DEEPCUBE: Explainable AI Pipelines for Big Copernicus Data

by Chiara Gervasi, Alessia Ferrari, Ioannis Papoutsis, Souzana Touloumtzi

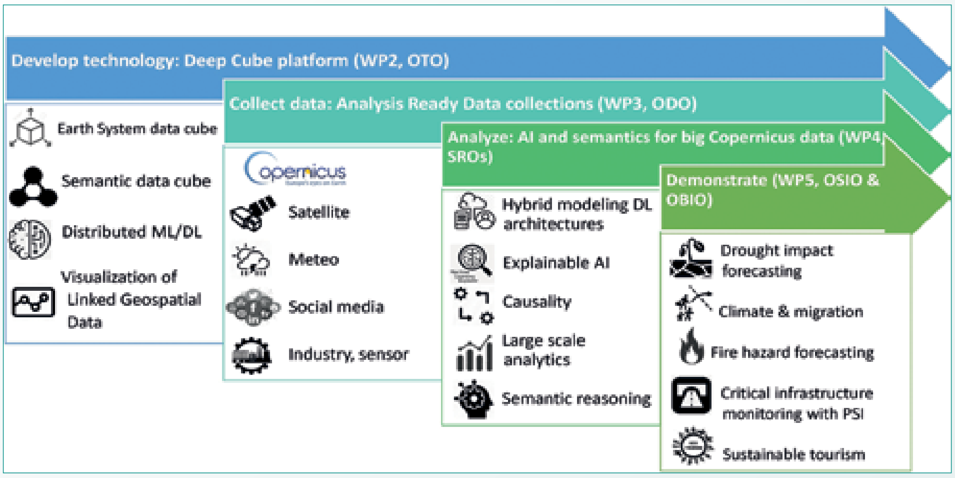

Figure 1: DeepCube implementation workflow from technology development to demonstration in Use Cases.

DeepCube is a Horizon 2020 project implemented by 9 European partners coordinated by the National Observatory of Athens. It will unlock the potential of big Copernicus data with Artificial Intelligence and Semantic Web technologies, aiming to address problems of high socio-environmental impact.

According to the recently published (19 June 2020) European White Paper on Artificial Intelligence (https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf), Artificial Intelligence (AI) is a strategic technology that, among other things, "offers important efficiency and productivity gains that can strengthen the competitiveness of European industry and improve the wellbeing of citizens", while "it can also contribute to finding solutions to some of the most pressing societal challenges, including the fight against climate change and environmental degradation, and the challenges linked to sustainability and demographic changes". The same paper acknowledges Earth Observation (EO) and space as domains in which the race for global leadership is open.

DeepCube (https://deepcube-h2020.eu/) is a 3-year (January 2021 - December 2023) Horizon 2020 project that leverages advances in the fields of AI and Semantic Web to unlock the potential of the vast amounts of data produced by the Copernicus program. DeepCube has the goal to address problems of high environmental and societal impact while enhancing our understanding of Earth's processes correlated with Climate Change, in order to respond to the urgent challenges addressed by the EU Green Deal. To achieve this, the project develops an open, scalable and interoperable platform, integrating novel Information and Communication Technologies such as the Earth System Data Cube, the Semantic Cube, the Hopsworks platform for distributed Deep Learning, and state-of-the-art data visualization tools, providing solutions for all phases of an EO-based AI pipeline, from data ingestion and organization in data cubes, to feature engineering, semantic reasoning and visualization. In addition, DeepCube tests a hybrid modeling approach for Earth System Science, combining data-driven modeling bound by physical parameters, further enhanced through explainable AI and Causality for "physics-aware" AI applications.

Contextual background

DeepCube contributes to the European Commission's Destination Earth, an initiative that aspires to create a digital model of the Earth's physical resources, to help us plan and prepare for major environmental degradation and disasters due to Climate Change. To support this, the European Space Agency has put forward the concept of Digital Twin Earth to develop a dynamic, digital replica of our planet. A model powered by observations that provides an accurate representation of the

Figure 2: TRE Altamira's wide area monitoring over different countries around the world - the UK, Denmark, France and Japan.

