EOPEN: OPEN INTEROPERABLE PLATFORM FOR UNIFIED ACCESS AND ANALYSIS OF EARTH OBSERVATION DATA

*Guido Vingione[1], Gabriella Scarpino[1], Laurence Marzell[1], Tudor Pettengell[1], Ilias Gialampoukidis[2], Stelios Andreadis[2], Stefanos Vrochidis[2], Ioannis Kompatsiaris[2], Bernard Valentin[3], Leslie Gale[3], Woo-Kyun Lee[4], Wona Lee[4], Michael Gienger[5], Dennis Hoppe[5], Vasileios Sitokonstantinou[6], Ioannis Papoutsis[6], Charalampos Kontoes[6], Francesco Baruffi[7], Michele Ferri[7], Hoonjoo Yoon[8], Ari Karppinen[9], Ari-Matti Harri[9] **

[1]Serco S.p.A., [2]Information Technologies Institute, Centre for Research and Technology Hellas, [3]Space Applications Services NV/SA, [4]Korea University Environmental GIS/RS Centre, [5]University of Stuttgart – High Performance Computing Center Stuttgart, [6]National Observatory of Athens, [7]Autorità di bacino distrettuale delle Alpi orientali, [8]Sundosoft Limited, [9]Finnish Meteorological Institute

ABSTRACT

EOPEN (https://eopen-project.eu/) is a project which has received funding from the European Union's Horizon 2020 research and innovation programme under the topic EO Big Data Shift in 2017 and has a duration of 3 years, starting from November 2017. In this work, we present the concept of the project, its objectives and the lessons learnt after almost one year of project lifetime, as a follow-up to our previous project presentation at ESA BiDS'17 in Toulouse.

***Index Terms*—** Earth Observation, Copernicus, data fusion, interoperability, decision making, visual analytics.

1. INTRODUCTION

Earth Observation data access through the Copernicus data distributor systems has paved the way to monitor changes on Earth, using Sentinel data. One of the main objectives of EOPEN [1] is to fuse Sentinel data with multiple, heterogeneous and big data sources, to improve the monitoring capabilities of the future EO downstream sector. Additionally, the involvement of mature ICT solutions in the Earth Observation sector shall address major challenges in effectively handling and disseminating Copernicus-related information to the wider user community, beyond the EU borders. Relevant projects include the openEO [2], BETTER [3], PerceptiveSentinel [4] and CANDELA [5].

To achieve the aforementioned goals, EOPEN fuses Copernicus big data content with other observations from non-EO data, such as weather, environmental and social media information, aiming at interactive, real-time and user-friendly visualisations and decisions from early warning notifications. The fusion is also done at the semantic level, to provide reasoning mechanisms and interoperable solutions, through the semantic linking of information. Processing of large streams of data is based on open-source and scalable algorithms in change detection, event detection, data clustering, which are built on High Performance Computing infrastructures. Alongside this enhanced data fusion, a new innovative, overarching Joint Decision & Information Governance architecture is combined with the technical solution to assist decision making and visual analytics in EOPEN. EOPEN will be demonstrated in real use case scenarios in flood risk monitoring, food security and climate change monitoring, as also shown in Figure 1.

