DIGITAL EARTH OBSERVATION INFRASTRUCTURES AND INITIATIVES: A REVIEW FRAMEWORK BASED ON OPEN PRINCIPLES

M. Di Leo[1]*, M. Minghini[2], A. Kona[2], N. Spadaro[3], A. Kotsev[2], J. Dusart[4], S.Lumnitz[5], C. M. Ilie[6], C. E. Kilsedar[1], A. Tzotsos[7]

[1] Arcadia SIT, under contract with the European Commission, JRC, Ispra, Italy - (margherita.di-leo, candan.kilsedar)@ext.ec.europa.eu

[2] European Commission, Joint Research Centre (JRC), Ispra, Italy - (marco.minghini, albana.kona, alexander.kotsev)@ec.europa.eu

[3] Fincons S.p.A., under contract with the European Commission, JRC, Ispra, Italy - nicholas.spadaro@ext.ec.europa.eu

[4] European Commission, DG for Research and Innovation (RTD), Brussels, Belgium - jean.dusart@ec.europa.eu

[5] European Space Agency (ESA-ESRIN), Frascati, Italy - stefanie.lumnitz@esa.int

[6] Terrasigna, Bucharest, Romania - codrina.ilie@terrasigna.com

[7] Open Source Geospatial Foundation (OSGeo), Athens, Greece - tzotsos@osgeo.org

KEY WORDS: Digital Infrastructure, Earth Observation, Interoperability, Openness, Research Infrastructure, Review

ABSTRACT:

Recent years have seen a tremendous increase of digital Earth Observation (EO) infrastructures, which provide web-based environments for accessing and processing data in a highly automated and scalable way. However, the current landscape of EO infrastructures and initiatives is fragmented, with various levels of user on-boarding and uptake success. The current work aims to make sense of this complex landscape by providing two main contributions. First, it offers a classification scheme used to review and analyse more than 150 EO infrastructures and initiatives. Then, adopting a user-centric perspective, the main limitations and obstacles currently faced by users when working with the existing EO platforms are identified. For each of these limitations, we propose a number of good practices that could benefit, from a user point of view, the design and functioning of EO platforms. Some technological enablers, i.e. specific resources (such as software components, standards and data encodings) that emerged from the analysis as holding a great potential for improving the usability of existing EO platforms, are finally listed. The work aims to provide a first scientific insight on how to best design and operate EO platforms to maximise the benefits of their user communities.

1. INTRODUCTION

In recent years, the democratisation of access to Earth Observation (EO) data, in parallel to the increased volume and variety of such data, have led to the paradigm shift towards "bringing the user to the data" (European Space Agency, 2016). Such a democratised access is exemplified by the European Union's (EU) Copernicus Programme, which on a daily basis makes available terabytes of high-quality, openly-licensed EO data suitable for several research and commercial applications (Harris and Baumann, 2015). EO data, encompassing both remotely-sensed satellite data and in-situ data, represent a key data source for the new generation of spatial data infrastructures (Kotsev et al., 2021). The computational power required to work with these large amounts of data, the need for large storage volumes as well as the ease of data access and fast distribution of results were met with the rise of cloud-based digital infrastructures and services. These provide environments that can be readily instantiated and equipped with the necessary data and processing tools accessible through a common user interface, in a highly automated and scalable manner, to leverage EO data proximity. Several such infrastructures as well as other initiatives (the latter also including services and components that offer specific capabilities) have been developed, either as a byproduct of single companies leveraging enormous hyperscale computing power (such as Google Earth Engine, Microsoft Planetary Computer and Earth on Amazon Web Services) or as projects funded and operated by international consortia that are primarily driven by specific policy objectives.

Among the latter, a key international initiative is the Global Earth Observation System of Systems (GEOSS), launched by the Group on Earth Observations (GEO, `https://www.earthobservations.org`) to integrate existing EO platforms and infrastructures to strengthen environmental data sharing and improve the monitoring of the state of the Earth. Europe is delivering its regional contribution to GEO with the launch of the EuroGEO initiative and its corresponding digital European infrastructure contributing to GEOSS. It has the aim to connect existing European EO assets, including data, sensor networks, analytical methods and models, computing infrastructures, products and services that support European policy objectives. As part of this initiative, the European Commission's Joint Research Centre (JRC) is developing a prototype of such a European infrastructure for testing and monitoring the interoperability of existing EO assets. These derive from the substantial investments made by the EU — through its research and innovation funding programmes (Horizon 2020 and Horizon Europe) in the development of EO prototypes, architectures, demonstration products and services. Popular examples are the Copernicus Data Space Ecosystem, which merges the four existing Data and Information Access Services (DIASes) Sobloo, Mundi, ONDA and CreoDIAS, and the ESA's Thematic Exploitation Platforms (TEPs) (Gomes et al., 2020).

