GSTP-DAME

Data Intensive Technologies for Multi-mission Environments

Executive Summary

Doc Version No.	V1.1	
Document Id.:	pke120-29	
Checked by:	Cristoforo Abbattista	30/07/2018
Verified by:	Giulio Ceriola	30/07/2018
Contract Reference:	ESA Contract No. 4000117760/16/I-SBO	23/06/2016

Revision History

Log list			
Version	Author	Date	Note
1.0	Giulio Ceriola	12/07/2018	First issue
1.1	Giulio Ceriola	30/07/2018	Improved version for FR. Improved sections: 2.3.3 and 2.3.4. Rewritten section 2.4 Perspectives.

Index

1. Introduction

This document is the Executive Summary for the "GSTP DAME" project, and corresponds to the deliverable ES according to ESA Contract 4000117760/16/I-SBO.

The present document is released for the Final Review, in accordance with the Project Management Plan [AD3].

The present document summarises the findings of the project and it is be suitable for non-experts in the field and has been prepared so to be appropriate for publication.

1.1. Applicable Documents

AD1.   ESA Contract 4000117760/16/I-SBO and all the annexes and reference documents included inside.

AD2.   GSTP DAME, Contractor's Proposal (PK01-355708968-39)

AD3.   Project Management Plan (PK01-355708968-7), v1.1

1.2. Reference documents

[RD 1]   …

1.3. Acronyms and abbreviations

Code	Description
ASI	Italian Space Agency
Coll-IT	Italian Collaborative Ground Segment
C-TEP	Coastal Thematic Exploitation Platform
DHuS	Data Hub Servce Open Source software
EO	Earth Observation
ESA	European Space Agency
INFN	Istituto Nazionale di Fisica Nucleare (National Institute of Nuclear Physics)
OGC	Open Geospatial Consortium
OWS Context	OGC Web Services Context Document
NRT	Near Real Time
TEP	Thematic Exploitation Platform
VAP	Value Added Product
WCS	Web Coverage Service
WFS	Web Feature Service
WMS	Web Mapping Service
WPS	Web Processing Service

2. Executive Summary

2.1. Scope of the DAME project

The Data Intensive Technologies for Multi-mission Environments (DAME) has been conceived under the ESA General Support Technology Programme (GSTP) in order to support the ASI and ESA vision on the impact and benefits that the new IT technologies can bring to the Earth Observation and in particular to the ground segment architectures and their evolution as big data management infrastructure, in accordance with the paradigm "move the user to the data, not the data to the users".

Under this vision, the scope of the project is the development of the prototype of a software system for the exploitation of Earth Observation data, the setup of this system in the context of the Italian Collaborative Ground Segment (Coll-IT) infrastructure, as part, thus, of a federation of collaborative facilities in support of the Thematic Exploitation Platform (TEP) concept at national level, and the execution of a proof-of-concept demonstration of its operational concept through a selection of use cases, aimed at highlighting the benefits that the establishment of such a system can bring to different user communities.

2.2. Objectives and implementation strategy of DAME project

The high level objectives DAME are:

➢ the development of the prototype of a software system for the exploitation of Earth Observation data;

➢ the setup of this system in the context of the Italian Collaborative Ground Segment infrastructure;

➢ and the execution of a proof-of-concept demonstration of its operational concept through a selection of use cases, aimed at highlighting the benefits that the establishment of such a system can bring to different user communities.

These high level objectives have been translated into 5 technical objectives, listed here together with the strategy followed during their implementation:

❖ Evolution of the Data Hub Service (DHuS) Open Source Software

In terms of data management capabilities, the DHuS have been extended with a set of plugins for the handling of data related to the missions of interest for the scenario defined (COSMO-SkyMed; Prisma; Pleiades; Landsat 5/7/8) and the Value Added Products generated through the DAME processing facilities.

In terms of integration, the DHuS has been further customized, achieving the best level of integration with the other software components.

❖ Benchmark of data intensive technologies for archive organization, distribution, access

An optimal configuration of the DHuS has been implemented, achieving the best performances possible in the target environment in terms of archive organization, data access and distribution.

❖ Technology development and adoption of the necessary standards

Technology development has been ensured by the inclusion, in the DAME architectures, of software components developed within two important projects that Planetek is leading, the European INSPIRE Geoportal (JRC) and Fast4Map (ASI).

