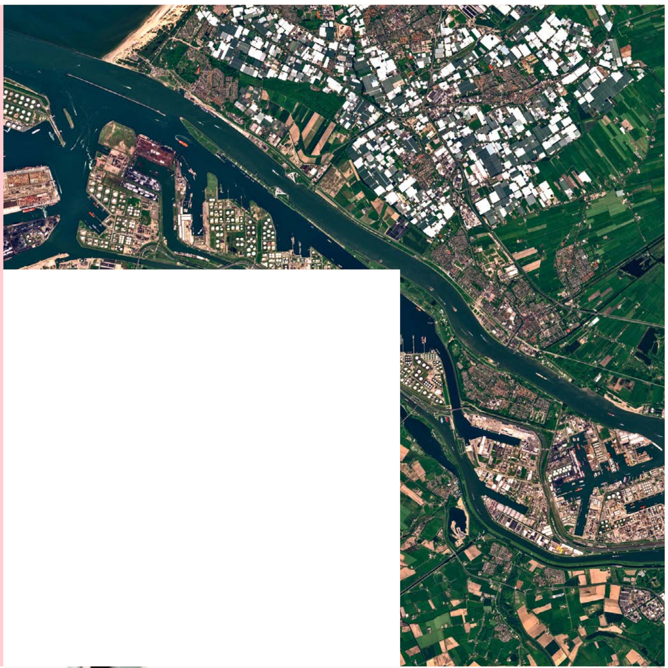

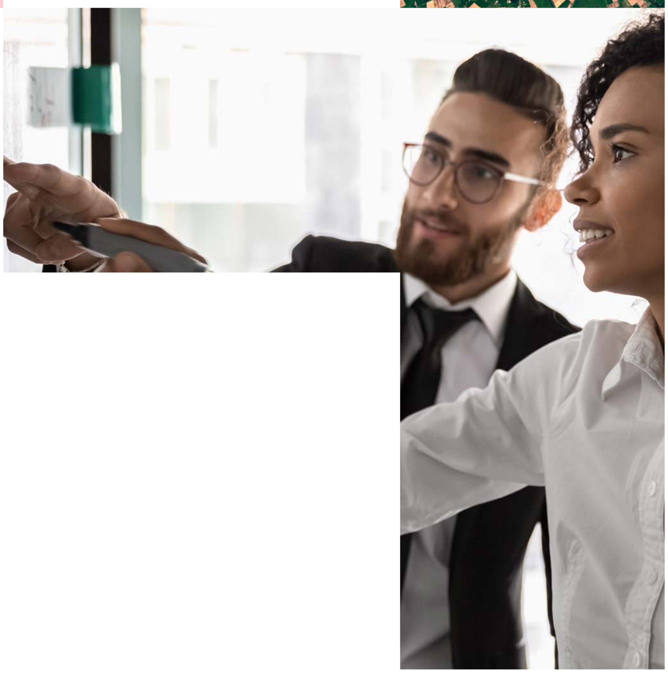

# EO4SD Lab

## How earth observation can support global sustainable development activities using cloud-based platforms

CGI

# Partnering for success

## EO4SD Lab

How earth observation can support global sustainable development activities using cloud-based platforms

Funded by

# Contents

# Earth Observation for Sustainable Development

**Since 2016, ESA's Earth Observation for Sustainable Development (EO4SD) initiative has aimed to facilitate a step change in the uptake of satellite-based environmental information in International Financial Institution's (IFIs) regional and global programs. CGI has partnered with GeoVille to develop EO4SD Lab, an earth observation (EO) processing and e-collaboration environment dedicated to Development Assistance (DA) to support this.**

With significant improvements in EO technology, the continuously growing volume and quality of data presents unique opportunities for Earth science and scientific communities whilst simultaneously posing a major challenge in reaching the full potential of exploiting the data. This is particularly prominent in extending this technology from purely scientific use to a level where it can be used as an operational source of environmental information in a range of non-EO specialist domains. For instance, organisations concerned with food security may require timely information on crop production statistics; urban planners developing new infrastructure requiring frequent monitoring of the local environmental impact; or natural resource management teams needing hydrological network mapping and water quality monitoring over wide-spread regions. A lack ICT resources and technical expertise will limit the extent to which these actors can access or produce value added EO information to meet these needs. This is in conjunction with the community challenges of the ever-growing diversity and complexity of development project needs and an increasing need for cooperation to make sense of the wealth of data. Cloud based services and big data visualisation offer the solution to support the discovery, access, and exploitation of this vital information.