The work presented in this paper originates within this context and is mainly driven by the acknowledgement of the high level of fragmentation of the current landscape of digital infrastructures and initiatives for accessing and processing EO data, both at the European and global level, which feature varying levels of user on-boarding and uptake success — see e.g. Wagemann et

---

* Corresponding author

al. (2021). To make sense of such a complex and ever-changing landscape, after this introduction Section 2 proposes a classification scheme for digital EO infrastructures and initiatives as well as a review framework, based on a set of open principles, identifying the most critical limitations faced by users when working with the existing EO platforms. This is followed by Section 3, where the classification of existing digital EO infrastructures and initiatives is presented together with the identification of a number of good practices to address the identified limitations. Some particularly good resources, referred to as technological enablers, which emerged from the review are also presented. Section 4 closes the paper with a discussion of the main findings and some potential future research directions.

2. METHODOLOGY

We developed a classification scheme for digital EO infrastructures and initiatives according to the services they offer. The classification scheme entails the following categories:

- Data providers: they make EO datasets available within infrastructures;

- Cloud-based geoprocessing platforms: computational capacity may also be offered by data providers in line with the paradigm "bring the user to the data";

- Brokers and catalogues: they offer discovery services by harvesting data from existing catalogues;

- Thematic hubs and Research infrastructures: they incorporate EO data relevant to specific thematic domains, such as agriculture, biodiversity and atmosphere;

- Data cubes: they implement a multidimensional array structure, on which one can load several bands and perform slicing and algebraic operations;

- Virtual infrastructures: they place additional layers on top of existing platforms with the goal to facilitate data access and increase the discoverability of and interoperability between such platforms;

- Initiatives and programmes: EO-related, publicly funded programmes.

The way to classify existing EO infrastructures and initiatives according to this set of categories is based on their most prominent characteristics, described on the related web pages. The authors acknowledge the difficulty arising when attempting to map the current landscape of EO infrastructures, due to the overlap of segments that form it. For example a DIAS can be considered a standalone cloud provider, but also a functional part of a Data cube or a Thematic Exploitation Platform. Many other examples of potential overlaps could be similarly identified. In this complex landscape, the categories established for the purpose of this paper can be framed according to a hierarchy, depicted in Figure 1. Based on this hierarchy, Data providers (such as Copernicus, NASA, JAXA, Airbus and Planet) and Cloud-based platforms (such as CreoDIAS, Google Earth Engine and Earth on Amazon Web Services) are the two main segments, enabling (through technology) the evolution of all the other segments of the EO infrastructural landscape to best serve the needs of the EO community. In Figure 1, such segments are depicted in green, cyan and red depending on the aspect they mostly pertain to (data, infrastructure and technology, respectively). Initiatives and programmes do not fall into any of these aspects and are therefore considered as an overarching category.
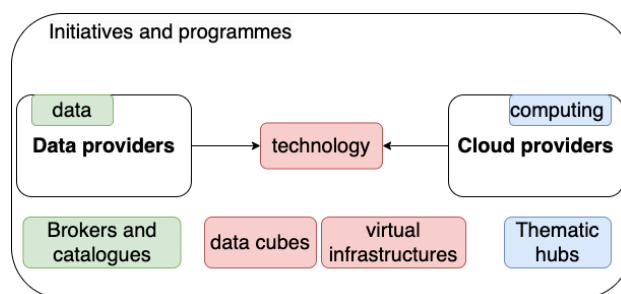
Figure 1. Hierarchy of the categories of EO infrastructures and initiatives used in this paper.

2.1 Identification of user needs

In a parallel work (Di Leo et al., in press), we analysed user needs and mapped them to the current offer of EO platforms, with the aim to identify overlaps and gaps in the existing ecosystem to steer future developments (with a special focus on the European infrastructure contributing to GEOSS). Users of digital EO platforms (mainly scientists, professionals and decision-makers) have a wide range of needs associated with the steps of creating, starting from raw data, EO products and services that provide actionable insights. While the available EO platforms may target only specific parts of the full lifecycle of service/product development, in our work we adopted a hands-on approach to identify the development steps of this lifecycle that are harder to complete using the available infrastructures.

In addition, we analysed — through both a review of available documentation and dedicated interviews — the use cases and pilots from the "EuroGEO Showcases: Applications Powered by Europe" (e-shape) project (Ranchin et al., 2021), a flagship European project showcasing the European contribution to GEO, in order to identify which EO digital platforms were successfully used in practice and whether currently available services cover the full lifecycle of the project needs (Di Leo et al., 2022; Voidrot-Martinez, 2022).

2.2 Review framework

From the process of analysing user needs described above, we identified the following limitations (as seen from a user perspective) in the uptake of (some of) the existing platforms:

- Fragmentation, leading to discoverability and interoperability issues: multiple EO platforms have been implemented over time, sometimes to serve the very same communities (i.e. having the same scope and field of application), with the result that it is sometimes hard for users to discover and compare services offered by different platforms, which often are also not interoperable with each other.

- Steep learning curve: the amount of time users need to become familiar with digital EO platforms is often underestimated and may be very large.

- Difficulty to understand what the services offered are and whether they fit the user needs (e.g. sometimes access is granted behind a paywall or is subject to filling a form); in addition, pricing is often not transparent and users find it difficult to compare the offers from different providers.