The Copernicus program
The Copernicus Earth Observation program, providing free, open, and high quality data about our planet at large scales, is believed to be a game changer for both science and the industry. Today, Copernicus is producing 15 terabytes of data every day, while every product is downloaded on average 10 times. However, the availability of the sheer volume of Copernicus data outstrips our capacity to extract meaningful information. That is why the Earth Observation community is in need of technology enablers to propel the development of entirely new applications at scale.

past, present and future changes of our planet to support us in preparing a better response to future challenges.
We consider DeepCube as a showcase of the Digital Twin Earth potential, addressing all of its elements from data to infrastructure, to technology, to Research & Development, to new business models, with Artificial Intelligence and big Copernicus data at its core.
DeepCube is driven by the scientific and business questions behind its Use Cases. What makes its applications different is that they serve non-traditional use cases, penetrate untapped markets, exploit unique datasets and employ new AI architectures.

The DeepCube Use Cases

Five Use Cases (UCs) have been designed by DeepCube's partners to showcase the innovative technologies used in this project: two on business, two on earth system sciences, and one on migration.
UC1: Forecasting localized extreme drought and heat impacts in Africa
UC2: Climate induced migration in Africa
UC3: Fire hazard short-term forecasting in the Mediterranean
UC4a: Automatic volcanic deformation detection and alerting
UC4b: Deformation trend change detection on PSI time-series for critical infrastructure monitoring
UC5: Copernicus services for sustainable and environmentally-friendly tourism
[*UC4 is under the umbrella of SAR (Synthetic Aperture Radar) interferometry and consists of the smaller-scale UC4a on volcanic risk reduction and the larger-scale UC4b that integrates measurements from in-situ geodetic and other information with InSAR (Interferometric Synthetic Aperture Radar) time-series*].

The UCs are the core of DeepCube and drive the development on both the technology and the research sides. The project approach is to co-design

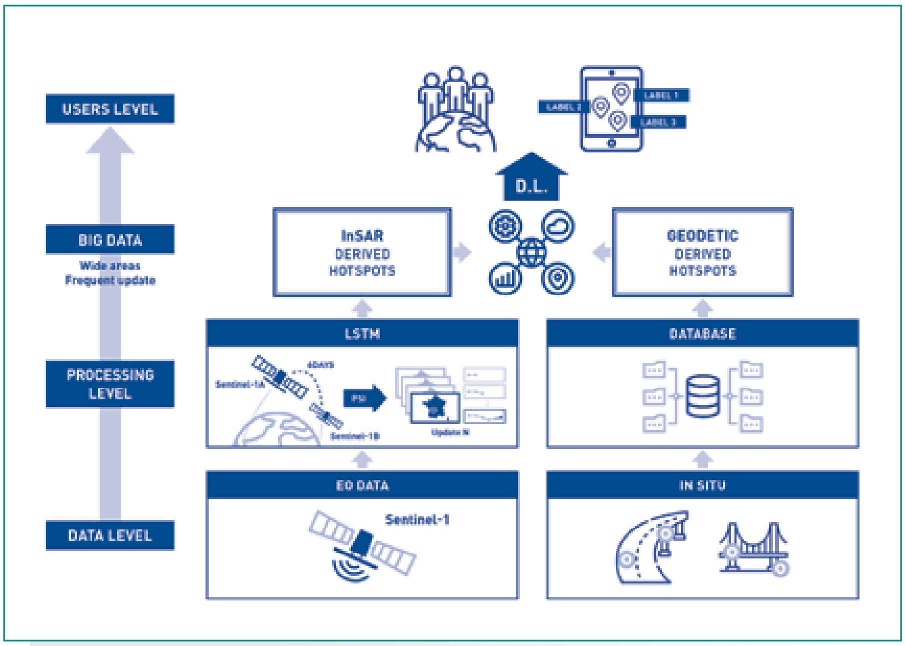

Figure 3: TRE Altamira's UC4b schematic.