The overall interoperability of DAME platform received great benefit from the experience that Planetek has matured using the INSPIRE specifications, built on the top of OGC standards. An interoperability exercise has been also implemented, interconnecting with the actual C-TEP infrastructure.

❖ Procedure and tools for incorporation of new algorithms and products using a marketplace of processors

DAME includes a marketplace, where a user can upload an own processor (interface specifications are available). Once the processor is approved by the administrator, the processor is available to all the users which can run or download it. Furthermore the processor can be also accessible from outside the platform through the WPS standard.

❖ Innovative procedures for user interaction and user services

DAME architecture includes also a Collaborative Working Environment which relies on social networking capabilities to facilitate the interactions among users (e.g. for sharing products, processors or analysis) and between users and the DAME operational team.

2.3. Achievements of DAME project

2.3.1. DAME platform implementation

The DAME platform was fully implemented, successfully tested and demonstrated towards relevant users. A simple picture of its architecture is showed below.
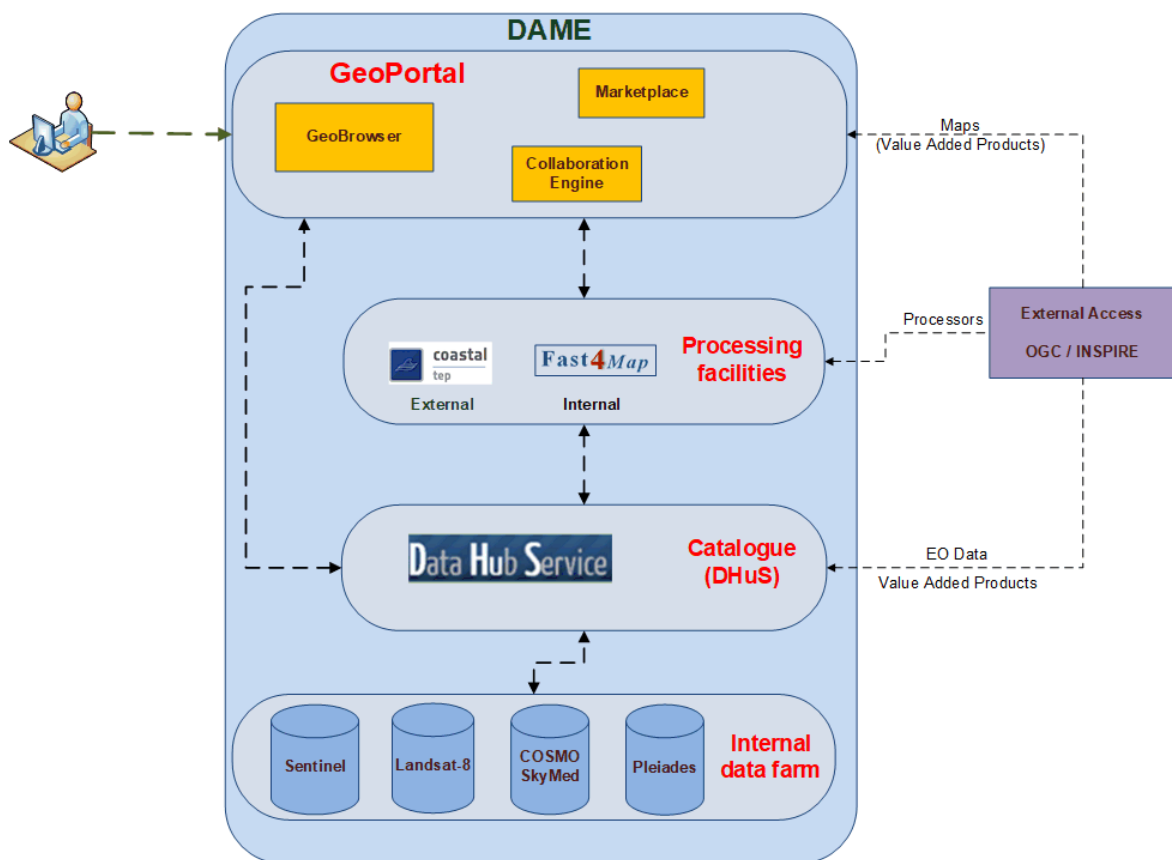

Figure 1. DAME platform architecture

Four main components can be identified.

Geoportal

This is the web interface for external users which allows to access all DAME functionalities through three modules: the GeoBrowser, the Collaboration Engine and the Marketplace.

From the GeoBrowser the user can search, visualize, download, select and use EO data or already generated products (indicated as Value Added Products or VAPs) to launch a given processor.