- Processing workflows, especially within EO research environments, are often not customisable and open for the users to modify and adapt to their own needs.

- Vendor lock-in: once users start developing on a certain EO platform, moving their code to another platform may be a lengthy process.

- EO platforms may not facilitate code sharing and reuse.

- Lack of assurance about the sustainability of the EO platforms, especially those that are project-based (these are usually developed and implemented by consortia of private companies) after the initial project funding has concluded.

- Internal policies prohibiting the publication of commercial added-value code/algorithms, with regard to data processing and knowledge extraction, to external EO platforms, unless there are very robust security policies in place.

Based on the identified limitations, we built a user-centred review framework where we suggest potential approaches and solutions on how these can be better addressed.

3. RESULTS

3.1 Classification of EO infrastructures and initiatives

In this section, we present the preliminary results of the classification performed using the scheme presented in Section 2, which was applied to more than 150 existing digital EO infrastructures and initiatives. This review is not meant to be exhaustive, since additional platforms exist in the EO landscape and new ones will continue to appear in the future. Also, we did not have the chance to perform a hands-on experimentation of all the platforms listed in the following, which would require a consistent investment of time and resources. Nevertheless, we believe that our classification scheme, together with the preliminary analysis of the existing platforms (extending previous analyses restricted to small subsets only, e.g. the one by Gomes et al., 2020), can: i) provide a significant benefit to the EO user community in the identification of gaps and synergies; and ii) usefully inform platform providers on how they may improve existing services and drive the development of future ones. As mentioned in Section 2, the boundaries between the identified categories are sometimes blurred and the same platforms may in principle belong to more than one category; in these cases, the category selected in the classification was the one corresponding to the main function or feature of the platform, as described on its web page. For each identified category, the following subsections illustrate the main highlights of the classification.

3.1.1 Data providers The platforms that, as their most prominent service, offer discovery and accessibility of EO data, are classified as Data providers. Table 1 lists providers of several kinds of EO data that are relevant to multiple categories, such as: i) the collection type: in-situ, modeled, remotely-sensed; ii) the societal/policy domain: hydrology, agriculture, sea, etc.; iii) the relevance to specific geographical regions, such as the Arctic or the Antarctic; and iv) the level of data processing (L2, L3, L4). Among these, the Copernicus services are publicly-funded programs providing open access to datasets and products addressing the domains of atmosphere, marine environment, land, climate change, security, emergency and ground motion. The Copernicus Open Access Hub is perhaps the most well-known provider of Sentinel data. An important tool that we used to discover data providers is the Yellow Pages of the GEOSS Portal (`https://www.geoportal.org/yellow-pages`).

Name	URL
Arctic and Antarctic Data Archive System	`https://ads.nipr.ac.jp`
AfroMaison	`http://afromaison.net`
Aspen Global Change Institute	`https://www.agci.org`
Abu Dhabi Global Environmental Data Initiative	`https://agedi.org`
Moldovan Agency for Land Relations and Cadastre	`https://www.geoportalinds.gov.md/geonetwork`
Argo profiling floats global network	`https://argo.ucsd.edu`
AtlantOS	`https://www.atlantos-h2020.eu`
Conservation of Arctic Flora and Fauna (CAFF)	`https://www.caff.is`
Copernicus Atmosphere Monitoring Service	`https://atmosphere.copernicus.eu`
Copernicus Global Land Service	`https://land.copernicus.eu`
Copernicus Marine Service	`https://marine.copernicus.eu`
Copernicus Climate Change Service (C3S)	`https://climate.copernicus.eu`
Copernicus service for Security applications	`https://www.copernicus.eu/en/copernicus-services/security`
Copernicus Emergency Management Service	`https://emergency.copernicus.eu`
Copernicus Open Access Hub	`https://scihub.copernicus.eu`
EUMETCast	`https://www.eumetsat.int/eumetcast`
MAXAR Analysis Ready Data (ARD)	`https://www.maxar.com/products/analysis-ready-data`
Planet	`https://www.planet.com`
Earth Online portal	`https://earth.esa.int`
Validation Data Centre	`https://evdc.esa.int`
EarthData	`https://www.earthdata.nasa.gov`
Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici	`https://www.cmcc.it`
CEOS WGISS Integrated Catalog (CWIC)	`https://ceos.org/ourwork/workinggroups/wgiss/access/cwic`
Data Center - Climate Change Centre Austria	`https://data.ccca.ac.at`
German Aerospace Center	`https://www.dlr.de/EN/Home/home_node.html`
ArcGIS Living Atlas of the World	`https://livingatlas.arcgis.com/en/home`
Groundwater Assessment Platform	`https://www.gapmaps.info`
Global Biodiversity Information Facility	`https://www.gbif.org`
Global Mercury Observation System	`https://www.gmos.eu`
European Commission - Joint Research Centre	`https://data.jrc.ec.europa.eu`
Indian Space Research Organisation	`https://www.isro.gov.in`
OneGeology	`https://onegeology.org`

