | Web Processing Service |
| WPS-T | Transactional Web Processing Service |
| XACML | eXtensible Access Control Markup Language |

Chapter 2. Overview

TBD

Chapter 3. Design

3.1. Resource Catalogue Service

3.1.1. Overview and Purpose

The purpose of the resource catalogue service is to register metadata from either data sources or processing results where they are indexed for discovery and search by downstream applications.

3.1.2. Software Reuse and Dependencies

This service uses the COTS <https://pycsw.org> product with custom configuration.

3.1.3. Interfaces

This service provides the following interfaces:

1. OGC API - Records - Part 1.0, Core, version 1.0
2. OGC Catalogue Service for the Web, version 3.0.0
3. OGC Catalogue Service for the Web, version 2.0.2
4. OpenSearch with Extensions for Geo, Time and Earth Observation
5. Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)
6. Search/Retrieval via URL (SRU) search protocol

Request / response mechanisms are via HTTP GET (RESTful / KVP) and POST (XML/JSON). SOAP bindings are also supported.

3.1.4. Data

3.1.4.1. Configuration

This service uses an INI based configuration to allow configuration from various metadata sources parsed using various schemas into the backends.

3.1.4.2. Data flow

The resource-catalogue is implemented as a library as well as a server process. The library is implemented as a backend which is called by the registrar. When the registrar sends metadata to the resource catalogue, the metadata is processed according to the workflow and written to the underlying pycsw backend.

The resource-catalogue server process makes registered metadata discoverable via the interfaces described above.

The resource-catalogue supports the following item schemes:

- Sentinel-2 L2A: This expects an unpacked Sentinel-2 Level 2A SAFE package as the path input. It will parse all product and granule related metadata and generate an ISO 19115-2 document of the product as a single item
- STAC-Catalog: This scheme requires a path to a directory containing at least one STAC item file that will be read and parsed for metadata, generating an ISO 19115-2 document of the processing result as a single item

3.1.5. Applicable Resources

Chapter 4. Workflow

4.1. Backend ingest and metadata processing

The metadata repository is initialized via the configuration that is defined as part of the resource-catalogue. The registrar queries the backend to determine whether an incoming item has already been registered.

If the item is a product, both the `MTD_MSIL2A.xml` and `INSPIRE.xml` files are processed and used as input to generate an ISO 19115-2 document with spatial and temporal extents, platform and instrument information, as well as links to all data and granules.

If the item is a processing result, the STAC item is processed and used as input to generate an ISO 19115-2 document with spatial and temporal extents, platform and instrument information, as well as links to all results generated as a result of the process.

In both cases, if the metadata record already exists in the resource-catalogue, and if overwriting is enabled, the record will be sent as an update. Else, the record is sent as new metadata.

<< End of Document >>