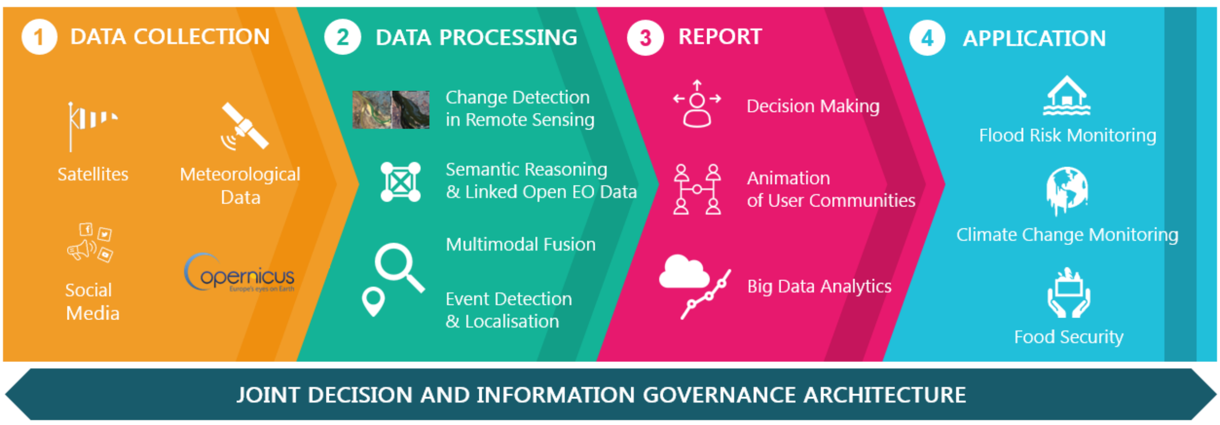

Figure 1: The EOPEN concept

* This work is supported by the EOPEN project, funded by the European Commission, under the grant agreement H2020-776019.

2. APPROACH

The overall objective of EOPEN is to provide a platform targeting non-expert EO data users (non-traditional user communities), experts and SME community that reveals and makes Copernicus data and services easy to use for Big Data applications by providing EO data analytics services, decision making and infrastructure to support the Big Data processing life-cycle allowing the chaining of value adding activities across multiple platforms.

To successfully address the Big Data challenges and to benefit from the services provided by ICT companies for accessing and processing Copernicus data, we are developing the EOPEN platform which delivers Copernicus data to non-traditional user communities, applying data compression and storage of EO and non-EO data (i.e. meteorological data, social media, linked open data), using cloud infrastructure and high performance computing (HPC), in order to fuse data from diverse sources and from different modalities (e.g. visual, textual or spatiotemporal).

Indexing Sentinel high resolution (HR) images will be performed to ensure fast access to their related content and assists pattern recognition and machine learning techniques by boosting their performance when they rely on solid multimedia indexing techniques. Data management techniques, community detection and tracking on the network for visualisation of usage activities and network analytics to reveal the key players (public or private users) and groups of users (communities) in the EO domain will be performed. All these heterogeneous sources of information are combined, through multimodal fusion, to semantically interpret the content of EO data resulting in efficient decision-making and visualisation in line with the proposed Joint Decision and Information Governance Architecture. The overall approach of EOPEN, aligned with the operational timeline is shown in Figure 2.

3. APPLICATION

Three use case scenarios are foreseen in EOPEN. In the following, we present the challenge that EOPEN deals with, in each use case considered.

3.1. Flood risk assessment and prevention

The pilot area, within the Italian Eastern Alps river District, comprises all the municipalities of the Local Risk District of Vicenza in Italy. This area is regularly affected by critical flooding from the Bacchiglione River and its tributaries. Planned flood defences remain largely unfinished, and a high risk of flooding therefore persists. Flood in the cities led to high levels of water in the streets, causing many problems such as the drowning of people, building damage and traffic problems. As indicated in the Flood Directive (2007/60/CE) water authorities should plan measures in order to aim at reducing risks by minimizing the possible damages effects and losses that may result.

In the Local Risk District of Vicenza AAWA provide flood forecasts warnings by running its Flood Forecasting System (AMICO) based on traditional meteorological data. The emergency phase is coordinated by the Mayors of the cities with a slow response time due to the lack of a data processing structure able to monitor in real time the evolution of the flooding in the territory (in term of flooded areas evolutions, impacts, damages). Currently these information come to decision-making authority through radio communications by people (civil protection volunteers) distributed in the territory. There is a need to