Japan Agency for Marine-Earth Science and Technology	`https://www.jamstec.go.jp/e`
National Land Survey of Finland (Etusivu)	`https://www.maanmittauslaitos.fi/en`
Sustainable Caucasus	`https://sustainable-caucasus.unepgrid.ch`
OpenStreetMap (OSM)	`https://www.openstreetmap.org`
EUBUCCO	`https://eubucco.com`
Microsoft Building Footprints	`https://github.com/microsoft/GlobalMLBuildingFootprints`
Plan4all	`https://www.plan4all.eu`
Protected Planet	`https://www.protectedplanet.net/en`
Map of life (MOL)	`https://mol.org`
OpenGeoHub	`https://opengeohub.org`
Africa Knowledge Platform	`https://africa-knowledge-platform.ec.europa.eu`
Natural Earth	`https://www.naturalearthdata.com`
Digital Earth Australia	`https://www.dea.ga.gov.au`
Digital Earth Africa	`https://www.digitalearthafrica.org`

Table 1. Data providers.

Name	URL
CreoDIAS	`https://creodias.eu`
WEkEO	`https://www.wekeo.eu`
OpenEO Platform	`https://openeo.cloud`
Actinia Mundialis	`https://actinia.mundialis.de`
Google Earth Engine	`https://earthengine.google.com`
Microsoft Planetary Computer	`https://planetarycomputer.microsoft.com`
Earth on Amazon Web Services	`https://aws.amazon.com/earth`
European Open Science Cloud (EOSC Hub)	`https://www.eosc-hub.eu`
European Weather Cloud	`https://www.europeanweather.cloud`
Earth Observation Data Centre (EODC)	`https://eodc.eu`
Terradue	`https://www.terradue.com`
SEPAL	`https://sepal.io`
TerraScope	`https://terrascope.be/en`
EOX	`https://eox.at`
ADAM	`https://adamplatform.eu`

Table 2. Cloud-based geoprocessing platforms.

3.1.2 Cloud-based geoprocessing platforms This category includes data providers offering computational capabilities such as cloud computing, High Performance Computing (HPC) facilities, GPUs, etc. (see Table 2). Without having to download the data, these platforms allow users to process it in the cloud and then download the results. This category includes, among others, commercial hyperscalers such as Google Earth Engine, Microsoft Planetary Computer and Earth on Amazon Web Services. Out of the 5 initially established DIASes, CreoDIAS, Mundi, Sobloo and ONDA are currently evolving into a new platform called the Copernicus Data Space Ecosystem, while WEkEO remains a standalone platform. The OpenEO Platform is a continuation of the EU-funded "openEO" Horizon 2020 project, whose goal was the development of a three-layered Application Programming Interface (API) that would allow users to consistently and seamlessly find and access EO data from different providers as well as processing it using R, Python and JavaScript. The OpenEO API is an important example of a technological enabler, which helps overcome vendor lock-in and allows seamless migration from one endpoint to another.

3.1.3 Brokers and catalogues Data catalogues, which are listed in Table 3, are brokering metadata from existing data providers, allowing users to browse and discover data from one single entry point. As an exception, the GEO Knowledge Hub offers a catalogue of "knowledge" (as opposed to just datasets) that also includes software, documents and other products.

3.1.4 Thematic hubs and Research infrastructures Research infrastructures (see Table 4) are "facilities that provide resources and services for the research communities to conduct research and foster innovation in their fields" (European Commission, 2023). The ESA Thematic Exploitation Platforms offer environments tailored to specific areas with easily accessible data and specialised, ready-to-use software. Other platforms like Gaiasense, which offers access to data, infrastructure and services related to agriculture in Greece, are tailored to specific

regions and not only to thematic areas. The European Plate Observing System (EPOS) and the GRID-Geneva are wider in scope and offer infrastructure and services on different thematic areas, with GRID-Geneva being part of the Early Warning and Assessment Division of the United Nations Environment Programme's global group of environmental information centres.

3.1.5 Data cubes Data cubes (see Table 5) are platforms offering Analysis Ready Data (ARD) and/or access to other types of data, which can be processed in a multi-dimensional array (including the spatial components, the bands and the time component). They can be focused on a certain domain or on a specific geographical region. Data cubes can also refer to software components integrated as tools in other platforms, which generally offer access to a broader variety of datasets.