The EO4SD Lab responds to the challenges presented using an Infrastructure as a Service (IaaS) approach to provide a scalable, low cost, e-collaboration framework to facilitate easy access to European EO mission data as well as immediate exploitation, analysis, sharing, mining and visualisation with high-level products. Critically, the platform presents an architecture designed for targeting specified user levels, lowering the entry threshold for all expertise levels and therefore supporting increased awareness and acceptance of EO within development assistance activities (e.g. project design, planning, implementation, policy-making, strategic planning, monitoring and safeguards). In addressing these factors the EO4SD Lab offers facilities to capitalise on the inherent benefits of EO data and contribute to building a long-term sustainable use of EO-derived information services to support development assistance activities.

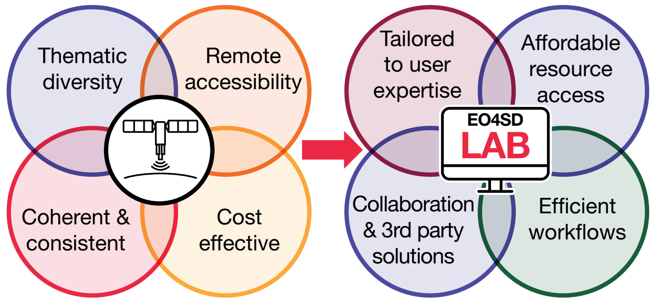


Figure 1: Key benefits of Earth Observation for sustainable development, made accessible by the key benefits of EO4SD Lab



# The challenge

Continuous evidence-based monitoring of international development progress is integral to meet the needs of a sustainable process and drives the need for readily usable information that can be integrated into policy making, planning & monitoring and evaluation procedures. Currently, a variety of data varying in complexity are implemented to support this task, given the wide variety of stakeholders, multifaceted domains and the distinct goals worldwide development activities entail. This strives for a greater level of coordination and efficiency between varied development aid user communities from governmental organisations in developing countries to international non-governmental (IFIs and NGOs) to address the given challenges and carry out economic development in an environmentally sustainable manner.

What is needed are independent, quantifiable data sources on development status, specifically in relation to development activities dependent on natural resources and affecting the local and/or global environment. This is where satellite EO technology has the potential to facilitate international development work in a globally consistent manner. Data produced is comprehensive, accurate, repeatable and timely, and has the flexibility to be tailored project specific requirements while maintaining a high level of reliability.

Despite this, there several limiting factors which have restricted the large-scale uptake of EO data in the international development sector. Often these limitations begin with the inherent complexity of the technology and a lack of expertise in the domain, either in exploitation of data or even at the level of not knowing what can currently be achieved with EO data, at what scales, with which costs and to what accuracy. To those with intermediate expertise, disparity between services and the non-localisation of data sources limit the efficiency to which data can be accessed, processed and produce meaningful outputs. With an unprecedented amount of EO data now available, the process of storing and analysing satellite data on local computing infrastructure has become increasingly limiting.

Development assistance needs call for both a generalised solution to monitoring and tracking project goals, and address the reality of implementing EO practices as such a solution:

- **N1:** Reduced complexity and greater uniformity in the acquisition methods of development tracking data spanning a wide variety of sustainable development projects.
- **N2:** Independent, quantifiable and reliable data sources relating to the tracking of sustainable development projects at all stages of the development lifecycle.
- **N3:** Varied and specialised entry points to EO data exploitation, tailored to user expertise levels.
- **N4:** Realistic and cost-effective access to technology and infrastructure for large scale EO adoption.
- **N5:** Streamlined and efficient workflows when acquiring, exploiting and sharing EO data.