The geobrowser has also functionalities to display into the map the VAPs and to save the map composition as workspaces to be later re-loaded or shared. The VAPs and their metadata (all INSPIRE compliant) are available for visualization through OGC WMS protocol and can be downloaded using the OGC WFC and WCS protocols. Workspaces are imported/exported as OGC Web Services Context Document (OWS Context).

The second module is the Collaboration Engine, which allows to share a previously shared workspace with other DAME users. Each user to whom a workspace is shared with can also add its comments and/or answer to comments from other users.

The third module is the Marketplace. Through this module the user can provide a processor developed by itself in order to add it among the available processors (into the Processing Facility) and/or to share it with other users, allowing for commenting.
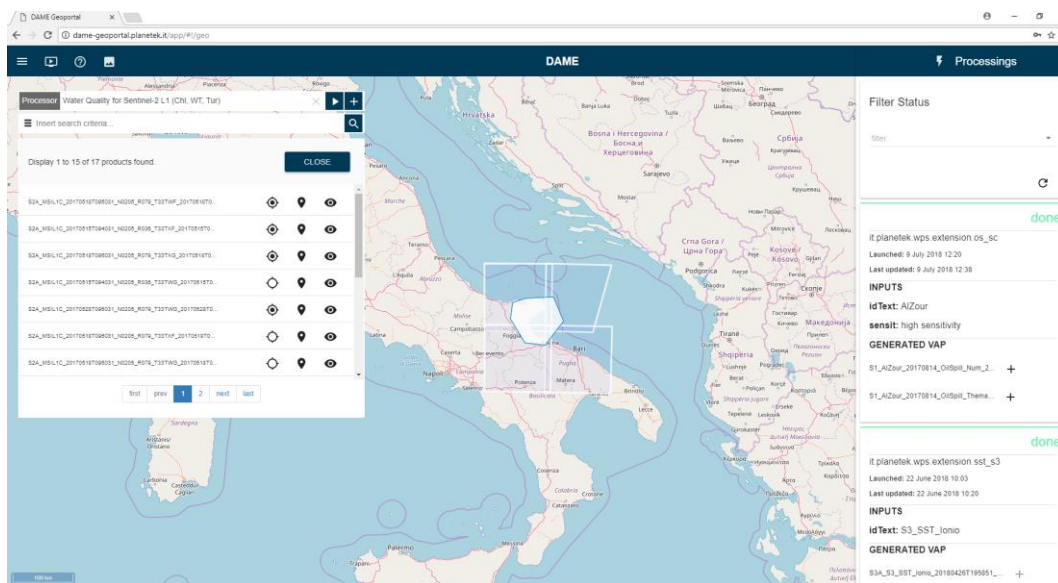

Figure 2. Screenshot of the Geobrowser.

Processing facilities

The processors that the users can run from the GeoBrowser have been implemented on two different processing facilities, by exploiting the OGC WPS protocol.

Fast4Map, part of the DAME platform, hosted all the processors but one, which was implemented on the C-TEP. All of them have been implemented as WPS, which allowed to interface in the same way a processor on Fast4Map internal facility and a processor on C-TEP external facility. The former can exploit the proximity with the EO data, while the latter exploited the virtual environment available on the C-TEP to easily develop an algorithm and pulish it as a WPS.