3.1.6 Virtual infrastructures Many Virtual infrastructures are developed, with the aim of interconnecting services provided by existing infrastructures. Some examples are listed in Table 6. Gaia-X is a European initiative focused on creating a federated data infrastructure that enables secure and trustworthy data exchange between different organisations and systems. An implementation of Gaia-X federation services is available at `https://gitlab.com/gaia-x/data-infrastructure-federation-services`. The International Data Space Association (IDSA) provides a set of technical specifications and guidelines for creating secure data exchange ecosystems based on the principles of data sovereignty, data privacy, and data security. The Eclipse Dataspace Connector proposal (`https://projects.eclipse.org/proposals/eclipse-dataspace-connector`) provides a connector framework for sovereign, interorganisational data exchange based on Gaia-X and IDSA specifications. The Earth Observation Exploitation Platform Common Architecture (EOEPCA) establishes a consensus of best practices for EO Exploitation Platforms based on open standards, and is developing a reference implementation of building blocks as free and open source software. The GEOSS Platform Plus, currently under development, aims at facilitating access to tailor-made information and actionable knowledge through a

Name	URL
GEOSS Portal	`https://www.geoportal.org`
NextGEOSS	`https://catalogue.nextgeoss.eu`
Geo-Cradle	`http://geocradle.eu/en`
Centre of EO Research & Satellite Remote Sensing (BEYOND)	`http://beyond-eocenter.eu`
GEO Knowledge Hub	`https://gkhub.earthobservations.org`
European Environment Agency (EEA) geospatial data catalogue	`https://sdi.eea.europa.eu/catalogue/srv/eng/catalog.search#/home https://sdi.eea.europa.eu/catalogue/doc/api/index.html https://github.com/geonetwork/core-geonetwork`
INSPIRE Geoportal	`https://inspire-geoportal.ec.europa.eu`
The official portal for European data	`https://data.europa.eu/en`
STACIndex	`https://stacindex.org`
Ecopotential In-Situ Data Catalogue	`https://www.ecopotential-project.eu`
Federated EO Missions Support Environment	`https://fedeo-client.ceos.org/about`
IAEA Environment Laboratories	`https://www.iaea.org`
International Centre for Integrated Mountain Development	`https://www.icimod.org`
CEOS International Directory Network	`https://idn.ceos.org`
World Soil Information (ISRIC)	`https://www.isric.org`
The World Bank Data Catalog	`https://datacatalog.worldbank.org/home`
Planerary Data Access (PANDA)	`https://panda.copernicus.eu`
UN Spider Knowledge Portal	`https://www.un-spider.org`
Climate-ADAPT indicators	`https://climate-adapt.eea.europa.eu/en/knowledge/c-a-indicators`
European Environment Agency (EEA) indicators	`https://www.eea.europa.eu/ims`

Table 3. Brokers and catalogues.

Name	URL
SEPAL	`https://sepal.io`
eLTER	`https://elter-ri.eu`
TWIGA	`https://website.twiga-h2020.eu`
AgriStack	`https://pib.gov.in/PressReleasePage.aspx?PRID=1883173`
Tembo Africa	`https://www.developmentaid.org/organizations/awards/view/410505/transformative-environmental-monitoring-to-boost-observations-in-africa-tembo-africa`
GBIF	`https://www.gbif.org`
Life Watch ERIC	`https://www.lifewatch.eu`
ANAEE	`https://www.anaee.eu`
ESA Thematic Exploitation Platforms (TEPs)	`https://www.anaee.eu — https://urban-tep.eu/#! — https://www.coastal-tep.eu/portal — https://f-tep.com — https://geohazards-tep.eu/#! — https://portal.polartep.io/ssoportal/pages/login.jsf?faces-redirect=true — https://foodsecurity-tep.net`
European Plate Observing System (EPOS)	`https://www.epos-eu.org`
Eionet	`https://www.eionet.europa.eu`
Gaiasense	`https://www.gaiasense.gr/en/gaiasense`
AGdata-hub	`https://agdatahub.eu/en`
Actris	`https://www.actris.eu`
CUAHSI	`https://www.cuahsi.org`
Swiss Federal Institute for Forest, Snow and Landscape Research WSL	`https://www.wsl.ch/en/index.html`
EOMAP	`https://www.eomap.com`
GRID-Geneva	`https://unepgrid.ch/en`
GROW Observatory	`https://growobservatory.org`
HydroShare	`https://help.hydroshare.org`
International Institute for Applied Systems Analysis	`http://www.iiasa.ac.at`
International Long Term Ecological Network	`https://www.ilter.network`
Monitoring of the Environment for Security in Africa	`https://www.rcmrd.org/en/about-servir/40-past-projects/25-mesa`
National Atmospheric Research Laboratory	`https://www.narl.gov.in`
Recetox	`https://www.recetox.muni.cz/en`
Swedish Meteorological and Hydrological institute	`https://www.smhi.se/en`
Open Foris	`https://openforis.org`
TAHMO	`https://tahmo.org`
Biopama Ris	`https://rris.biopama.org`

layer of services for collecting, organizing, and distributing EO data, including in-situ data and models. Finally, this category also includes the Copernicus Data Space Ecosystem (CDSE), already mentioned in Sub-section 3.1.2. This open ecosystem, expected to be fully deployed by mid-2023, is formed by two components: i) a public service data portal that provides access to a wide range of data and services from the Copernicus Sentinel missions; and ii) a commercial service that offers payable services such as computational services, infrastructure as a service, etc.

3.1.7 Other Initiatives and programmes Table 7 groups other Initiatives and programmes that are connected with the above-mentioned infrastructures, but do not fall into any of the previous categories. These include, among others, the aforementioned EU Copernicus Programme and the e-shape project.