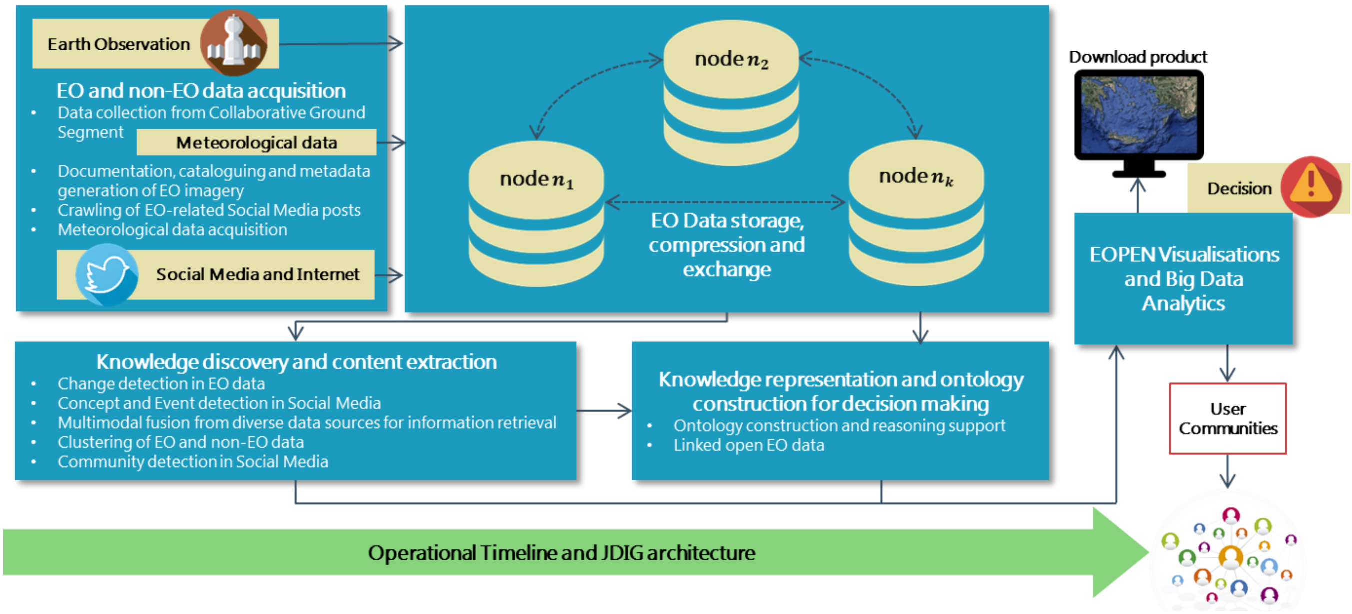

Figure 2: The EOPEN approach

provide faster and more effective emergency responses to extreme weather by increasing the speed of risk analysis.

3.2. Food Security through Earth Observation

"Food Security" is a denomination introduced by the Food and Agriculture Organization (FAO) of the United Nations. The problem is really complex and comprises several different components (food access, distribution, food supply stability, use of food). There are many recent examples that show the problem and more precisely food crises, for instance famine in the horn of Africa (2011) and the critical need to deliver timely food security information to decision-makers. Satellite data have been used to detect and monitor severe agricultural events since 1972 on the occasion of the extreme drought that took place in Russia. Improvements have been made in the spectral, spatial and temporal resolutions of Earth observation (EO) systems since then. Copernicus program and Sentinels' missions are the most ambitious Earth observation initiative and have a great impact and contribution also in the field of food security.

3.3. Monitoring Climate Change through Earth Observation

The climate can be defined as average weather conditions in terms of the mean and variance of temperature, precipitation and wind over a period of time. In climate studies, the averages of these parameters are normally calculated through an averaging filter spanning 30 years. Currently, it is scientifically clear that the climate is changing and the temperature is rising. Precipitation patterns and the frequency of occurrence of storms are changing as well.