These fundamental limitations on applying EO to sustainable development have called for the provision of on-demand cloud resourcing and systematic services and archives of EO data. By widening the access to the relevant technological capacities, users should be able to easily create their own individual data products by focusing on their particular location, area of interest and timelines.

# Benefits of earth observation data

In recent years, ESA's EO4SD initiative has been working towards mainstream EO-based products and services to be used as 'best-practice' environmental information in the working processes of international development. It is through a range of unique benefits that EO data has the potential to further support the full lifecycle of development aid activities, either by serving as direct monitoring indicators or helping augment and validate national statistics.

## Thematic diversity

Today's diversity of in orbit EO instruments provides global remote sensing across the electromagnetic spectrum. The correct choice of band or band combination allows for the extraction of a multitude of downstream products tailored to discern distinct information about a region of interest. This opens EO data up to an extensive list of applicable thematic areas used as indicators in sustainable development activities (N1). In the case of EO4SD initiative this has included agriculture and rural development, water resource management, urban development, climate resilience, disaster risk reduction, and supporting states affected by fragility. Furthermore, well defined services and products can provide many interdisciplinary applications, for instance a NDVI product may be utilised as an indicator for crop health, forestry degradation or urban green mapping. These products can easily be scaled to national, sub-national and regional levels depending on project specifications.

## Improved accessibility and a non-intrusive technology

Combined with data from long-term EO archives, the expansion of the European Commission's Copernicus program over the past 6 years has greatly enhanced the operational capability of global monitoring from space. Commercial operators are further complementing this with VHR commercial data. The acquisition of EO data is also not limited to in-orbit instruments, often acquiring ancillary data from drones or aircraft for greater resolution of localised areas. The extent of accessible EO data is therefore no longer a limiting factor in applying Earth Observation to sustainable development activities. The wealth of available data means (given the correct tools for access and processing) any region of interest is accessible for remote monitoring and the non-intrusive method provides an objective and independent data source. This vastly reduces the complexity of acquiring da ta from remote or difficult to access regions (N1). Sometimes this may be the only feasible method of collecting information from such regions (e.g. conflict areas), for instance in a recent EO4SD initiative case assessing the impact of armed conflict on the Syrian agricultural sector[i]. In many cases, the availability of collected data can be within days or even hours of acquisition, vital in cases of disaster response.

## Cost-effective

Satellite imagery may be the only cost-effective technology able to provide data on a global scale and can propagate to cost-savings within the entire DA implementation cycle. Although traditional alternatives should be used in combination with EO, the expense associated with bespoke surveying, ground teams and census data can be significant, especially when requiring repeat monitoring at regular frequencies and over vast regions of interest. The adoption of EO informed monitoring allows repeat large area measurement in a timely manner, for instance with Sentinel-2 having a global revisit period of just 5 days. Although VHR commercial data (<1m) comes at considerable expense, prices for commercial EO data is set to fall in the coming years[ii]. Regardless, with Copernicus being the world's largest environmental monitoring programme, the vast majority of use cases are adequately supplied with free and open access data (N4).


Figure 2: Thematic areas of ESA's EO4SD initiative

## Coherency and consistency

EO data provides a globally coherent and consistent source of information, both spatially and temporally through archived data. This allows for informed decision making and development of policies in the early stages of Design and Preparation, from analysis of data acquired preceding the initiation of a given development activity (N2). The long-term continuity of data allows the reliable cross-comparison needed for development tracking of considerable timescales, and similarly global monitoring facilitates comparisons of any country or region in the world regardless of the country's location, data reliability or GDP. The consistency further offers increased transparency, objectivity and accountability in the monitoring and evaluation stages, the remote sensing instruments have a well-defined and controlled range of error, making them less susceptible to biases that may be inherent in smaller scale local measurements (N2). This reliability is especially significant when combined with ground truthing data, leading to increased efficiency of existing operations and increasing the impact of delivery stages. Furthermore, any acquired EO data is processed and organised according to defined industry standards and to specified levels of radiometric/geometric correction and resampling, allowing users immediate analysis (N5).

