SwissEnvEO: A FAIR National Environmental Data Repository for Earth Observation Open Science

Giuliani, Gregory; Cazeaux, Hugues; Burgi, Pierre-Yves; Poussin, Charlotte; Richard, Jean-Philippe; Chatenoux, Bruno

How to cite

GIULIANI, Gregory et al. SwissEnvEO: A FAIR National Environmental Data Repository for Earth Observation Open Science. In: Data Science Journal, 2021, vol. 20, n° 22. doi: 10.5334/dsj-2021-022

This publication URL: https://archive-ouverte.unige.ch/unige:152005

Publication DOI: 10.5334/dsj-2021-022

RESEARCH PAPER

SwissEnvEO: A FAIR National Environmental Data Repository for Earth Observation Open Science

GREGORY GIULIANI
HUGUES CAZEAUX
PIERRE-YVES BURGI
CHARLOTTE POUSSIN
JEAN-PHILIPPE RICHARD
BRUNO CHATENOUX

*Author affiliations can be found in the back matter of this article

ABSTRACT

Environmental scientific research is highly becoming data-driven and dependent on high performance computing infrastructures to process ever increasing large volume and diverse data sets. Consequently, there is a growing recognition of the need to share data, methods, algorithms, and infrastructure to make scientific research more effective, efficient, open, transparent, reproducible, accessible, and usable by different users.

However, Earth Observations (EO) Open Science is still undervalued, and different challenges remains to achieve the vision of transforming EO data into actionable knowledge by lowering the entry barrier to massive-use Big Earth Data analysis and derived information products. Currently, FAIR-compliant digital repositories cannot fully satisfy the needs of EO users, while Spatial Data Infrastructures (SDI) are not fully FAIR-compliant and have difficulties in handling Big Earth Data.

In response to these issues and the need to strengthen Open and Reproducible EO science, this paper presents SwissEnvEO, a Spatial Data Infrastructure complemented with digital repository capabilities to facilitate the publication of Ready to Use information products, at national scale, derived from satellite EO data available in an EO Data Cube in full compliance with FAIR principles.

CORRESPONDING AUTHOR:
Gregory Giuliani
University of Geneva, Institute for Environmental Sciences, GRID-Geneva, Bd Carl-Vogt 66, CH-1211 Geneva, CH; University of Geneva, Institute for Environmental Sciences, EnviroSPACE lab., 66 Boulevard Carl-Vogt, 1205 Geneva, CH
gregory.giuliani@unige.ch

KEYWORDS:
Data Sharing; Interoperability; Earth Observations; Data Cube; Environmental monitoring; Switzerland; Open Science; Reproducibility

TO CITE THIS ARTICLE:
Giuliani, G, Cazeaux, H, Burgi, P-Y, Poussin, C, Richard, J-P and Chatenoux, B. 2021. SwissEnvEO: A FAIR National Environmental Data Repository for Earth Observation Open Science. *Data Science Journal*, 20: 22, pp. 1–18. DOI: *https://doi.org/10.5334/dsj-2021-022*

1. INTRODUCTION

Environmental sustainability is now recognized as a critical global issue after that the World Economic Forum (WEF) has classified for the first time five environmental risks (e.g., climate change, biodiversity loss) at the first five positions of global risks in terms of likelihood (World Economic Forum, 2020). This requires global coordinated efforts from countries to adequately managed the limited natural resources available and the increasing pressures to meet the needs of a growing population (IPBES, 2018).

Governments have national and international reporting commitments and obligations as well as national environmental programs. To efficiently and effectively manage their natural resources, they need to have adequate information and knowledge about the limits of the planet so that they can decide how best to use these limited resources (Breuer et al., 2019; Fritz et al., 2019; Nativi et al., 2019). Consequently, they need information that is consistent, spatially explicit, and sufficiently detailed to capture anthropogenic impacts, and national in scope as recommended by the United Nations 2030 Agenda for Sustainable Development (Saito et al., 2017; *The future we want*, 2012).

To support environmental monitoring, satellites are providing since 1972 continuous, synoptic, multi-spectral observations of our planet (Wulder et al., 2012). These remotely sensed Earth Observations (EO) data can significantly contribute to establish the baseline for determining trends, defining present conditions, and informing future evolution (Estoque, 2020). Hence, satellite EO data have the potential to drive progress against key national and international development agendas providing new insights and support better policy making across diverse issues of environmental sustainability (Andries, Morse, Murphy, Lynch, & Woolliams, 2019; Dhu et al., 2019; Gregory Giuliani, Egger, et al., 2020; Kavvada et al., 2020).