Name	URL
Vision on technology for a better world — VITO	https://vito.be
Hydrology (IHP)	https://en.unesco.org/themes/water-security/hydrology
Webservice-Energy	http://www.webservice-energy.org
World Resources Institute	https://www.wri.org

Table 4. Thematic hubs and Research infrastructures.

Name	URL
FAIRiCube	https://fairicube.nilu.no
ENVISION Data Cube	https://envision-h2020.eu
Eco Data Cube	https://ecodatacube.eu
AD4GD	http://www.ad4gd.eu
B3 (B cubed)	https://cordis.europa.eu/project/id/101059592
Euro Data Cube	https://eurodatacube.com
Harmonia	https://harmonia-project.eu
Open Data Cube	https://www.opendatacube.org https://github.com/opendatacube
Earth System Data Cube (ESDC)	https://deepesdl.readthedocs.io/en/latest/datasets/ESDC
Swiss Data Cube	https://www.swissdatacube.org
Austrian Data Cube	https://acube.eodc.eu
EarthServer	https://earthserver.eu/index.php
EarthCube	https://www.earthcube.org

Table 5. Data cubes.

3.2 User limitations and good practices to address them

As mentioned in Section 2.2, we identified the following limitations in the uptake of (some of) the existing EO platforms: (1) fragmentation, leading to discoverability and interoperability issues; (2) steep learning curve; (3) difficulty to understand what the services offered are and whether they fit the user needs, and pricing being often not transparent; (4) processing workflows being not customisable and open; (5) vendor lock-in; (6) code sharing and reuse not facilitated; (7) lack of assurance about sustainability; (8) internal policies prohibiting the publication of commercial added-value code/algorithms to external EO platforms. For all these limitations, we identified — based on the analysis of the previously classified platforms as well as from the authors' experience — a number of good practices that may help address these limitations.

- For limitations (1), (4), (5), (6) and (7), we suggest the following good practices: i) Releasing software under open source licenses fosters collaboration, reuse and growth of products that are considered to be useful by the community; ii) Adopting open standards and/or open APIs with publicly released specifications (so that other implementations can reuse the software) facilitates reuse and increases interoperability; iii) Federating services (e.g. authentication) reduces fragmentation among available platforms; iv) Open

Name	URL
Open Earth Monitor	https://earthmonitor.org
Earth Observation Exploitation Platform Common Architecture (EOEPCA)	https://eoepca.org — https://github.com/EOEPCA
GEOSS Platform Plus	https://geossplatformplus.com
Copernicus Data Space Ecosystem (CDSE)	https://dataspace.copernicus.eu
Green Deal Data Space Foundation (GREAT Project)	https://www.greatproject.eu
Urban Data Space for Green Deal (USAGE Project)	https://cordis.europa.eu/project/id/101059950
Destination Earth	https://digital-strategy.ec.europa.eu/en/policies/destination-earth
Gaia-X	https://www.gaia-x.eu
Sen4Cap	http://esa-sen4cap.org https://github.com/Sen4CAP
GEANT	https://geant.org
European Open Science Cloud (EOSC)	https://www.eosc-hub.eu
Indian Geo Platform of ISRO	http://bhuvan.nrsc.gov.in/home/index.php
Radiant MLHub — Open Geospatial ML Library	https://mlhub.earth
Scent – Smart Toolbox for Engaging Citizens into a People-Centric Observation Web	https://scent-project.eu

Table 6. Virtual infrastructures.

Name	URL
All Data 4 Green Deal (AD4GD)	https://www.ogc.org/initiatives/ad4gd
e-shape	https://e-shape.eu
Horizon Cloud	https://h-cloud.eu
European Strategy Forum on Research Infrastructures (ESFRI)	https://www.esfri.eu
Climate Change Initiative	https://climate.esa.int/en
Copernicus	https://www.copernicus.eu/en
NASA's Earth Observing System	https://eospso.nasa.gov/content/nasas-earth-observing-system-project-science-office
ESA's Living Planet Programme	https://www.esa.int/Applications/Observing_the_Earth/FutureEO/Living_Planet_programme_introduction
CBERS Programme	https://en.wikipedia.org/wiki/China%E2%80%93Brazil_Earth_Resources_Satellite_program
Disaster Monitoring Constellation	https://www.sstl.co.uk/space-portfolio/the-disaster-monitoring-constellation