# Exploitaion platforms in support of EO for sustainable development

## Introduction to EO4SD Lab

EO4SD Lab is a cloud based EO data exploitation platform (EP) sponsored by ESA to provide a coherent source of environmental information and help bridge the gap between the idealised solution of EO adoption and what is currently achievable by defined user groups. In tandem with the thematic diversity EO data can provide, the EP provides services to cover a range of thematic topics including forestry, agriculture, integrated water management, drought management, hazard mitigation and general supervision management (N1). The architecture of the Lab builds upon and unites existing technologies, data and related components, to provide the following key benefits:

- Efficient access to data products and services: Data and services in the same virtual environment allow for efficient workflows.
- Scalable infrastructure as a service: Users have access to computing resources they could not afford on their own.
- Tailored to user expertise: Defined entry points serve the needs of varied users from novices to EO experts.
- Collaboration: Awareness raising, information exchange, sharing and promotion about EO based products and services amongst the community.
- Unique third-party data and solutions: End users and value adders can add processing functions and data to the EP, to adapt with the changing needs of the user community.

CGI was well placed to develop an EP with these benefits, being able to draw on experience with previous ESA initiatives such as the thematic exploitation platforms (TEPs) for forestry (F-TEP) and food security (FS-TEP). While intended to serve specific themes, these platforms address many of the same challenges as the lab, and also provide access to a variety of user types. This offered a strong basis for the EO4SD lab architecture and its components with a key difference of needing to more clearly separate its user entry points.

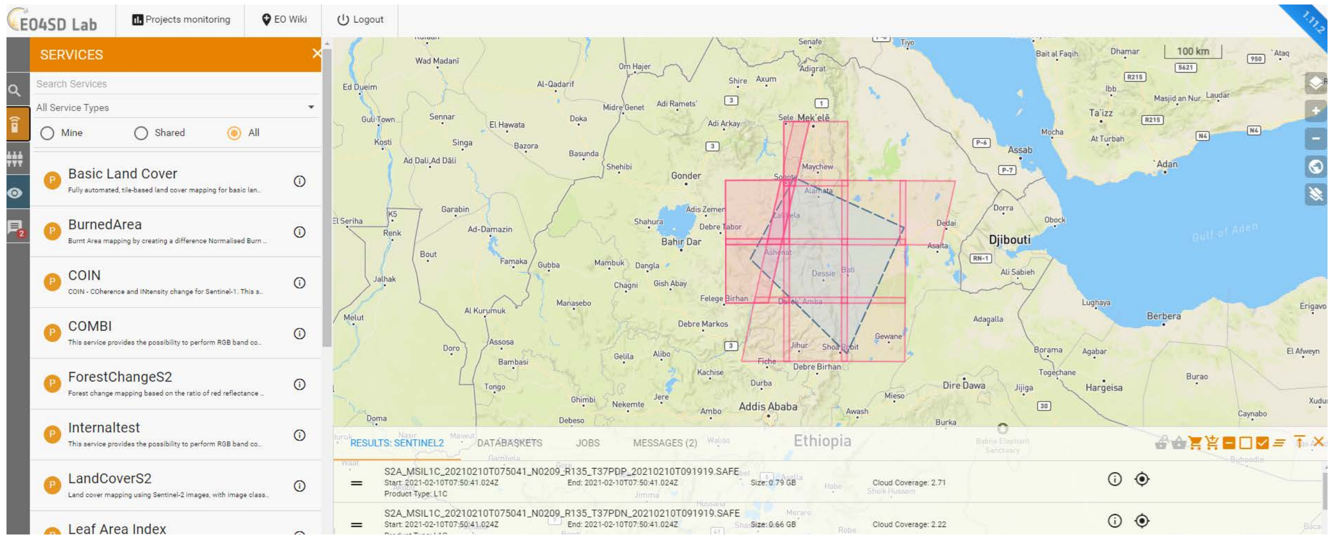


Figure 3: The EO4SD Lab platform – searching for Sentinel 2 data and relevant processing services



# How EO4SD Lab meets the needs of EO user groups

## Tailored to user expertise

A major barrier in the uptake of EO data in sustainable development communities is down to accessibility of the domain for non-expert users. Typically, the higher the users are in the decision line the less they are involved in the technical details, this can lead to a lack of consideration on the effectiveness of the available methods. The identified user groups across sustainable development stakeholders and organisations (i.e. the EO and R&D community, NGOs, IFIs, and national ministries & departments) and their associated needs are defined below.

