H2020 EOPEN EASING COPERNICUS DATA & SERVICES EXPLOITATION

Leslie Gale (1), Bernard Valentin (1), Hakim Boulahya (1), Guido Vingione (2), Gabriella Scarpino (2), Laurence Marzell (2), Tudor Pettengell (2), Ilias Gialampoukidis (3), Stelios Andreadis (3), Stefanos Vrochidis (3), Ioannis Kompatsiaris (3), Dennis Hoppe (4), Li Zhong (4), Michele Ferri (5), Daniele Norbiato (5), Francesco Zaffanella (5), Ari Karppinen (6), Jani Tyynela (6), Petteri Karsisto (6), Haris Kontoes (7), Ioannis Papoutsis (7), Vasileios Sitokonstantinou (7), Thanassis Drivas (7), Hyun-Woo Jo (8), Hoonjoo Yoon (9)

[1] Space Applications Services, NV/SA Leuvensesteenweg 325, 1932 Sint-Stevens-Woluwe, Belgium
[2] Serco S.p.A. Via Sciadonna, 24, 00044 Frascati, Rome, Italy
[3] Information Technologies Institute, Centre for Research and Technology Hellas, Thessaloniki, Greece
[4] Höchstleistungsrechenzentrum Stuttgart, Universität Stuttgart, Nobelstraße 19, 70569, Stuttgart, Germany
[5] Autorità di Bacino Distrettuale delle Alpi Orientali, Cannaregio 4314, 30121 Venezia, Venice, Italy
[6] Finnish Meteorological Institute, Climate Research Programme & Space and Earth Observation Centre, P.O. Box 503, FI-00101 Helsinki, Finland
[7] Institute for Astronomy & Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens, Metaxa & Vas. Pavlou, GR-15236 Athens, Greece
[8] Korea University Environmental GIS/RS Center, Anam-ro 145, Seoul, Republic of Korea
[9] Sundosoft, Ltd., 1701-1706, 181 Gasan Digital 1Ro, Geumcheon-Gu, Seoul, Republic of Korea

Email: leslie.gale@spaceapplications.com; bernard.valentin@spaceapplications.com; hakim.boulahya@spaceapplications.com; guido.vingione@serco.com; gabriella.scarpino@serco.com; Laurence.Marzell@serco.com; Tudor.Pettengell@serco.com; eliasgj@iti.gr; andreadisst@iti.gr; stefanos@iti.gr; ikom@iti.gr; hoppe@hlrs.de; li.zhong@hlrs.de; michele.ferri@distrettoalpiorientali.it; daniele.norbiato@distrettoalpiorientali.it; francesco.zaffanella@distrettoalpiorientali.it; Ari.Karppinen@fmi.fi; Jani.Tyynela@fmi.fi; petteri.karsisto@fmi.fi; kontoes@noa.gr; ipapoutsis@noa.gr; vsito@noa.gr; tdrivas@noa.gr; ndeavor4a1@gmail.com; hjyoon@sundosoft.co.kr

KEY WORDS: Cloud Platforms, Big Data, Downstream, Interoperability, Distributed Processing

ABSTRACT: Pursuing Copernicus data and Service exploitation as part of the EO downstream sector requires innovative application of mature ICT solutions; they are essential to address the related "Big Data" issues involved in both EO and non-EO (such as meteorological data and data stemming from social media) data processing, also characterized by the Big Data four Vs: volume, velocity, variety, veracity.

Main objectives of EOPEN (EOPEN, 2019) include the provision of an easy to use environment to implement user services and capabilities to fuse Sentinel data with multiple, heterogeneous and Big Data sources; additionally, the involvement of mature ICT solutions in the Earth Observation sector addresses major challenges in effectively handling and disseminating Copernicus-related information to the wider user community, beyond the EU borders.

The EOPEN framework provides a scalable solution exploiting the back offices providing services to downstream service providers and consultancy organizations performing Big Data analytics. EOPEN is part of the Intermediate layer and complies with the DIAS requirements.

The EOPEN platform pilots comprise three Use Cases addressing crucial contemporary issues of flood risk assessment and prevention, food security and climate change. Practical examples related to the above topics have been considered for the implementation of the use cases, involving stakeholders for the definition of a supporting tool in the framework of a Joint Decision & Information Governance (JDIG) Architecture.

EOPEN is realized as a multi-platform layering framework to support public entities and decision managers aimed at scholars and professionals with various levels of expertise in usage of geospatial data as well as at the general public to assist them in their working activities. The capability of serving users in a wide range of diverse purposes resides in EOPEN's intrinsic flexibility; developer support is provided for the implementation of a given application. Still, with appropriate training, applications can be managed with a reduced effort of both end-users and service providers - the latter being among the main candidates to benefit from the platform. Moreover, using a single interoperable tool reduces the effort needed when dealing with new application areas, favouring an interdisciplinary approach and cooperation.

EOPEN will deliver to the user community a core set of open source products and services. During the project, EOPEN can be made freely available to implement use cases.

The paper will present the agnostic nature of EOPEN with respect to data platform, cloud environment, and application requirements. The paper will also include examples of usage of EOPEN through the presentation of the EOPEN Use Cases. Details on the EOPEN platform, its architecture and applied ICT solutions and methods is the focus of an EOPEN dedicated (open) session.

1 INTRODUCTION

Copernicus is the European Union's Earth observation programme coordinated and managed by the European Commission in partnership with the European Space Agency (ESA), the EU Member States and EU Agencies. The Copernicus programme delivers extensive insight about our planet by collecting data from a dedicated constellation of satellites, known as "Sentinels", as well as tens of third-party satellites known as "contributing space missions", complemented by "in situ" (meaning local or on-site) measurement data. The data and information products delivered by the Copernicus services are put at the disposal of citizens, public authorities and policy makers, scientists, entrepreneurs and businesses. By making the vast majority of its data, analyses, forecasts and maps freely available and accessible, Copernicus contributes towards the development of new innovative applications and services.