The following characteristics of satellite EO data can bring significant benefits to support directly or indirectly environmental indicators (Merodio Gómez et al., 2019):

(1) *Spatial resolution:* capacity to provide information potentially every 10 m × 10 m;
(2) *Temporal resolution:* capacity to capture data at different frequency of revisit (e.g., 5 days for Sentinel-2, 16 days for Landsat, up to 2.9 days in combined use mode and closer to the equator) (Li & Roy, 2017);
(3) *Scale:* capacity to provide information for scales ranging from local up to global;
(4) *Time-series:* capacity to provide continuous data starting as early as 1972 (e.g., Landsat);
(5) *Multi-spectral:* provides measurement in different wavelengths (e.g., visible, thermal near-infrared) to capture different information on various environmental components (e.g., land, water);
(6) *Consistency:* gives the ability to compare generated information in a consistent manner at various scales;
(7) *Complementarity:* data can be validated using additional sources such as sensors or crowd-sources data; and
(8) *Availability:* Landsat and Copernicus, the largest satellite EO data providers, assure access to data for the next two decades.

However, the vision of EO data-driven decision making has not been fully addressed, challenges remain, and therefore the full information potential of EO data has not been yet realized (Andries, Morse, Murphy, Lynch, Woolliams, et al., 2019). Effective and efficient use of EO data freely and openly available (e.g., Landsat, Advanced Very High Resolution Radiometer (AVHRR), Moderate Resolution Imaging Spectroradiometer (MODIS)) from different data repositories (e.g., NASA's Earth Observing System Data and Information System (EOSDIS) and its Distributed Active Archive Centers (DAACs) – *https://earthdata.nasa.gov/*) in policy and decisions-making processes are hampered by Big Data challenges (e.g., Volume, Variety, Velocity) and associated complexities (e.g., data preparation, model integration & prediction) (Guo, 2017). This poses several issues in terms of Volume (e.g., data volumes have increased by 10 in the last 5 years); Velocity (e.g., Sentinel-2 is capturing a new image of a given place every 5 days); and Variety (e.g., different type of sensors, spatial/spectral resolutions). For example, the Copernicus Sentinel-2 generates 22'000 images to cover the Earth every 5 days which corresponds to more than 1'600'000 images per year (Sudmanns et al., 2019). Copernicus is the largest EO data provider in the World distributing 250TB daily, the archive currently holds 250PB of data and the daily growth rate is approximately 220TB (Drusch et al., 2012). Therefore, classic approaches to

EO data handling, distribution and processing have constraints (e.g., data size, heterogeneity and complexity) that impede their extensive use and analysis (Guo et al., 2020).

In recent years, the increasing production of Analysis Ready Data (ARD) (i.e., data processed to a minimum set of requirements allowing immediate analysis) (Lewis et al., 2018; Simonis, 2019; Zhu, 2019) by satellite data providers together with the emergence of tools and cloud-based analysis platforms have significantly reduced to work required to exploit and analyze large volumes of EO data. The Earth Observations Data Cube (EODC) concept has emerged as promising technical solution for lowering barriers in EO data handling, facilitating the connection between data, application and users (Dwyer et al., 2018; Zhu, 2019), and offering new significant possibilities to deliver applications and products harnessing satellite EO data for tackling environmental, economic and societal challenges (P. Baumann, 2018; Peter Baumann et al., 2019). Using EODC and ARD, the spatio-temporal analysis of large volumes of EO data will dramatically be eased by providing direct access to satellite data over a given region that are already pre-processed (e.g., geometrically, and atmospherically corrected).