The next picture illustrates the concept aadopted in DAME and the workflow implemented to allow a user to run a processor.
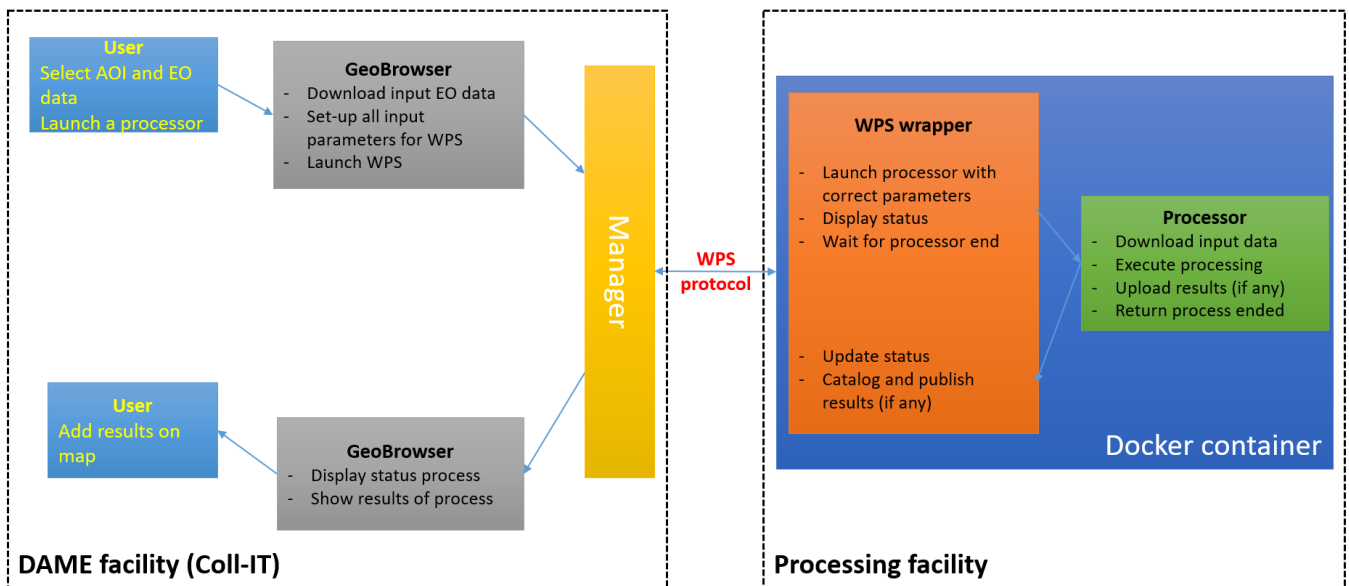
Figure 3. Workflow of the execution of a processor in DAME. Exploiting the WPS concept, the processing facility can be located anywere. If it is within the DAME facility then the download of the input data from the processor will be very fast.

Data Hub Service (DHuS) and Internal data farm

The Data Hub Service handles the data access task and is responsible, therefore, of ingesting from external and internal repositories and make searchable and available for processing EO data of the selected type (along with VAPs) in conformance with interoperability concepts based on the Open Data and on the OpenSearch protocols. This functional layer is directly connected to the Internal Data Farm, with inclusion of all the repositories to be internally setup with the data necessary for the demonstration of the system capabilities.

It included data from all the three Sentinel missions, from Landsat 8, COSMO SkyMed and Pléiades. Apart from the Sentinel data, ad-hoc add-ons have been developed within the DAME project in order to ingest, catalogue and serve the other EO data/missions[1]. Similarly and add-on for the VAPs have been developed.

2.3.2. Deployment at the Coll-IT

At the time of the project execution a provisory version of the Coll-IT was deployed at the ReCaS facility of the National Institute of Nuclear Physics (INFN) in Bari[2].

Consequently, with the support of ASI an agreement was done with INFN in order to allocate proper resources on the ReCaS cloud for the deployment of DAME.

All the components of the platform have been then successfully deployed as a set of docker containers which implement singularly the main DAME components. After deployment the system was run and demonstrated for about 2 months.

---

[1] The add-ons developed within DAME have been later published under the GNU Affero General Public License like the core DHuS software.

[2] See: https://www.recas-bari.it/index.php/en/

2.3.3. Demonstrated use cases

The use cases implemented within DAME and demonstrated to the users and corresponding actors are illustrated in Table 1

Utilisation and basic scenarios, identified at the start of the project, are also mapped. Basic scenarios – defined into the initial Statement of Work – are technology driven. Utilisation scenarios are user driven.

Use case	User category(ies)	Utilisation Scenario	Basic Scenario(s)
Analysis of water quality trends and anomalies	Scientific (non-EO) Institutional	Research	1. Data ingestion and cataloguing 3. On-demand processing chains
Analysis of coastline and coastal subsidence	Scientific (non-EO)	Research	1. Data ingestion and cataloguing 3. On-demand processing chains
Water quality at high spatial resolution	Service Providers Scientific	Service Development	1. Data ingestion and cataloguing 3. On-demand processing chains
Water quality from Sentinel-3, including NRT	Service Providers Scientific	Service Development	1. Data ingestion and cataloguing 2. NRT processing chains 3. On-demand processing chains
Oil spill detection, including NRT	Service Providers Scientific	Service Development	1. Data ingestion and cataloguing 2. NRT processing chains 3. On-demand processing chains

Table 1: Use cases implemented and demonstrated in DAME.