the UCs together with their beneficiaries and ensure that the outcomes fit their needs and decision-making processes.
Drought and heat waves, due to their frequent occurrence in the last decade, are expected to become even more frequent in the future, as the corresponding persistent weather situations become more and more probable. UC1 puts on the table the following questions to respond, using AI on big EO data: *Which are the spatial factors that yield impact susceptibility versus resilience to meteorological drought and heat waves? Can we predict localized impacts given coarse scale meteorological information? Can we employ advanced ML solutions to model the spatio-temporal drought impact by combining this with a physical approach in a hybrid model? What are the anticipated long-term effects of drought and heat?*
UC2 aims to uncover how extreme weather conditions cause internal displacement and whether any other drivers are playing a role, related or not, to Climate Change. Through modern observational causal inference models, UC2 will be designed to understand, quantify, and predict migration flow effects from socio-economic contextual information as well as from environmental variables extracted from EO data.
Wildfires are a very impactful hazard affecting natural ecosystems and manmade infrastructures, causing in extreme cases human life losses. It is expected in the future that fire danger will increase, affecting even northern latitudes and evergreen tropical forests. The goal of UC3 is to model fire hazards in multiple temporal scales, providing short-term (e.g. hourly), mid-term (e.g. weekly, monthly) and long-term (e.g. seasonal) predictions of the best possible accuracy to interested users.
At any given moment, several volcanoes worldwide are erupting, while there are more than 1,500 volcanoes capable of reawakening and creating severe or even catastrophic impacts to society. Historical data from eruptions indicate that they are almost always preceded by volcanic unrest. Therefore, early warning based on detected volcanic unrest could be of great importance for civil protection authorities, enhancing their response effectiveness and allowing for scientists to deploy critical in-situ monitoring equipment to assess volcanic hazards more accurately. UC4a adopts a novel DL method, which has not been explored before in EO problems, to detect deformation associated with volcanic activity using SAR complex imagery.
Tourism is one of the pillars of the modern economy. It constitutes more than 10% of global GDP. The number of international tourists is forecasted to rise to 1,8 billion in 2030, making it crucial to find efficient ways to handle this growth, preserve the fragile destinations, and adapt to the increasing demand over limited hospitality infrastructures. Additionally, more than 65% of European travelers have declared that they are striving to make their travels more sustainable but do not find the right information or the possibility to assess their environmental footprint.
UC5's goal is to design a pricing engine for hotel rooms and tour packages purchases independent from the major reservation platforms and incorporate the environmental dimension for sustainable tourism.

UC4b- Deformation trend change detection on PSI time series for critical infrastructure monitoring

TRE Altamira (TREA) is a global leader in InSAR derived services and a pioneer in developing new products from satellite SAR data. Born as an independent company in 2000, TRE was funded by Politecnico of Milan to market the first PSI (Persistent Scatterer Interferometry) technique worldwide and in 2016, it merged with Altamira, an InSAR service company too. Since then, TREA employs advanced InSAR techniques to measure ground motion and structural movement from space. In more than 20 years of experience, TREA has analyzed over 5,000,000 km2 in the world and carried out 1,000+ projects in various market sectors such as civil engineering, oil & gas, mining, and natural hazards. Several critical infrastructure projects have been conducted all over the world to support engineers during the phases of planning, designing, and construction. Today, using its proprietary SqueeSAR® algorithm, TREA can provide surface displacement maps over large or small areas.
When asked to join the DeepCube project, TREA was thrilled to contribute to the challenge of exploiting the mass amount of Sentinel-1 SAR data combined with in-situ geodetic and other measurements with the final goal of creating a commercial service to monitor critical infrastructure at large scales.
The rationale behind:
The Copernicus Sentinel-1 (S1) SAR mission has turned out to be a game changer for the EO community. This twin platform constellation provides wide-scale, systematic (every 6 days) and free-access imagery over most of

the globe. Any single S1 image covers a 250x250 km2-area and is delivered within one hour from acquisition. This can be considered as a significant improvement over existing SAR systems.
Advanced InSAR technology is used to produce deformation maps from satellite SAR imagery with millimetric-precision. Thanks to S1 revisit time, deformation maps can be delivered to end-users on a regular basis, providing average velocity displacement rates of Persistent Scatter points and their associated displacement time series. Each deliverable consists of a layer of hundreds of thousands of measurement points and can be compared to previous layers to detect ground surface instability over large areas. The availability and accessibility to mass amounts of S1 data, combined with cloud-based solutions, enable TREA to deliver nationwide InSAR measurement databases providing millions of measurement points and their associated time series of displacement at each update.