Different countries have already adopted the EODC concept to investigate how to benefit and integrate information generated from EO data into their policies and information systems (Ariza-Porras et al., 2017; Asmaryan et al., 2019; Dhu et al., 2017). Switzerland has initiated the Swiss Data Cube (SDC) project in late 2016, an initiative supported by the Federal Office for the Environment (FOEN) and is developed, implemented and operated by four research institutions: the United Environment Program (UNEP)/GRID-Geneva in partnership with the University of Geneva (UNIGE), the University of Zurich (UZH), and the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) (Chatenoux et al., n.d.). The SDC is a cloud-based analytical platform based on the Open Data Cube (Killough, 2018; Rizvi et al., 2018) that includes 36 years of Analysis Ready Data of optical (e.g., Landsat 5-7-8, Sentinel-2) (G. Giuliani et al., 2018; Gregory Giuliani, Chatenoux, et al., 2017) and radar (e.g., Sentinel-1) (Truckenbrodt et al., 2019) imagery over the entire swiss territory accounting for a total volume of 6TB, corresponding to more that 1000 billion observations, and updated on a daily basis. It minimizes the time and scientific knowledge needed to handle and analyze large volumes of satellite EO data with consistent, spatially aligned, calibrated observations. The SDC is a unique asset to track environmental changes nationwide using EO data, allowing developing adequate responses to problems of national significance (Gregory Giuliani, Chatenoux, et al., 2017). It has already allowed to generate information products on Sustainable Development Goals (SDG) (Gregory Giuliani, Chatenoux, et al., 2020; Honeck et al., 2018) and snow cover evolution (Poussin et al., 2019; Salzano et al., 2019). However, none of these products are available in the geographical information platform of the Swiss Confederation Federal Administration (*https://www.geo.admin.ch*) nor other information products derived from satellite EO data. This situation is caused by the lack of awareness on the potential of EO-based information products and their availability. Consequently, there are opportunities to improve how EO data are used to support decision-making and increase the societal value of information and data products (Virapongse et al., 2020). In particular, to enhance the EO value chain from data to knowledge (Rowley, 2007) it is critical to lower the entry barriers to facilitate uptake and use of products by adhering to standards that increase interoperability. Therefore, making interoperable products of downstream services is an essential prerequisite (Gregory Giuliani, Masó, et al., 2019). Traditionally, geospatial data are managed through Spatial Data Infrastructure (SDI) that are platforms for facilitating the exchange, share, use and integration of geospatial data based on interoperability standards (G. Giuliani et al., 2011, 2014; Gregory Giuliani, Nativi, et al., 2017). SDI can be local (e.g., Geneva State SDI – *http://www.sitg.ch*), national (e.g., the Canadian Geospatial Data Infrastructure (CGDI) – *https://www.nrcan.gc.ca/science-and-data/science-and-research/earth-sciences/geomatics/canadas-spatial-data-infrastructure/10783*), regional (e.g., America's SDI) (Merodio Gómez et al., 2019), global (e.g., Global Earth Obsetvation System of Systems – *http://www.geoporal.org*), thematic (e.g., Global Risk Data Platform) (G. Giuliani & Peduzzi, 2011) or institutional (e.g., United Nations SDI) (Henricksen, 2007). However, SDI are not well suited to deal with Big Earth Data (Innerebner et al., 2017) and there is a demand for enhancing SDI to properly deal with the vast amount of geospatial and EO data (Gomes et al., 2020; Guo et al., 2020). In conjunction, the scientific community is calling for more open and reproducible science to improve reliability, efficiency and credibility of scientific research (Munafò et al., 2017). Making data available through common guidelines is not yet a widely adopted approach in many scientific communities (Stall et al., 2019). To broadly enhance the reusability of scientific data, the Findable, Accessible, Interoperable and Reusable (FAIR) principles have been proposed and promoted (Wilkinson et al., 2016). Nonetheless, openness is not yet fully embedded in environmental research process and FAIR-compliance is not sufficient for open science. FAIR-compliance makes it easy to find, access and reuse data, but free and open data and information policies are essential for open science. Sharing of data, methods and tools are happening more on a case-by-case basis than on fully open platforms. In addition, the need to support long-term preservation of data, code and documentation is recognized as a major obstacle to make sense of data (Science Business, 2019) that has been recently addressed by ISO with the standard 9165-2:2020 (Geographic information — Preservation of digital data and metadata — Part 2: Content specifications for Earth observation data and derived digital products) (ISO, 2020; Peng et al., 2021).

Taking these elements into consideration, the objective of this paper is to present SwissEnvEO, a data repository of environmental information products produced with the Swiss Data Cube, delivered in compliance with the FAIR Data Sharing principles and supported by SDI, to facilitate the discovery, access and use of these products and ultimately make them available and used by the widest audience possible.

2. METHODOLOGY & IMPLEMENTATION

2.1 EARTH OBSERVATIONS DATA MANAGEMENT

EODC is an effective means to operate Big Earth Observations Data (or Big Earth Data) produced by satellites and made available under open data licenses from various data providers (Harris & Baumann, 2015). Various implementations are available ranging from software libraries like gdalcubes (Appel & Pebesma, 2019) or dtwSat (Maus et al., 2019); complete software stacks such as the Open Data Cube (ODC) (Killough, 2018) or RasDaMan/EarthServer (Peter Baumann et al., 2016, 1997); processing platforms like e-sensing (Camara et al., 2017) or the JRC Earth Observation Data and Processing Platform (JEODPP) (Soille et al., 2018) ; or cloud-based processing facilities such as the Copernicus Data and Information Access Services (DIAS) (European Commission, 2018), the Google Earth Engine (GEE) (Gorelick et al., 2017) or the System for Earth Observation Data Access, Processing and Analysis for Land Monitoring (SEPAL) (Tondapu et al., 2018). These diverse approaches contribute to progress towards the vision of an Open and Reproducible Earth Observation Science with the ultimate goal to unlock the information power of satellite EO data by transforming them into decision-ready products and actionable knowledge (Gregory Giuliani, Camara, et al., 2019).