Copernicus has shifted EO data processing into the realms of Big Data requiring the innovative application of mature ICT solutions, state of the art processing methods and algorithm development. Additionally, the trend is to use both EO and non-EO (such as meteorological data and data stemming from social media) in applications delivering services to end-users. With increasing complexity the need for support has grown. If we are to expect EO experts and especially non-EO experts to effectively use the data to create services, new infrastructure and services are required to ease the exploitation of Copernicus data and services.

EOPEN is responding to the needs of the Copernicus programme and users by providing a framework for easy creation of platforms to implement and deliver user services and capabilities to fuse Sentinel data with multiple, heterogeneous and Big Data sources. EOPEN provides innovative solutions to integrate and combine mature ICT solutions to address major challenges in effectively handling, processing and disseminating Copernicus-related information to the wider user community in Europe and beyond.

2 THE EOPEN FRAMEWORK

EOPEN facilitates the use of Copernicus data and services by providing an infrastructure to support the Big Data processing life-cycle, allowing the chaining of value adding activities across multiple platforms. The platforms include the Thematic Exploitation Platforms (TEP) of ESA, commercial suppliers and the Copernicus Data and Information Access Services (DIAS).

The EOPEN Framework builds on and extends the Automated Service Builder (ESA ASB, 2017; Valentin, 2017). It provides a framework of core components and application specific extensions to allow easy creation of an Open Interoperable Platform for unified access and analysis of Earth Observation data in combination with data from other non-EO sources. To get away from the need to have EO experts, ICT experts and large IT support teams to create and use a platform EOPEN has placed an emphasis on "low coding" solutions coupled with access through APIs to support the more software programming oriented community. In this way, EOPEN caters to the needs of a broad user community, offering support for both expert and non-expert users.

Designed to be application agnostic as well as implementation platform agnostic an EOPEN implemented platform can be hosted on desktop, servers or cloud computing resources. Nominally, EOPEN is hosted on a cloud to take advantage of the EOPEN features for parallelisation and scaling of processing and access to Big Data lakes of Sentinel imagery and other data. EOPEN offers a transparent unified interface so that users can select the processing environment from a list of configured environments. In the EOPEN project, all EU DIAS platforms (5 in total) as well as the HPC/HPDA environments of the Stuttgart University are available to users.

EOPEN provides services to end-users, service developers, algorithm contributors and ultimately organisations who want to create their own exploitation platform. It provides a front-office for accessing other exploitation platforms and resource providers (e.g. DIAS, TEPs, MEPs) and a back-office as it allows other platforms to use services created and operational on EOPEN (processing using published algorithms, meteorological data, tweets, etc.).

Where EOPEN distinguishes itself strongly from other frameworks is that it has services and provides capabilities for all logical layers of the Big Data architecture (Data sources, Data preparation & storage, Analysis and Consumption). It also provides users of EOPEN with the capability to access other platforms offering services in the layers. In fact, in a single workflow, data and services can be accessed and used from multiple platforms providing users with a multi-platform layering solution.

3 EOPEN WORK PLAN

EOPEN has established a work plan (Figure 1) for supporting users in their implementation of a service offering support to service developers and end users. The process comprises three steps:
1. JDIG – identifying exactly what information is useful and how that usefulness can be applied (Section 4);
2. Service Development – working with the use case developers to implement additional core components, application specific extensions and the pilot use cases (PUCs) (Section 5);

3. Service Exploitation – performing the pilots to investigate how EOPEN has improved the capabilities of the service providers of the use cases and meets the goals set by the project (Section 6).
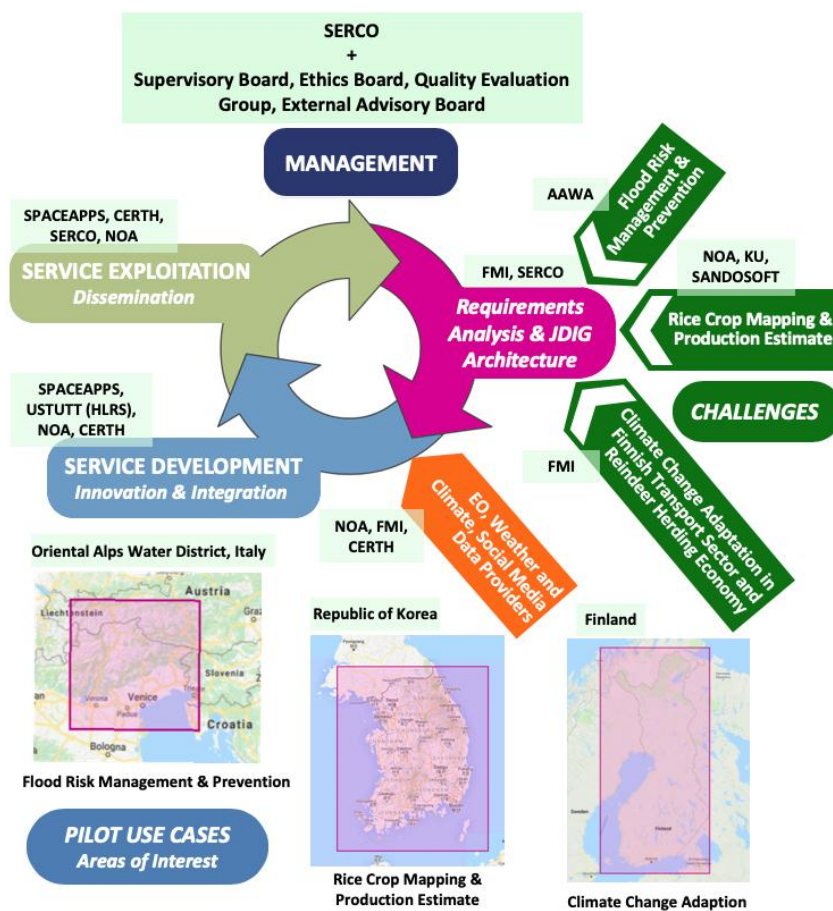

Figure 1, EOPEN Work Plan

4   JDIG APPROACH

As data needed by decision makers becomes more varied, more frequent and more easily accessible, the complexity that surrounds it has a corresponding increase. Even with the advances in machine learning and Big Data processing, those using the data to provide services or to make critical decisions, are often left with the significant challenge of identifying exactly what information is useful and how that usefulness can be applied, amongst the vast and dizzying array of data available. To address this challenge in EOPEN, the Joint Decision & Information Governance Architecture Framework (JDIG) has been developed.