In order to define such use cases, the users and stakeholders of the DAME platform have been grouped in three categories: Scientific, Service Providers and Institutional.

The category of **scientific users** includes scientists working in research centres, universities and also in the research branch of private and public bodies, which are experts in the environmental domain and their duty is to study the methodologies and analytics for assessing the status of the environment in order to identify possible measures for intervention. Such scientific users are often not experts in the EO domain, but may benefit from the information derived from that in the execution of their analysis and modelling.

The category of **service providers** concerns added-value companies and private bodies which provide services and products based on EO data. Their activities include the definition and implementation of algorithms which process satellite data in order to generate different types of products related to the measure of variables linked to the marine environment in order to provide further elaborated information to their customers. Such users are mainly interested in the technological side of the EO added-value chain, starting from the access

serco

planetek
italia

and use of EO data, to the capability of processing a large amount of them, up to the provision by flexible systems.

Institutional users are public authorities which are entitled at different levels to enforce the national and European legislation, mainly affecting environmental issues, and to provide the due reporting for it. The Water Framework Directive (WFD), the Marine Strategy Directive (MSFD) and the Habitat and Birds Directive (Natura 2000) are examples of European directives that are usually implemented by local authorities.

These institutional users are the final "target" of the information derived from the EO data. In this respect, even if they can be considered potential end users of some of the DAME services, in the value added chain they should be considered as users/customers of service providers or research institutes which provide them further elaborated products (e.g. indices, analytics, etc.) to be direcly used for their reporting duties.

2.3.4. Benefits provided to the users

The DAME platform intends to provide various benefits to the two main involved users' categories, scientific and technological users. The main ones are:

- ✓ Easy and uniform access to a catalogue of different EO data sensors, which can be searched and selected by means of the DHuS using a standard protocol (OpenSearch) and then used as input to processors.
- ✓ Cataloguing of products generated by a processor, which allows to use them, by means of the DHuS, as input to another processor or to download and/or to share them.
- ✓ Easy and straightforward way to add a processor, by means of the marketplace in order to use it directly or to schedule its execution through the Geobrowser or to run it remotely using the WPS standard.
- ✓ Easy sharing of maps (context sharing)
- ✓ Social functionalities in order to exchange feedback among the users' community.

During the demonstration, two scientific users and two service providers tested and evaluated the different use cases implemented. The overall feedback received was highly positive and all the DAME concepts seem to be understood and well appreciated. Also the usability of the platforms in terms of easiness of use and as potential support to users' practices were judged very high and all improvements reported go towards a further capability of integration with such practices.

Concerning the thematic content of the platform also in this case all users provided a very positive feedback on the processors available and further highlighted the advantage of being able to use EO data from free missions like all Sentinel and Landsat ones as well as very high resolution commercial missions like COSMO SkyMed and Pléiades. In particular most appreciated processors were the ones based on SAR (in particular coastline and coastal subsidence) and the ones which allows to devire further information from the generated VAPs (statistical processors).

2.4. Perspectives

The successful demonstration of the DAME platform and the illustrated achievements, paves the road to possible further exploitations of the platform as a whole or of the derived know-how from single components.

In the original scope of the project, the DAME platform was conceived a possible evolution of the Coll-IT, as recalled by the objectives described in section 2.2. However the final DAME platform went beyond it. It demonstrated two main interoperability concepts, that allows to implement the Big Data approach:

- o the use of OGC and INSPIRE standards to made available EO data, EO derived products, processors and elaborated maps in a smooth and uniform way;
- o processors accessible using standard WPS interface and running on processing facilities based on the docker technology, allowing to move such facilities near the EO data.

This allows to consider DAME – in an EO ecosystem view (see Figure 4) – as a component of a "Platform Service Layer", that is an exploitation platform which provides basic functionalities like cataloguing, EO data searching and processing, VAP processing & archiving, marketplace, collaboration environment, etc. on which to build a tailored Exploitation Layer depending on the targeted community. So in this respect DAME platform can have a scope similar to the one of the current ESA Exploitation Platforms.