The climate change manifests itself most visibly and rapidly at high latitudes. Hence a regional pilot area is established around Finnish borders. The most prominent and clear indicators of climate change are the atmospheric temperature and the average annual snow cover, which is consistent with warmer global temperatures. These parameters have been monitored by the Finnish Meteorological Institute (FMI) for decades, even for more than 100 years. Hence we use the ground-based climate observations of the Finnish Meteorological Institute dating back for more than 100 years (FMI Climate Services Archives) and satellite observations, e.g., by EUMETSAT, ESA, NASA and NOAA. A major project asset is the access to all FMI data which have been made avilable to support open access policy.

4. ILLUSTRATIONS

In this section we present some indicative applications that EOPEN offers to its end user community. EOPEN collects EO data and combines them with weather forecasts and Twitter posts. A social media image is automatically annotated to extract knowledge (concepts), using a technique based on Deep Convolutional Neural Networks (DCNNs), as shown in Figure 3.

Figure 3: Concepts extracted from a social media image

Moreover, the location of a tweet is rarely included (less than 5%) in the stream of posts, hence EOPEN considers automatic estimation of an event location from Twitter content and positioning on a map through an external ontology and linking of data. This functionality supports the situational awareness of an authority that would like to fuse citizen observations with EO products (Figure 4).

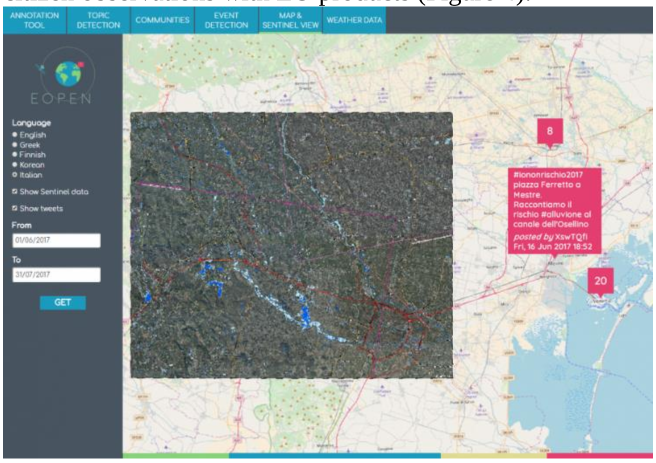

Figure 4: Localised events from social data add value to EO data products, which can both appear in a GIS view.

Satellite images often include a road network or a part of it. In case of an extreme weather event, such as a severe flood event, passing from one part of the road to another is not possible. EOPEN offers the possibility to estimate passable or not passable parts of a road network, with high accuracy, using a Residual Neural Network (ResNet), where some training took place on the annotated dataset of the Multimedia Satellite task of MediaEval2018. The analysed images keep their original georeferenced information, so as to be visualised as a GIS layer.

Figure 5: Various inputs to the road passability service

doi:10.2760/848593

The plan is to employ a Region proposal Neural Network to first detect the road parts (Figure 5) and then apply a binary classification algorithm to infer whether the road is passable or not.

Finally, EOPEN clusters similar content into groups of similar items. Similarity can be defined in many different ways (cluster by concept, location, event, user, etc.) even if refers to EO or non-EO data streams. The management of large and highly heterogeneous content requires scalable techniques that may take advantage of the existing European HPC resources, being also in line with the recent advances in the development of the four Data and Information Access Services (DIAS) and among which the ONDA DIAS service (https://www.onda-dias.eu/cms/) will be used for supporting the project validation. EOPEN develops algorithms using parallel programming techniques and libraries to boost scalability of the user applications and being executed on a HPC infrastructure.

5. IMPACT

The societal, technical and scientific and economic impact of EOPEN is eminent and briefly described as follows.