Envisat Programme	https://earth.esa.int/eo gateway/missions/envisat
GOSAT Project	https://www.gosat.nies.g o.jp/en
MetOp Programme	https://www.esa.int/Appl ications/Observing_the_E arth/MetOp_overview
EUMETSAT Programme	https://www.eumetsat.int
ESA's Proba Programme	https://www.esa.int/Enab ling_Support/Space_Engin eering_Technology/Proba_ Missions
SPOT Programme	https://spot.cnes.fr/fr
eartH2Observe	https://www.earth2observ e.eu
Earth Challenge 2020	https://earthchallenge20 20.earthday.org
Ecopotential	http://www.ecopotential -project.eu
Global Facility for Disaster Reduction and Recovery	https://www.gfdrr.org/en
NextGEOSS	https://nextgeoss.eu
Regional Centre for Mapping of Resources for Development	http://rcmrd.org/en
World Climate Research Programme (WCRP)	https://www.wcrp-climate .org
WorldPop Research Programme, University of Southampton	https://www.worldpop.org
Pangeo	https://pangeo.io/index. html
ESA Digital Twin Earth	https://www.esa.int/Appl ications/Observing_the_E arth/Working_towards_a_D igital_Twin_of_Earth

Table 7. Initiatives and programmes.

governance, i.e. building a community of users and developers around the project may lead to more people taking a stake in the curation of the project and its outcomes, safeguarding its sustainability over time.

- For limitation (2), we suggest the following good practices: i) Sandbox environments allow users to easily experiment and evaluate whether the offer of a certain platform meets their needs; ii) Training materials (documentation, tutorials and videos) and events (e.g. webinars) are a useful means to provide users with a quick introduction to the main functionalities.

- For limitation (3), we suggest the following good practice: i) Provision of a comprehensive and transparent list of the services offered by the platform and the related costs. The presence of a tool comparing costs and services offered by several platforms (see Section 3.3) partially addresses this limitation.

- For limitations (2), (3), (4) and (7), we suggest the following good practices: i) Embracing co-design, i.e. the process of actively involving users in all phases of design and implementation of the services, including adjusting them according to user feedback (Barbier et al., 2021); ii) Setting up and providing active support in helpdesks, forums and mailing lists; iii) Cultivating the growth of a com-

munity around the project; iv) Adopting open principles in software development and project governance.

- For limitation (8), we suggest the following good practice: i) Adopting permissive licenses enabling the reuse of code and algorithms, including for commercial purposes.

3.3 Technological enablers

Based again on the results of the analysis of digital EO infrastructures made in Section 3.1, we identified a first set of technological enablers, i.e. specific resources such as standards, APIs, data encodings and software components that promote interoperability, usability, and reuse of digital EO platforms. We argue that the adoption of these enablers should be promoted into EO initiatives such as GEOSS and the European infrastructure contributing to it. The following list of technological enablers is not meant to be exhaustive, as it only reflects the authors' opinion and is subject to future updates.

- The NoR Portal (`https://nor-discover.cloudeo.gr oup`) addresses the cost transparency issue, allowing users to compare available pricing offers from various platforms all in one place;

- The Yellow Pages tool (`https://www.geoportal.or g/yellow-pages`) is a catalogue of data providers, which addresses the problem of discoverability; a meta-catalogue that also includes other types of services and platforms would be extremely useful to have;

- The OpenEO API (`https://openeo.org`) addresses the issue of interoperability among EO cloud solutions and prevents vendor lock-in;

- Analysis-Ready Cloud Optimized (ARCO) datasets (Stern et al., 2022) and Cloud Optimised GeoTIFF (`https:// www.cogeo.org`) maximize the usability of EO data in a cloud environment.

- The Open Geospatial Consortium (OGC) APIs (`https:// ogcapi.ogc.org`) are a new generation of OGC standards designed to make it easy for anyone to serve and consume geospatial data on the web.

- The SpatioTemporal Asset Catalogs (STAC, `https://st acspec.org/en`) provides a standardized way to expose collections of spatio-temporal data that improves indexing of assets.

4. DISCUSSION AND CONCLUSIONS

This paper addressed the proliferation and fragmentation of digital infrastructures and initiatives, which currently characterises the EO landscape. The work offers two main contributions. First, we proposed a classification scheme to make sense of the complex landscape of existing platforms for publishing, discovering, processing and combining EO data and services. This scheme was further applied to classify more than 150 infrastructures and initiatives, which — to the authors' knowledge — represents the largest such classification available in literature. Second, by adopting a user-centric approach we distilled the main limitations currently preventing an optimal uptake of existing EO platforms, which were further translated into actionable recommendations for platform managers and developers.

The starting point for our analysis was the task to conceptualise a prototype architecture for the European EO data infrastructure contributing to GEOSS. The review of existing EO platforms performed within this context allows us to make some general reflections. First, platforms and ecosystems provided by the private sector (e.g. Google Earth Engine, Microsoft Planetary Computer, Earth on Amazon Web Services and Digital Earth Africa) seem to be more successful in terms of user communities involved compared to (European) publicly-funded infrastructures. One of the main reasons for this lies exactly in the fragmentation of the current EO landscape, where a large number of infrastructures and initiatives exist, which often overlap in their objectives and/or offer very similar services, sometimes even to the same communities. The information that users typically need for developing solutions cannot be found in just one place, forcing them to deal with different platforms (e.g. for data access, infrastructure setup, account management and design of processing workflows) that often present a steep learning curve and an associated huge effort to get started, in addition to the need for separate authentications and the lack of interoperability. Faced with the task to navigate and integrate this complex landscape, developers may be discouraged from using them, preferring instead to rely on the integrated services offered by hyperscalers and other private sector providers.