The JDIG supports decision makers and other volume data users with the means to identify how they can improve their services and capabilities, through targeted improvements in their data and information identification, collection, communication and storage, and the governance activities that surround and enable it. Based upon an enterprise architecture framework: TOGAF 9.2 (TOGAF, 2019), and refined through previous EU funded H2020 projects UNITY (Marzell, 2019) & AUGGMED (H2020 AUGGMED, 2019), the JDIG takes a whole system approach to identifying the shared risks, challenges, objectives and outcomes of cross functional teams in tackling pertinent, collective and contemporary challenges. A Target Operating Model (TOM) is elicited through a scenario and use case driven approach that identifies the information needed to optimize different decisions, tasks and activities within a service, through upfront and continued dialogue with stakeholders across the decision-making chain as described through the scenarios.

Information is then traced through the JDIG journey and use-case value streams to identify whether it currently exists in some form but is being blocked by aspects or constraints relating to Governance, People, Process or Technology in the Current Operating Model (COM); or, whether it is new information that needs to be attained. The subsequent gap analysis between the COM and TOM, defines areas for organisational improvement; or requirements to procure/develop a system to obtain this information. In the context of EOPEN, this has been conducted to facilitate the three Pilot Use-Cases: flood risk assessment and prevention, food security and climate change that are briefly illustrated in Section 6. Figure 2, below, shows an example analysis taken from the flood risk assessment and prevention use-case.
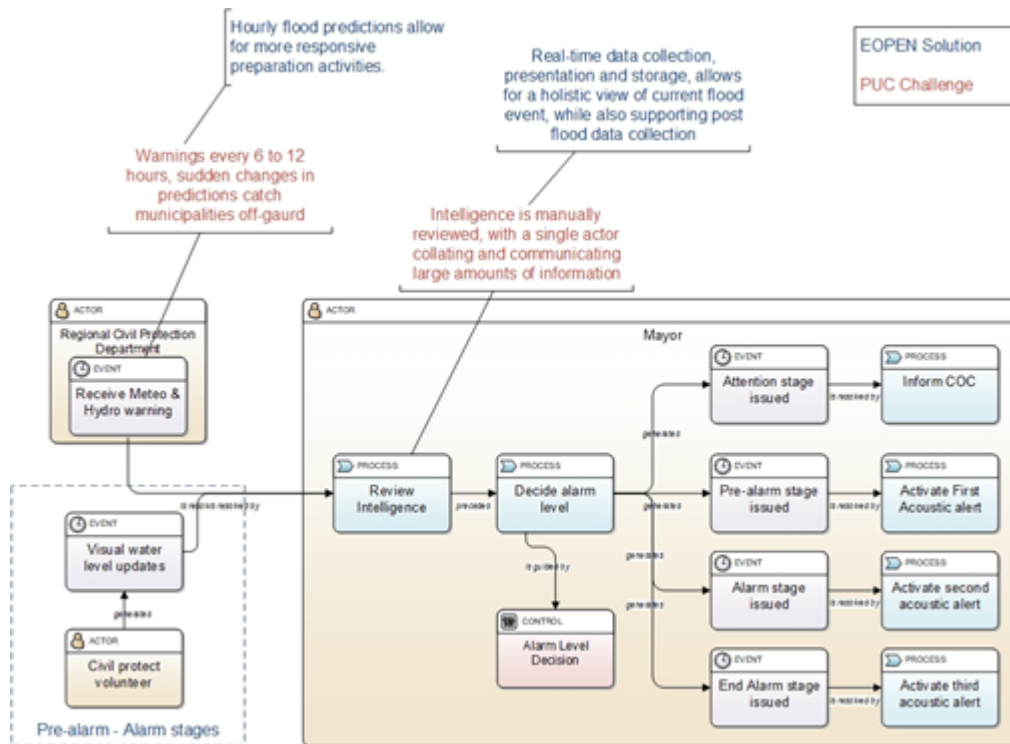
Figure 2, JDIG Example - PUC1 - Mayor Alarm Level Decision

5 SERVICE DEVELOPMENT

EOPEN falls into the category of integration Platform as a Service (iPaaS). It provides tools and technologies that support the execution of workflows, the development and lifecycle to manage the integration of user applications, management and monitoring of application execution, and essential cloud features such as multitenancy, elasticity and self-provisioning. EOPEN is an open architecture built on OGC Web Processing Services standards allowing the exchange and use of services between users, service providers, and integrators.

5.1 EOPEN Core Components

The EOPEN platform integrates a number of core components whose goal is to support the developers in importing, testing, and publishing services, and the end-users in discovering and executing these services. The steps for creating a service are the import of the algorithms (packaged as "processes") that implement specific processing tasks, their integration within a processor workflow, and the publication of this processor as a service. The service may then be discovered and executed by the end-users. All these operations are performed using Web interfaces.

Importing a new algorithm as a Process: The creation of a custom process by uploading, packaging and registering a new algorithm is made possible using the Process Importer. The tool is responsible for packaging the algorithm files together with the selected software dependencies (e.g. Python, R, SNAP Toolbox, Machine Learning libraries) within Docker images, pushing and registering the Docker image in a Docker registry where it will become available for remote deployment and execution, and registering the new process in the EOPEN platform, making it available for integration in the processor workflows.