However thanks to its capability of archiving VAPs and of cataloguing multi-mission EO data, DAME can be also seen as a contributor to a "Resources Tier Layer", by re-ingesting its EO data and/or user-generated products (VAPs) into the EO eco-system. In this case only the mentioned component of DAME would be exploited within existing facilities like a DIAS.
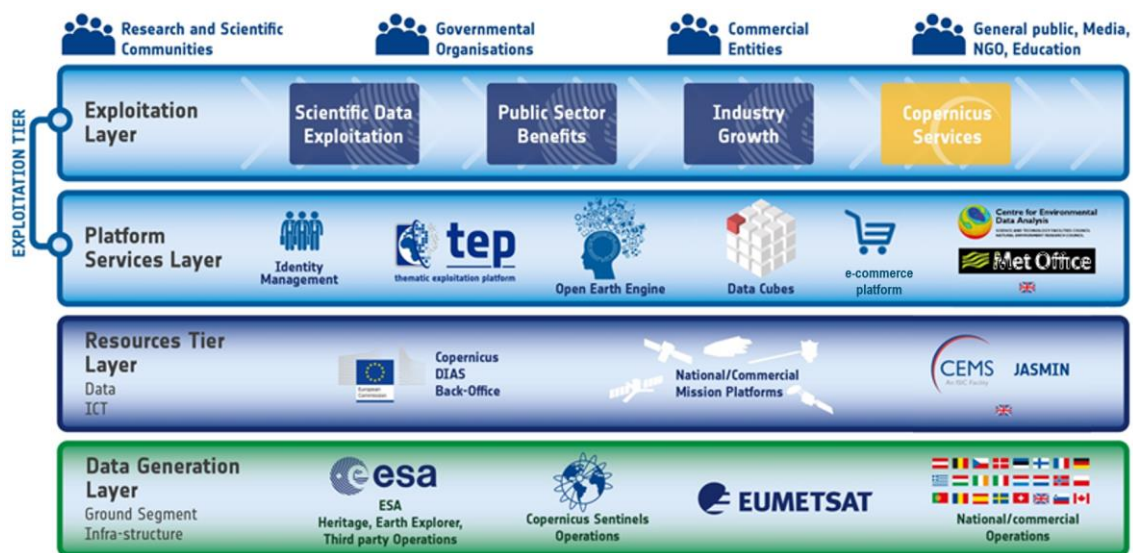
Figure 4. The EO Ecosystem: a platform-based economy, as presented by ESA in the EO Science for Society Information Day (22nd March 2018).

So the value of DAME in a platform-based economy relies both on DAME as a whole, but also as the use of specific components. So depending on the considered view, some possible DAME perspectives are the following ones.

➢ Use of the DAME platform at the final Coll-IT

The final Coll-IT will be released within the 2018 in the ASI premises in Matera. Due to the modularity of the platform as docker containers, DAME can be easily deployed in another environment, including the final Coll-IT. Planetek offers its availability to perform the deployment once the final environment is available. DAME will be immediately able to exploit the whole catalogue of data available (including also COSMO SkyMed mission), being the Coll-IT catalogue based on the DHuS.

Leveraging on the experience gained by Planetek due to its participation to the ongoing ASI CosteLAB project (development of an Italian Platform for the coastal risks), and with the collaboration between the DAME and C-TEP platforms, a couple of possible scenarios for setting-up a collaborative environment between DAME and COSTELab could be envisaged: in a first scenario DAME platform can be used as a backbone of COSTELab, for providing functionalities like access to multi-mission EO data, processing and cataloguing generated products and collaborative features. On this backbone both the virtual laboratory and the research activities and derived processors can be built up so evolving the DAME platform into the intended Italian thematic platform on Coastal Risks. In a second scenario DAME can be seen as a stand-alone platform at the Coll-IT that COSTALab can exploit to allow for including external processors (through WPS), for expanding the base of EO data and products that can be used, etc..

➢ Re-use of the DAME platform for future activities/projects

Considering the benefits that the DAME platform brought to the users, as mentioned in previous section, the whole platform or parts of it could provide added value to future activities and/or projects which require to setup a service and/or a sysem with functionalities of cataloguing EO data and products and of using them to run pre-defined or customized processors.

In particular, since the DAME platform will be already available to the CollIT, DAME could be used "as is" in CosteLAB, exploiting its DhUS interface and its docker modularity; the actual choice, among reuse part of DAME platform, or built from scratch another platform, will depend on ASI and the prime contractor of CosteLAB.

➢ Exploitation of demonstrated concepts and gained know-how is already in place within Planetek service portfolio, in particular for the proprietary Rheticus® platform[3].

The first two mentioned perspectives starts and build on ASI initiatives like the Coll-IT and the COSTELab project, however similar perspectives could be imagined in other contexts as for example projects in other countries or at European scale.

---

[3] See at: https://www.rheticus.eu/