The success of private commercial EO platforms seems indeed to derive from a quick adaptation to the changing user needs, in addition to the extensive documentation provided (both written and available through webinars or videos), the high level of support offered to the community, the smart options for discovering and accessing data, the possibility to share code and datasets, the simplicity of access (immediately after signing up users can start building their own products), etc. To achieve the same success, publicly-funded initiatives such as the European EO infrastructure contributing to GEOSS, conceived as a virtual digital ecosystem integrating multiple existing European contributions, will need to incorporate a user-driven approach in its design. This would require identifying and addressing the limitations that current users face when creating solutions from the interplay and integration of multiple elements.

There are several opportunities to extend the current work. First, as mentioned in Section 3.1, our analysis is in no way meant to be exhaustive and to cover the whole spectrum of existing EO platforms. On the contrary, we believe ours is an initial effort that could hopefully stimulate other researchers to update or complement our review. For this reason, we currently plan to setup a dedicated page on a public wiki where the list of classified platforms can be collaboratively improved and extended over time. We still suggest interested researchers to make use of the GEO website and the GEOSS Portal as the entry points for the discovery and investigation of the existing infrastructures.

Similarly, we believe the good practices and the technological enablers we identified (in Sections 3.2 and 3.3, respectively) constitute useful resources for those who are faced with the task of designing and coordinating EO platforms as well as managing their user communities. The specific impact of implementing any of such good practices and technological enablers represents additional fruitful ground for future research.

DISCLAIMER

The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

REFERENCES

Barbier, R., Yahia, S. B., Le Masson, P., Weil, B., 2021. Expanding usages of Earth Observation data: a co-design approach to grow an ecosystem of efficient service designers. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, IEEE, 296–299. `https://doi.org/10.1109/IGARSS47720.2021.9553914`.

Di Leo, M., Kilsedar, C., Spadaro, N., Dowell, M., Kotsev, A., Soler-Garrido, J., 2022. Analysis of e-shape showcases: a JRC perspective on leveraging e-shape results. *EuroGEO Workshop 2022*, Athens, 7-9 December 2022. `https://www.greekgeo.noa.gr/egw2022/images/2022/EGW_Pres/e-shape_Friday/EuroGEO_eShapeSession_JRC_final.pdf` (20 May 2023).

Di Leo, M., Kona, A., Spadaro, N., Minghini, M., Kotsev, A., in press. Pairing user's needs to the current offer of platforms – the case of big Earth Observation. *EuroGEO Workshop 2023*, Bolzano, 2-4 October 2023.

European Commission, 2023. Research infrastructures. `https://europa.eu/!xcw7KD` (20 May 2023).

European Space Agency, 2016. Request for Information "EO Exploitation Network of Platforms Architecture". `https://shorturl.at/bftxF` (20 May 2023).

Gomes, V. C., Queiroz, G. R., Ferreira, K. R., 2020. An overview of platforms for big earth observation data management and analysis. *Remote Sensing*, 12(8), 1253. `https://doi.org/10.3390/rs12081253`.

Harris, R., Baumann, I., 2015. Open data policies and satellite Earth observation. *Space Policy*, 32, 44-53. `https://doi.org/10.1016/j.spacepol.2015.01.001`.

Kotsev, A., Minghini, M., Cetl, V., Penninga, F., Robbrecht, J., Lutz, M., 2021. INSPIRE: A Public Sector Contribution to the European Green Deal Data Space. EUR 30832 EN, Publications Office of the European Union, Luxembourg. `https://doi.org/10.2760/8563`.

Ranchin, T., Ménard, L., Fichaux, N., Reboul, M. et al., 2021. e-shape-EuroGEO Showcases: application powered by Europe contribution to EuroGEO and to the development of the EO industry. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, IEEE, 293–295. `doi.org/10.1109/IGARSS47720.2021.9553426`.

Stern, C., Abernathey, R., Hamman, J., Wegener, R., Lepore, C., Harkins, S., Merose, A., 2022. Pangeo Forge: Crowdsourcing Analysis-Ready, Cloud Optimized Data Production. *Frontiers in Climate*, 3, 179. `https://doi.org/10.3389/fclim.2021.782909`.

Voidrot-Martinez, M.-F., 2022. Status on the usability of the available European Platforms based on e-shape return of experience. *EuroGEO Workshop 2022*, Athens, 7-9 December 2022. `https://www.greekgeo.noa.gr/egw2022/images/2022/EGW_Pres/Towards_the_digital_age__platforms_and_data_spaces_for_the_EU_Green_Deal_-_Part_I/102_-_e-shape_-_Voidrot.pdf` (20 May 2023).

Wagemann, J., Siemen, S., Seeger, B., Bendix, J., 2021. Users of open Big Earth data–An analysis of the current state. *Computers & Geosciences*, 157, 104916. `https://doi.org/10.1016/j.cageo.2021.104916`.