Creating a new processor and configuring its workflow: Imported processes must be integrated in the workflow of pre-existing or new processors before being executed. This is accomplished by using the Web-based graphical Workflow Editor component. As shown on Figure 3, this is decomposed in several areas:
- On the left side are listed the available processes, including the built-in generic ones, and the ones imported by the service developers. The processes to integrate in the workflow must be selected from that list.
- The central area contains the canvas that allows building the workflow in an interactive manner. The boxes can be moved around and the data flow is configured by drawing connections from output parameters to input parameters. The editor verifies whether the connected parameters are of compatible data types or not.
- The right column contains two elements: A form that allows configuring the deployment constraints and the inputs of the process selected in the canvas: label, default value, read-only or editable, visible or not. Below this is provided a preview of the dynamically generated processor parameterisation form.
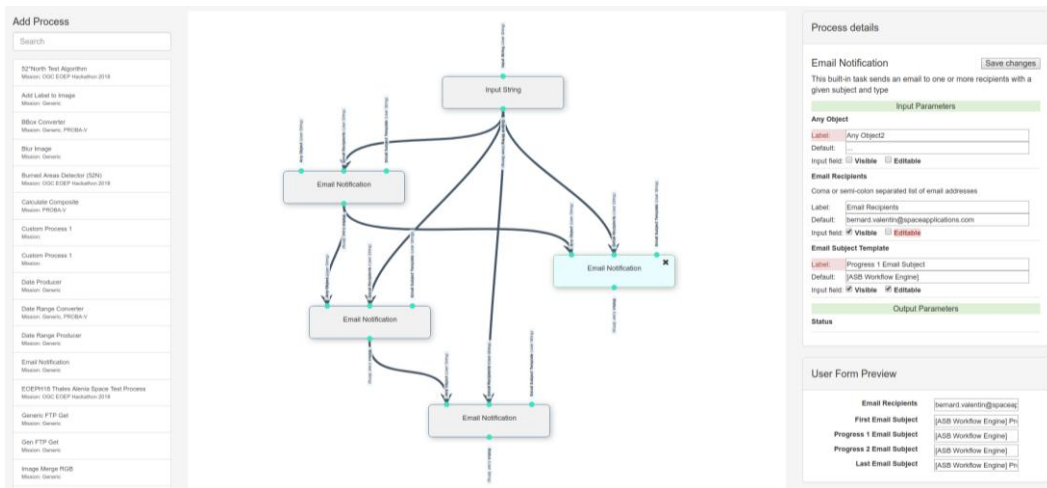
Figure 3, The EOPEN Platform Workflow Editor

Constraining process deployment to target platforms: In some situations, it may be the choice of the developers or the end-users to force specific processes to be deployed and executed in selected target platforms and worker nodes instead of letting the framework decide automatically based on built-in criteria. The reasons are multiple. For example, a process may need to access a specific type of data already available in a given platform, a platform may offer different hardware capabilities (such as providing GPUs), or the user may have special pricing conditions in certain platforms compared to others. The Workflow Editor includes a control that allows selecting, for each workflow process instance, one or more "Worker Selectors". Each of these Worker Selectors identifies a platform, a hardware requirement, or a data type requirement.

Sharing of processes and processors: The sharing of the processes and processors is realized using the role-based access control mechanism integrated in the EOPEN Platform. In this mechanism, each user is given a role in one or more "workspaces". A role indicates which actions the user may perform on the resources (including processes and processors) assigned to the same workspace.

A shared process may be integrated by other users in their own workflows. A shared processor may be selected and executed by the users that have the appropriate role in one of the processor workspaces.

Selecting and executing a processor: The End-User Interface component exposes to the users the resources they have the right to see and use. The component provides the tools for discovering, selecting, parameterising and executing the processors. The user may then monitor the progress, get the execution report generated when the execution is complete, and access the resulting products.

The component also includes a **Scheduler** that allows configuring the date(s) and time(s) at which a given processor must be executed. The available scheduling options include the single execution at a given future date and time, the repeated execution at a given time and interval (expressed in days), within a date range, and the repeated execution at a given time on specific month days.

Visualizing the data: A general-purpose yet customizable Web-based dashboard is being implemented in EOPEN which allows visualizing the collected and the generated data in various manners. In particular, the dashboard allows the users to select which graphical components they want to see on their screen and how they must be organised (editable layout).

Graphical components include GIS viewers to visualise geo-localised layers and features on background maps, generic graph builders to display data series, and notification tools that allow the users to communicate with their teammates and the system to instantly send notifications such as warnings to the users.

Non-visible key components: Other, non-visible core components also play key roles in the preparation and the execution of the processors. These include the Docker registry in which all the process Docker images are stored and from which they are deployed, the libraries for deploying and executing processes remotely in Cloud-based and High-Performance Computing (HPC)-based environments.

In EOPEN, distributed deployment and execution are performed in the five DIAS platforms as well as in the HPC environment provided by HLRS, Stuttgart.

To integrate with the DIAS platforms, and any other cloud-based environment, the EOPEN Platform makes use of the Mesos and Marathon container orchestration tools. These tools are aware of the available processing nodes and are thus able to automatically balance the load when new process executions are requested. They are also capable of

automatically discovering the changes in the processing environments, thus providing scalability to the platform. For example, as soon as a new node is detected, this is added to the pool of available resources and becomes available for running the next processes.

Mesos and Marathon also permit to implement the mechanism that constrains the deployment of processes on specific platforms and processing nodes. The mechanism matches attributes preliminary set on the processing nodes with the worker selectors that may be associated to each process in the processor workflows.