Through 4 key functionalities (EO Wiki, GeoWizard, EO Basket and EO Processor), the EP architecture is critically designed to segment users according to their potential level of expertise in EO:

- **Management users** can obtain easy access to comprehensive factsheets, guidelines and showcases through the EO wiki, providing an easy entry point into the domain, and progress to selecting suggested satellite monitoring components and connect with costed solutions through the EO Basket.
- **Intermediate users** can access at higher levels through the GeoWizard or EO Basket for data discovery, external data harvesting and GIS functionalities for performing simple geospatial analysis and gathering thematic information in relevant displays for decision making.
- **Expert user** needs are still accommodated via the EO processor allowing elastic processing capacities, and set up of custom algorithm integration for specialised processing jobs.

This results in a unique platform where novice users can extract quantitative information for project planning support and implementation monitoring without a steep learning curve, while more advanced users still have the power and flexibility to extract the information they desire and develop their own processing chains (N3).

**Management type users**

- obtaining up-to-date EO images
- sharing datasets
- learning about EO potential
- guidance/tutorials on the interpretation of products
- publication
- communicating results
- e-collaboration to refine know how

**Intermediate users**

- consulting existing EO products
- performing simple geospatial analysis
- consulting information for decision making
- country/thematic intelligence gathering
- estimation of costs for technical assistance (including EO/GIS data/services)

**Expert users**

- processing EO data
- persistent access to information for monitoring
- performing complex GIS analysis

Figure 4: User group/expertise segmentation according to expected relative number of users (green-most; orange-medium; red-low)

# Collaboration & sharing

A strong foundation for collaboration and sharing not only means a richer source of information, data and products for users at all levels but facilitates effective workflows between users across the industry. The cloud-based architecture of EO4SD Lab integrates GeoServer for geospatial data sharing and allows for a strong focus on e-collaboration at all user levels, especially between user groups. Meaning information can flow between users of different expertise levels and reach the relevant user groups.

Expert users also have the ability to dockerise and deploy their algorithms to make bespoke services available to lower level users, opening up connections between relevant parties and keeping available services relevant to unique or specific community needs. Overall, this leads the interaction flow among the development assistance user community away from the current understanding of a highly inefficient *many-to-many* relationship. EO4SD Lab offers an extremely more efficient and reliable communication flow with all actors interacting and receiving information via the same platform, therefore allowing efficient interorganizational workflows on a large scale (N5).