How can we help end-users best understand and make use of this mass of information?

How can we contribute to the creation of a novel service that meets the monitoring requirements of a specific market sector?

Through the MATTCH project (ESA Open Call for Science, 2019), TREA has already implemented Deep Learning based methodologies to mass amount of S1 data to detect changes in time series trends for data screening purposes. So far, hotspots can be identified but no reason about driving mechanisms for these trend changes is given to end-users.
With DeepCube and the UC, TREA intends to further advance the state of the art reached in the MATTCH project.
TREA will develop new Deep Learning architectures for trend change detection from dense InSAR point time series- this time combined with industrial geodetic (GNSS) and other measurements to identify clusters of points sharing key attributes or features for critical infrastructure monitoring at a large scale. An Italian civil engineering group will contribute to the development of the UC by providing in-situ measurements over sample areas along a strategic infrastructure in Italy.

Conclusion

DeepCube addresses problems that require quantitative estimation of geophysical variables. For this purpose, a hybrid modeling approach for geophysical parameters estimation will be tested, where data driven modeling bound by physical models, AI applications, and Causality are combined.
DeepCube is exploiting non-space data, linking other information sources such as social media, industrial and socio-economic data to satellite data in order to create new value chains.
Also, DeepCube is tackling AI problems that use satellite InSAR data. Sentinel-1 SAR archive is the richest asset that remains hugely underexploited by the scientific community when it comes to AI-based applications, and DeepCube is addressing this gap.
The DeepCube platform is a unique legacy that will be developed as an open source suitable to be deployed in different cloud environments. To our knowledge, this will be the first time that such an end-to-end platform will become available, tailored for the processing of big Copernicus data and made available to the community to reuse and extend.

ACKNOWLEDGMENTS

Partners:
National Observatory of Athens, Greece (https://www.noa.gr/)
Max Planck Institute for Biogeochemistry, Germany (https://www.bgc-jena.mpg.de/index.php/Main/HomePage)
University of Valencia, Spain (https://www.uv.es/uvweb/college/en/university-valencia-1285845048380.html)
Logical Clocks, Sweden (https://www.logicalclocks.com/)
National and Kapodistrian University of Athens, Greece (https://en.uoa.gr/)
Gael Systems, France (https://www.gael-systems.com/)
Tre Altamira, Italy (https://site.tre-altamira.com/)
Infalia, Greece (http://infalia.gr/)
Murmuration, France (https://murmuration-sas.com/)

METAKEYS

Artificial Intelligence, Deep learning, Machine learning, Earth observation, Climate change

ABSTRACT

DeepCube is a 3-year Horizon 2020 project that leverages advances in the fields of Artificial Intelligence and Semantic Web to unlock the potential of big Copernicus data. Its goal is to address problems of high socio-environmental impact and enhance our understanding of Earth's processes correlated with Climate Change. To achieve this, the project employs mature ICT technologies, integrating them into a scalable, open and interoperable platform that provides solutions for all phases of an Earth Observation based AI pipeline. The DeepCube technologies will be demonstrated in five Use Cases.

AUTHOR

Chiara Gervasi, TRE Altamira
(chiara.gervasi@tre-altamira.com)
Alessia Ferrari, TRE Altamira
(alessia.ferrari@tre-altamira.com)
Ioannis Papoutsis,
National Observatory of Athens
(ipapoutsis@noa.gr)
Souzana Touloumtzi,
National Observatory of Athens
(stouloumtzi@noa.gr)