To integrate the High-Performance Computing systems of HLRS (HPDA and HPC) into the EOPEN platform, the open source cloud orchestration framework Cloudify has been adopted (Cloudify, 2019). The use of Cloudify enables model applications and services and automates the entire life cycle, including deployment on any cloud or data centre environment, monitoring all aspects of a deployed application, detecting issues and failure, manually or automatically remediating such issues, and performing on-going maintenance tasks.

5.2   EOPEN Application Specific Extensions

Application Specific Extensions are processes and services developed by the EOPEN team and made available to all users of EOPEN. Users will be encouraged to contribute to the application specific extensions collaborating with other users and contributing to enriching the resources available in EOPEN.

Support in selecting imagery: Regarding the retrieval of EO products, a process uses the Umbrella API (more information below) in order to search in different Sentinel data hubs for Sentinel-1, Sentinel-2, Sentinel-3 and Sentinel-5P products, store their metadata in a database named Sentinel Product Catalogue and also suggest the best hub to download the products from. The suggestion is based on measuring the download speed of each sub and a separate process is responsible for this scoring. A third process checks the products that no longer exist in the Sentinel hubs and deletes them from the Product Catalogue.

Identifying water pixels of satellite imagery: A first step towards the detection of flooded areas is to identify the water pixels of satellite imagery. To this end, a processor has been deployed that comprises three steps: the discovery of relevant EO products, their downloading and the water body masks generation. The Discovery process searches for Sentinel-1 and Sentinel-2 products for given criteria, such as area of interest and date, exploiting the aforementioned Umbrella process. In case that there are results, the Download process is triggered to get the products and then two alternative processes, one for Sentinel-1 and one for Sentinel-2, generate the waterbody masks for each product. In case of Sentinel-1 products, an automatic thresholding of the VH (or VV) band is used to detect the deep valley of the histogram separating inundated from non-inundated areas. For Sentinel-2 products, the Modification of Normalized Difference Water Index (MNDWI) is utilized, based on the Green and SWIR bands. In both cases, a GeoTIFF raster with the water body mask is generated (Figure 4).
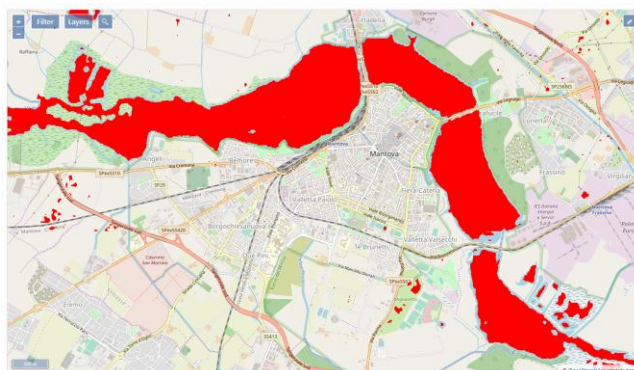

Figure 4, A generated waterbody mask as GeoTIFF raster

Social media crawling: A process constantly crawls new posts from social media, and in particular Twitter, that match predefined criteria, i.e. areas of interest, keywords, and interesting accounts. Before storing the tweets to a MongoDB database, the process also enriches them with information that is derived from analysis techniques about (i) the estimation of their validity to avoid fake news, (ii) the estimation of their relevance to the examined use cases (Moumtzidou, 2018), (iii) the detection of locations mentioned inside the text of tweets in order to geotag them, and (iv) the extraction of visual concepts from the images accompanying the posts.

Social media filtering: In order to present the social media data collected by the aforementioned process, a service has been launched that receives a set of filtering options, queries the database with the crawled tweets and returns all the necessary information. Figure 5 displays an example of how results are visualized in the EOPEN Dashboard, utilising this service. Available filters can be seen in the left panel, tweets that match the criteria in the middle (information includes text and image of the post, publication date, the author pseudonymised, the extracted visual concepts, and the estimated validity) and their extracted geolocation in the right.
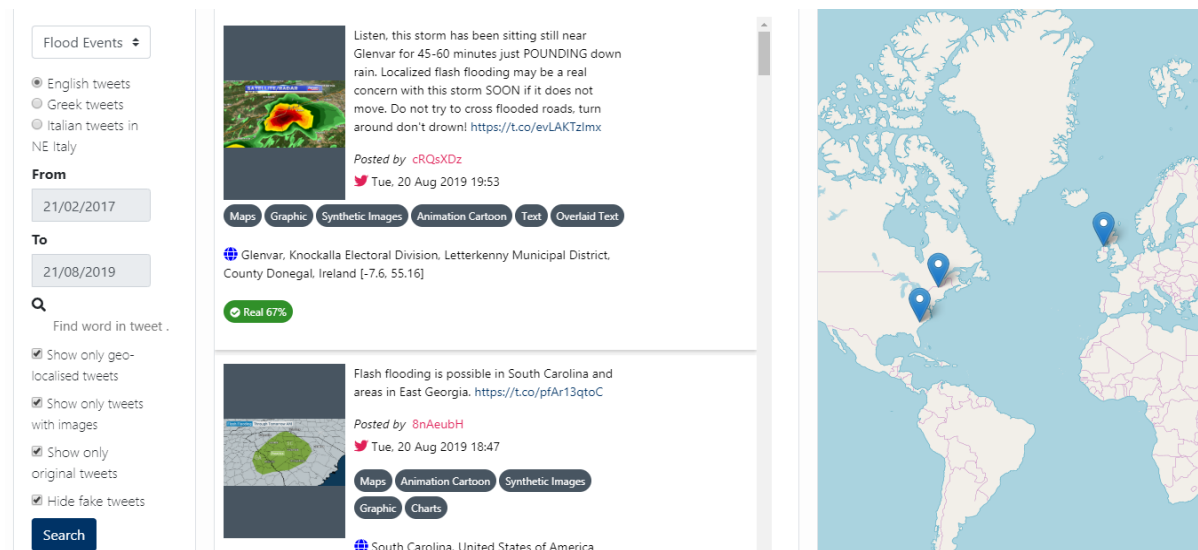
Figure 5, Filtering options and matching tweets

Social media event detection: Utilizing the collected tweets, the Event Detection process uses the fluctuation of the number of relevant tweets per day to discover events (Andreadis, 2018). Every day the z-score is calculated, which indicates how many standard deviations this day is from the sample's mean, assuming a Gaussian distribution (the sample is the previous thirty days). If a date's z-score is above a threshold of 3, then it is assumed that there is an event on that day and a JSON file is generated. Another process is able to convert this JSON file to a notification and instantly push it to the EOPEN Dashboard using the integrated Mattermost messaging service.