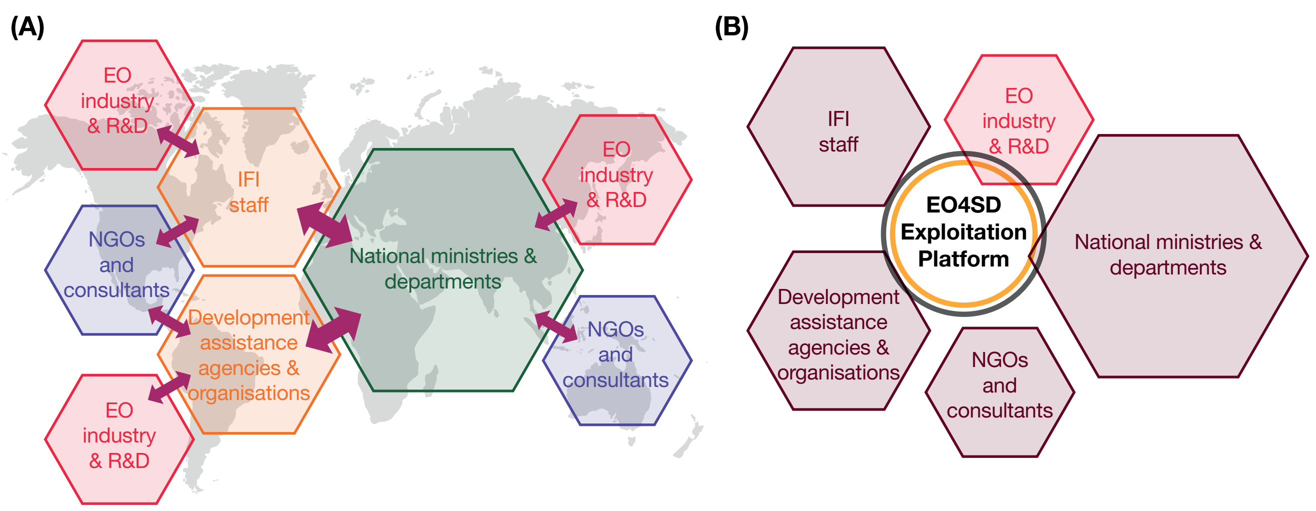


Figure 5: (A) Generic development assistance user community collaboration, interaction and information flow.
(B) EO4SD Lab e-Collaboration platform, sharing information across all sectors.

# Affordable access to data and processing

In the case of small or low-budget operations, it is the technological threshold required in the collection and processing of large data streams which prohibits the effective use of EO data. Previously, EO users were expected to download and analyse satellite data on their own computing infrastructure requiring suitable data storage and processing power. Due to the consistent increase in quantity and quality of satellite data, this is no longer a viable approach and becomes a significant issue in applying EO information in developing countries where IT infrastructure is more likely to be outdated and limited bandwidths may prevent effective access to large data streams. EO4SD Lab utilises scale cloud-based hardware such that the processing capability is collated in the same virtual environment as the EO data. Enabling open and simple-to-use access to users, through a web portal, this approach brings the users to the data, rather than the data to the users. This results in the significant advantage of the cloud-based approach; as long as the user has an internet connection they can access data/tools from any location without costly local hardware (N4). Furthermore, the on-demand scalability of the containerisation technology implemented allows for dynamic provisioning of processing capability only when it is needed, significantly reducing the ICT costs for end users.

In addition, the OSS nature of the platform provides low-cost access to tooling, extending the fundamental cost-effective nature of EO data acquisition down the processing chain (N4). Simultaneously, through the integration of these components, user efficiency becomes a key benefit of the Lab. With CREODIAS providing the majority of the EO data, users have fast direct access to over 20,000 TB of locally available data ready for immediate processing with provided processing applications. Alongside defined processing services and collaboration elements all within the same virtual environment, this is creating a systematic and streamlined individual user workflow, over traditional non-localised methods (N5).

# Architecture of EO4SD Lab

## Conceptual View

As described, the concept of the Lab is to be accessible to all users to allow information sharing while simultaneously providing different angles and perspectives through the 4 key functionalities of the Lab:

- **The EO-Wiki:** The wiki acts as the entry point to the lab and is concerned with what EO and geospatial information can do to support development assistance. It acts as a valuable resource in planning and design phases with showcases, factsheets and reports detailing assets that have been adopted in the past, how they have been used, and what benefits they have provided. In the interest of developing a community supported platform, users can further enrich the EO-Wiki by publishing their own additional material.

- **The EO GeoWizard:** The GeoWizard offers users without specific EO data processing expertise the statistical and analytical tools to produce thematic quantitative facts to answer project specific questions such as: what is the percentage of irrigated crops vs. non-irrigated crops in a specified area? What is the total area in $km^2$ of water bodies in a region? And has the aridity index increased in the last decade in this country? This information can be displayed in plots, tables, charts, histograms and time trends to make the most effective use of the results in decision support.

- **The EO Basket:** In the case where thematic layers are not addressed by the GeoWizard, the EO Basket allows for creation of required thematic layers to be used in subsequent analytics. A semantic-based search allows identification of suitable processing services by topic (e.g. forestry, agriculture, etc), geo-name or project phase (e.g. design, implementation, monitoring, etc). Through the primary Explorer interface, users can search for data, products and services relevant to their development activities and gain access to external service catalogues and portals.