5.1. Societal Impact

EOPEN provides a means to perform analysis that is not yet available. Three use cases demonstrate how EOPEN can be used to address societal challenges fully addressing EO-2-2017 call's requirement optimising the use of Copernicus data by non-traditional user communities to meet societal challenges. In particular, activities under the societal challenge for climate action, environment, resource efficiency and raw materials focus on GEOSS, much like EOPEN's use cases and services that stimulate past ICT and EO activities coupled with the new Sentinel data and knowledge from various sectors.

5.2. Technical and Scientific Impact

In comparison with the data distribution system of the US, the EU data and products distribution and sharing infrastructure system looks less effective, especially in terms of facilitating small companies to play with data and create marketable services and products out of it. Copernicus lacks a holistic approach to data management because data distribution is based on fragmented data sources. EU Copernicus data and products are spread in many different archives, formats and portals. The Copernicus core service segment (with 6 thematic areas and relative product portfolios), the EUMETSAT Satellite Application Facilities network (with 8 Application facilities and product catalogues) and the ESA Sentinels Open Access Hub portal represent just three examples of different, and sometimes overlapping, large product repositories, hosting huge amounts of information that are difficult to explore and navigate into. The lack of an adequate solution for combining data from multiple sites with non-space derived data has given rise to Thematic Exploitation Platforms and networks of them. The European Space Agency (ESA), on behalf of the European Commission, launched the DIAS services. The DIAS provides a scalable computing and storage environment for third parties. With references to the afore-mentioned architectures, the EOPEN combines state-of-the-art technical solutions from the EO domain with mature ICT technologies, to deliver an efficient orchestration of services and modules, infrastructure agnostic, which are offered to the end user, without his/her need to have experience in downloading and processing Copernicus EO products. In EOPEN, data harmonisation and standardisation have highest priority in order to foster all three use cases covering flood risk management, food risks from environmental factors and climate changes, serving as a worldwide solution inside and outside the EU.

6. CONCLUSION

In this work we briefly present the EOPEN concept and approach, both at the platform level and use case application. The purpose of this work is to demonstrate the current status of the development of EOPEN and to set a solid basis for the next two years of implementation, verification, validation, evaluation and demonstration.

REFERENCES

[1] EOPEN project: https://eopen-project.eu/
[2] openEO project: http://openeo.org/about/
[3] BETTER project: https://www.ec-better.eu/
[4] PerceptiveSentinel: http://www.perceptivesentinel.eu/
[5] CANDELA project: http://www.candela-h2020.eu/
[6] I. Gialampoukidis, A. Moumtzidou, M. G. Scarpino, G. Palumbo, S. Vrochidis, I. Kompatsiaris, F. Zaffanella, D. Norbiato, M. Ferri, & G.Vingione, "Earth Observation and Social Multimedia Data Fusion for Natural Hazards and Water Management: The H2020 EOPEN Project Paradigm", 2nd International Conference Citizen Observatories for natural hazards and Water Management, Venice, 27-30 November 2018
[7] G. Vingione, L. Marzell, E. Cadau., I. Gialampoukidis, S. Vrochidis, I. Kompatsiaris, B. Valentin, M. Melcott, L. Gale, W.-K. Lee, S. Woo, M. Gienger, D. Hoppe, C. Kontoes, I. Papoutsis, M. Ferri, F. Baruffi, J. Yoon, H. Yoon, A. Karppinen, A.-M. Harri, "THE H2020-EO EOPEN PROJECT", INSPIRE Conference 2018, 18-21 September 2018, Antwerp, Belgium
[8] G. Vingione, L. Marzell, E. Cadau., I. Gialampoukidis, S. Vrochidis, I. Kompatsiaris, B. Valentin, M. Melcott, L. Gale, W.-K. Lee, S. Woo, M. Gienger, D. Hoppe, C. Kontoes, I. Papoutsis, M. Ferri, F. Baruffi, J. Yoon, H. Yoon, A. Karppinen, A.-M. Harri. "The H2020-EO EOPEN project". ESA Conference on Big Data from Space 2017, 28-30 November 2017, Toulouse, France.

doi:10.2760/848593