Social media community detection and topic identification: Further information can be derived from tweets, so two more social media-related services have been developed. The first service targets to detect communities of Twitter accounts whose posts have been crawled, based on their "following" interaction (Andreadis, 2017a). The inputs are the use case of interest and a date range, while the response includes the pairs of connected users, the detected communities, and a list of the most influential users.

The second service clusters textual information from crawled tweets into groups, which can be perceived as topics that are popular on social media (Andreadis, 2017b). The input is the use case and language of interest and the response is the clusters detected in the 300 most recent relevant tweets, in order to capture the trending topics. Each topic refers to the tweets it comprises and some keywords that reflect the most common terms mentioned in these tweets.

Umbrella API: The Umbrella API is a service that implements the Product Catalogue that contains the metadata of the Sentinel missions. The RESTful API service design took into consideration multiple factors in order to provide metadata in a simple, lightweight, and fast way. First of all, the heterogeneity of the stored data and the hubs they have been harvested from has been considered. In addition, the API catalogue was built to provide an easy, user-friendly URL interface for filtering data. It also gives the potential for safe filtering based on model attributes, at the same time that it supports related fields; so that filtering can be applied to related models. Finally, the designed API allows customizable pagination style simplifying the returning paginated data process.

5.3   A Multi-Layer Platform

Figure 6 depicts the overall layered architecture of the platform made of the following elements:
- The ASB Controller Node runs software components that provide the management interfaces and are responsible for the management of the platform and the orchestration of the processing chains (workflows).
- The Worker Nodes located in the top-right corner of the figure run in a private cloud. There is a limited number of them but they can be configured to meet very specific needs such as include GPUs.
- The Worker Nodes located in the bottom-left of the figure run in one or more public clouds. Many instances can be started to meet the processing needs. Also, when deployed in environments such as Copernicus DIAS platforms, a fast, local access to Copernic data is available.
- On the right of the public cloud are represented HLRS' HPC and HPDA (Data Analytics) platforms. Each of these is divided in two main parts: the Login Node where the users and the client applications land when connecting to the platform, and the Compute Nodes where the processing is performed when initiated using ad-hoc HPC-specific tools.
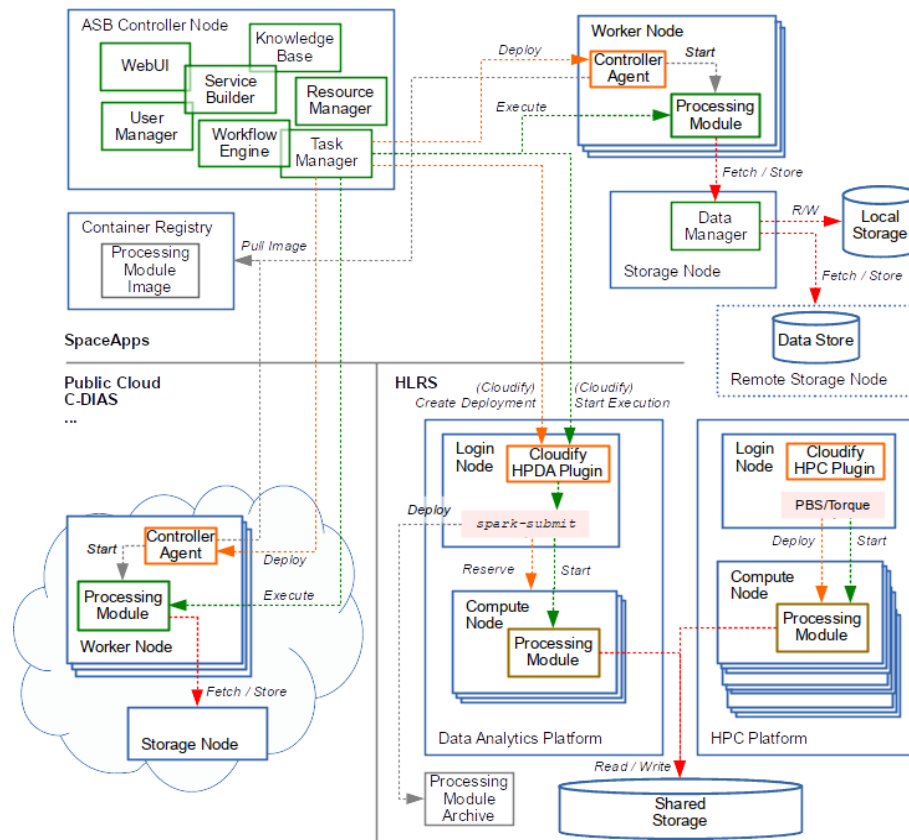
Figure 6, EOPEN Platform Deployment Diagram

6   SERVICE EXPLOITATION: THE PILOT USE CASES

6.1   PUC1 Flood Risk Assessment and Prevention

This Use Case is performed with the Eastern Alps River Basin District Authority (AAWA), responsible, among others, for applying European directives 2007/60/EC and 2000/60/EC. They have developed an early warning system (EWS) Flood Forecasting System of the predicted water level of the Bacchiglione river in Vicenza to inform regularly its Mayor. This use case addresses stakeholders from different offices, from local scale (municipality of Vicenza) to government scale: all authorities with specific

Figure 7, AAWA EWS Flowchart

roles during flood emergencies have been involved. Through EOPEN, AAWA EWS will access several additional input data for their hydrological model, available from Copernicus Global Land Services (CGLS) as well as an additional weather forecast, provided by the Finnish Meteorological Service, and maps of the flooded areas generated in the platform. Figure 7 is a flowchart of such a system.