- **The EO processor:** The EO processor offers a more technically advanced component, targeted at expert users and EO researchers to develop their processing chains or make them available to the community. The processor features a complete EO/GIS data analysis environment through standard tools such as QGIS, Jupyter Notebook and SNAP to support supervised data processing. Again, with EO4SD-Lab being a community supported platform, data produced using the EO Processor can be fed back to become an input source for the GeoWizard.

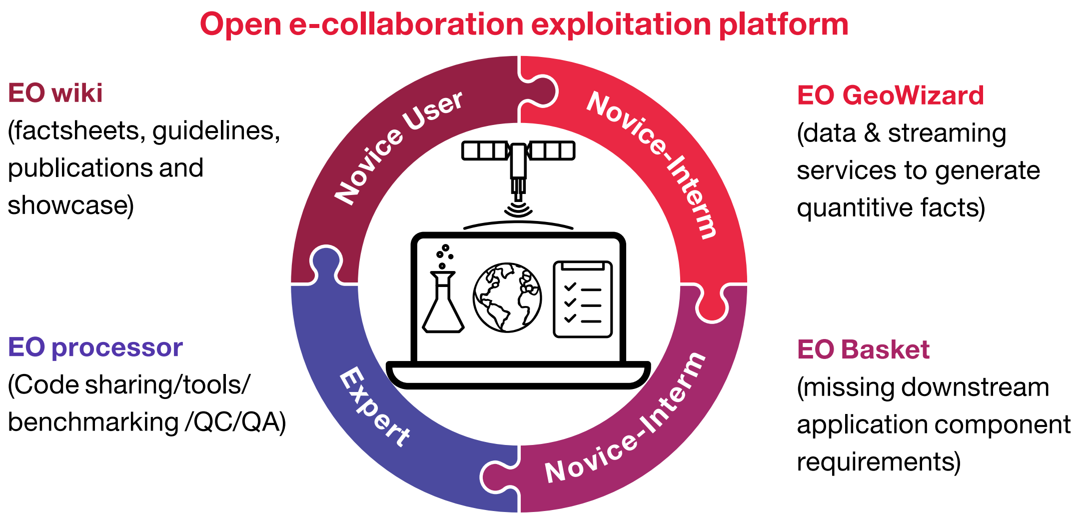


Figure 6: Conceptual view of EO4SD Lab.



## Technical View

Facilitating the 4 key functions, the EP is built through the integration of pre-existing open source software (OSS) tools, which are configured and expanded as necessary to meet the solution needs.

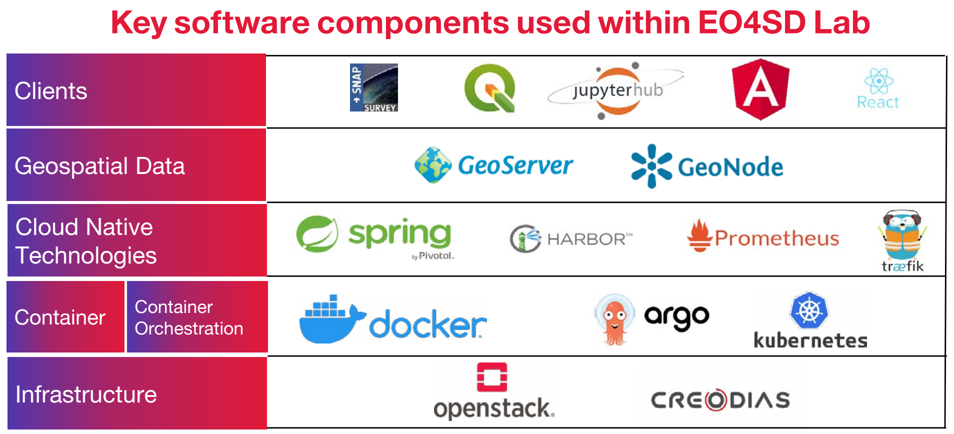


Figure 7: The OSS components integrated into the EOSD Lab technology stack.