Moreover, EOPEN will provide a scalable environment to run the AAWA hydraulic model to predict the magnitude of the flood, based on the discharge generated by the hydrological model, in addition to the EOPEN flood detected areas on specific areas of interest. Furthermore, AAWA carries out Flood Risk Management.

The map on Figure 8 shows the risk of floods from rivers, sea, surface water, groundwater and reservoirs for flood management.
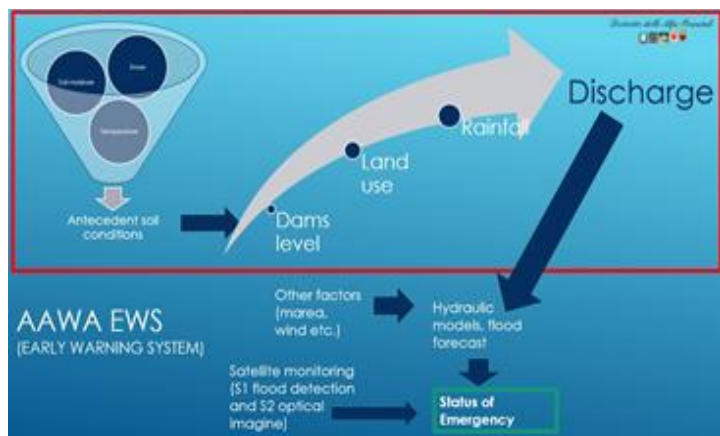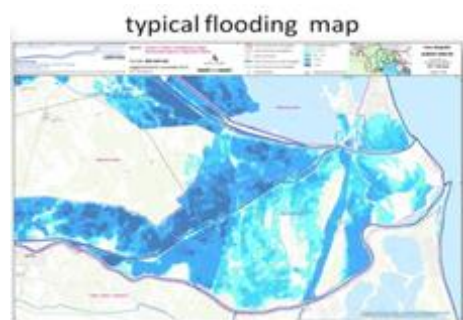

Figure 8, The PGRA (Flood risk management plan, DIR 2007/60/EC)

6.2   PUC2 Food Security

PUC2 deals with the food security monitoring in Republic of Korea focusing on the production of rice. Rice is systematically overproduced in the Republic of Korea, resulting in large storage costs, while at the same time there is underproduction of other major grains, making the country largely dependent on imports. Since 2000, rice has been constantly oversupplied and it is estimated that the annual supply of rice is over 280,000 tons. One of the main reasons for the overproduction of rice is the government's agricultural policy and active market intervention. The government would isolate the market if the price falls below the target price and compensate 85% of the price decreased. In addition local governments also have a lot of support policies for rice farmers. Producers prefer rice farming because of the relatively lower labour input due to high mechanisation and stable income compared to other crops.

Based on the collected user requirements and their subsequent translation into Earth Observation (EO) based products and services, it was decided to implement a food security monitoring system founded on the accurate mapping of rice and the succeeding estimation of rice yield for a given year of inspection. The main stakeholders of PUC2 include the Korea Rural Economic Institute, the National Institute of Agricultural Science, and Food Companies like CJ Group.

The first principal component of the implementation of PUC2 is the accurate, efficient and large scale mapping of rice in Republic of Korea. The rice mapping classification framework for PUC2 involves two different approaches to be compared in terms of effectiveness, efficiency and generalization. The first involves the utilization of a time-series of Sentinel-1 Synthetic Aperture Radar (SAR) images to train a Recurrent Neural Network (RNN) model using freely available land cover maps of past years, and then fit the model on a different Sentinel-1 time-series for the year of inspection. The second method uses a combined input dataset of both Sentinel-1 and Sentinel-2 (optical) imagery to train a Random Forest classifier based on a dynamically updated land cover map, through unsupervised clustering, for the year of inspection. Utilizing the rice mapping product and time-series of Sentinel derived vegetation indices, such as the Normalized Difference Vegetation Index (NDVI), linear regression models are applied to accurately estimate the rice yield for the year of inspection. Figure 9 illustrates the implementation workflow.

Figure 9, PUC2 implementation workflow

PUC2 attempts to develop a food security monitoring system that is transferable, reproducible and scalable. Therefore, within PUC2 we exploit exclusively open access data, models that generalize well, and develop distributed processing workflows, within a High Performance Data Analytics environment, to account for linear scaling-up and enable the efficient processing of big EO data.

Two EOPEN partners from the Republic of Korea are involved, namely the Korea University Environmental GIS/RS Center (KU-eGISRS) (Prof. Woo-Kyun Lee) and a private company, SUNDOSOFT; they have supported EOPEN participation to this conference; PUC2 is also the main focus of a dedicated EOPEN session.

6.3   PUC3 Climate Change

The focus for PUC3 is the Finnish Lapland and the challenges that climate change brings to the local reindeer herding livelihoods as well as for the infrastructure and transportation. The main stakeholders are the reindeer herders, reindeer researchers and the Finnish Transport Infrastructure Agency (FTIA).

The herders' livelihood depends directly on the environmental and seasonal variation in snow cover, snow depth, temperature and onset of snow melting. They need to adapt to sudden changes in these conditions, which can lead to changes in reindeer pastures or feeding problems. They utilize daily weather forecasts, but lack long-term estimates of snow depth and temperature. The reindeer researchers are also lacking long-term data of snow depth over time in herding-specific areas.