The underlying ICT infrastructure, including the dynamic processing capabilities, is provided by the CREODIAS[iii] Infrastructure-as-a-Service (IaaS). CREODIAS is one of the Copernicus Data and Information Access Services (DIAS) platforms, which allow users to discover, manipulate, process and download Copernicus data and information. Further to this, built using applications such as OrfeoToolbox[iv] and ESA SNAP[v], EO4SD Lab provides a data processing service portfolio for the generation of a range of derived products, such as land cover, vegetation indices and water parameters, extending the suite of tools to be used within a single web portal. Through its developer tab, the lab also has the ability for users to upload their own scripts written in a wide range of languages, enabling them to create and share their own bespoke services.

The challenge of efficient and scalable processing of remotely sensed data is addressed through the adoption of containerisation technology. The use of container-technology is particularly relevant as it allows distinct processing activities to be instigated and executed independent of other activities. This both reduces the limitations in terms of processing activities and also ensure resources are used efficiently. Within EO4SD Lab, OSS tools such as Argo and Kubernetes are used for container orchestration with containers themselves being created using Docker. With the developer tab facilitating the dockerisation of user developed bespoke services, developers can deploy them on the platform providing the wider user community with ease of access to the rich set of services produced by varied collaborators without the concern of dependencies or set-up.

The visualisation of geospatial data is enabled through two key OSS tools. GeoServer[vi] is a commonly used open-source server that allows users to share, process and edit geospatial data. It publishes data from any major spatial data source using open standards and is the key tool for the displaying geospatial data through the GeoBrowser. This works in synergy with GeoNode[vii], which is a web-based application and platform for developing geospatial information systems (GIS) and for deploying spatial data infrastructures (SDI). Within EO4SD Lab, GeoNode, through the EO Wiki interface, provides users with searchable access to the various geospatial data layers generated and the ability to combine these layers into maps for further analysis and sharing. Additionally, the lab provides users with access to traditionally native desktop software packages such as SNAP and QGIS[viii] on the cloud for further data analysis, visualisation and processing.

# What now?

Achieving sustainable development in the international development sector, in many cases, lacks the continuous evidence-based monitoring that is needed for effective decision support at all steps in project life-cycles. A need has arisen for easily accessible, independent, and reliable environmental data sources with the flexibility to be applied across domains, reducing complexity of project monitoring in the wider development community. Where the idealised solution of EO adoption addresses many of these user needs with objective, consistent and cost-effective data, there remain barriers (ICT resources, lack of expertise, workflow efficiency) which limit the user community from fully exploiting the wealth of available resources. EO4SD Lab addresses these issues through its cloud-based architecture and its distinct segmentation of user groups, reducing the knowledge and technological thresholds required and bridging the gap between the capabilities of novices and experts. Large datasets and a suite of OSS tools amalgamated in a virtual cloud-based environment are streamlining the user processing workflow, and a strong focus on data and service sharing provides the relevant connections in the wider community and drives further expansion of the platform. While the capability gap has been significantly reduced, issues still remain in being able to effectively provide the relevant quantitative information to decision makers and managers without any technical expertise (or without even needing to know it was created with EO data). This therefore remains a significant target in the evolution of the platform. With this, the platform continues to develop by addressing real world end-user priorities through the selection of service pilot projects, in collaboration with significant development community actors such as the World Bank, to further refine and serve the user needs.



# References:


i https://www.eo4idi.eu/sites/default/files/publications/eo4sd_syria_final.pdf

ii https://eo4society.esa.int/wp-content/uploads/2020/07/Caribou-Space_ESA-EO-for-Agenda-2030-v2.pdf

iii https://creodias.eu/

iv https://www.orfeo-toolbox.org/

v https://step.esa.int/main/toolboxes/snap/

vi http://geoserver.org/

vii http://geonode.org/

viii https://www.qgis.org/en/site/


# About CGI

Founded in 1976, CGI is among the largest IT and business consulting services firms in the world.

We are insights-driven and outcomes-based to help accelerate returns on your investments. Across 17 industries in 400 locations worldwide, our 76,000 professionals provide comprehensive, scalable and sustainable IT and business consulting services that are informed globally and delivered locally.

Our commitment: Insights you can act on.

www.cgi.com/uk/defence