FTIA is the Finnish national agency responsible for the maintenance of public roads and monitoring their conditions. They use road-specific criteria for snow removal and sanding, and utilise weather forecasts and cameras on weather stations. They also prepare their budget based on yearly trends in climate. Due to climate change, close to freezing temperatures have become more common in winters, which leads to a continuous cycle of freezing and thawing on roads. This increases the need for sanding and costs of maintenance. FTIA would benefit from road temperature and thawing information especially in low-class roads, which have a lower priority.
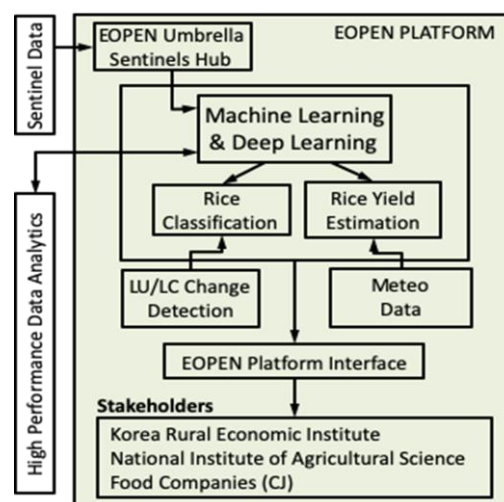
In EOPEN, we provide historical data of both air temperature and snow depth in Finland, which can be used to estimate trends in local areas. For monitoring purposes, we also provide continuous data on the land surface temperature through the Sentinel LST product and on the snow cover and thawing conditions through the FMI GlobSnow and Freeze/Thaw products. In addition to the data, the EOPEN provides tools to plot the data and estimate both temporal/areal statistics. Together, these enable the stakeholders and users to mitigate their problems and better adapt to climate change.

7 INTERMEDIATE RESULTS, CONCLUSIONS & FUTURE STEPS

The Pilot Use Cases are testing EOPEN's ability to provide a framework to leverage the core components and application specific extensions. We have demonstrated how EOPEN reduces the effort required to establish a service. Users benefit from being able to create a workflow, build processes to execute on a cloud, create systematic processing with scheduled or data driven events, use extensions to get (non-EO) data, pre-process and fuse the data into their services with a negligible amount of coding effort. Furthermore, the creation of dashboards is now supported by EOPEN allowing end users to access and execute services provided by the pilot use cases.

Now in its development phase, EOPEN aims to become an operational framework by the end of the project. Additional use cases are being sought. Indeed, an additional Use Case is currently being considered, using the EO satellite data (Sentinel 2) and weather forecast, available through the EOPEN platform, to run a Crop Water Demand Forecast module, which is an adaptation of a module implemented in the H2020 MOSES project (H2020 MOSES, 2018). In the concluding phase of the project, services developed in the project will be exploited and feedback from end users and stakeholders will be solicited. This activity will be supported by webinars and surveys, which will become available in the project website (EOPEN, 2019). Further dissemination and marketing actions will promote EOPEN's innovative approach to service implementation, in which data, products and computing resources are integrated as open sources tools, reducing the need for expert EO and IT staff and reducing computing effort making the EOPEN platform appealing to non-expert users who are then able to focus their effort in their fields of application expertise to deliver public and commercial services.

REFERENCES

- Andreadis, S., Gialampoukidis, I., Kalpakis, G., Tsikrika, T., Papadopoulos, S., Vrochidis, S. and Kompatsiaris, I., 2017a. A Monitoring Tool for Terrorism-Related Key-Players and Key-Communities in Social Media Networks. In 2017 European Intelligence and Security Informatics Conference (EISIC) (pp. 166-166). IEEE.
- Andreadis, S., Gialampoukidis, I., Vrochidis, S. and Kompatsiaris, I., 2017b, November. A topic detection and visualisation system on social media posts. In International Conference on Internet Science (pp. 421-427). Springer, Cham.
- Andreadis S., Gialampoukidis I., Fiorin R., Lombardo F., Norbiato D., Karakostas A., Ferri M., Vrochidis S. and Kompatsiaris I., 2018, November, Social Media Observations for Flood Event Monitoring in Italy over a One-Year Period. In 2nd International Conference Citizen Observatories for natural hazards and Water Management (COWM 2018).
- H2020 AUGGMED, The Serious Game Platform, Retrieved August 27, 2019, from http://auggmed-project.eu/.
- Cloudify, 2019. Product Overview: Accelerate your Cloud Native Transformation, Retrieved January 29, 2019, from https://cloudify.co/product/.
- ESA ASB, 2017. Automated Service Builder, Retrieved September, 2019, from https://www.spaceapplications.com/products/automated-service-builder-asb/.
- H2020 EOPEN project, 2017. Open interoperable platform for unified access and analysis of Earth Observation data, Retrieved September 2019, from https://eopen-project.eu/.
- H2020 MOSES, 2018. Satellite data ensures smooth flow of water resources, from https://cordis.europa.eu/project/rcn/196825/brief/en
- Marzell, L. & Brewster, B. 2017, October. Unity Framework for Trust, Retrieved August 27, 2019, from https://www.academia.edu/36137186/Unity_FW_for_Trust.
- Moumtzidou, A., Andreadis, S., Gialampoukidis, I., Karakostas, A., Vrochidis, S. and Kompatsiaris, I., 2018, April. Flood relevance estimation from visual and textual content in social media streams. In Companion Proceedings of the The Web Conference 2018 (pp. 1621-1627). International World Wide Web Conferences Steering Committee.
- TOGAF, 2019. The Open Group Architectural Framework Standard, Version 9.2 Overview, Retrieved on August 27th 2019, from https://www.opengroup.org/togaf-standard-version-92-overview.
- Valentin, B., Melcot, M., Gale, L., Mougnaud P., Iapaolo M, ESRIN, November 2017. ASB - A Platform and Application Agnostic Solution for Implementing Complex Processing Chains over Globally Distributed Processing and Data Resources. In: Conference on Big Data from Space - BiDS'17, Toulouse.