EODC has broadened the usage EO data to different communities enabling to convert large volume of satellite data into meaningful geophysical variables for monitoring environmental changes (Gregory Giuliani, Dao, et al., 2017). However, to further reduce the gap between user and decision-ready products additional efforts are required to facilitate the discovery, access and use of these information products (Peter Baumann, 2019). To reach this objective, a frequent solution is to adhere to a set of standards and best practices (Amorim et al., 2017). In recent year, the Findable-Accessible-Interoperable-Reusable (FAIR) principles have been promoted in various scientific communities providing guidance for scientific data management and stewardship (Stall et al., 2019; Wilkinson et al., 2016). These principles are addressing data producers needs to maximize the use of research data in the current digital science by facilitating sharing and accessibility to digital data (Bermudez, 2017; Ferrari et al., 2018; Wagner & Jonkers, 2017). To support such a systemic change of science practices towards an Open and Reproducible Science, infrastructures have emerged providing services to assist data publishers to comply with FAIR principles. In short, data can be deposited into Digital Repositories, such as Zenodo (*https://zenodo.org*) or Pangaea (*https://www.pangaea.de*), that provide capabilities to expose scientific data and make them *findable* (i.e., by anyone through common search tools), *accessible* (i.e., data and metadata can be explored), *interoperable* (i.e., data and metadata can be exchanged, used and integrated using standards) and *reusable* (i.e., by others with clear usage licenses and documentation). If this general approach can be useful to strengthen scientific data reusability, it is not fully convenient to address the specificities of EO data. Satellite imagery and products are part of the geospatial domain and usually geospatial data discoverability, accessibility, sharing and re-use is commonly achieved through Spatial Data Infrastructures (SDI) (Kotsev et al., 2020). SDI have been developed and implemented for more than two decades and have largely demonstrated their capabilities at different scales, for digital geospatial data production, management, analysis and diffusion (Gregory Giuliani et al., 2016; Guo et al., 2020). To achieve interoperability and facilitate data discovery and access of heterogeneous geospatial resources, SDI mostly rely on Open Geospatial Consortium (OGC) and the International Organization for Standardization (ISO) standards who defined a set of specifications addressing the needs of the geospatial community. Data can be documented with metadata using ISO19115 (resource metadata), ISO19139 (metadata encoding) and complemented by the OGC Catalog Service for the Web (CSW) specification defining an interoperable interface to publish, discover, search and query metadata. Vector and Raster data are published respectively with Web Feature Service (WFS) and Web Coverage Service (WCS) and both can be visualized with the Web Map Service (WMS) (G. Giuliani et al., 2016; Ramage, 2011). Currently, a lot of efforts are devoted to increase the interoperability of EODC, using OGC and ISO standards, both for upstream (i.e., data ingestion, data pre-processing) and downstream services (i.e., application or products development) (Gregory Giuliani, Masó, et al., 2019; Gomes et al., 2020; Maso et al., 2019; Nativi et al., 2017) to make EO data Science more open and reproducible (Gregory Giuliani, Camara, et al., 2019). Even if geospatial data managed through SDIs can be broadly considered as FAIR, they are not fully compliant with these principles. Therefore, the challenge to be addressed is to make EO data and products fully FAIR compliant (to ensure the needs of the broad scientific community) while at the same time preserving the specificities and needs of the geospatial domain. Specifically, a solution is required to benefit from the two disciplines increasing the openness, transparency, traceability and reproducibility in EO science, improving the sustainable use and increasing the value of EO digital resources (Coetzee et al., 2020; Curry & Moosdorf, 2019). ***Table 1*** is synthetizing identified issues related to EO and FAIR principles.

Table 1 Identified issues of SDI and FAIR digital data repositories when used with Big Earth Data.

FAIR PRINCIPLES	ISSUES WITH BIG EARTH DATA
Findable	
F1. (Meta)data are assigned a globally unique and persistent identifier	(1) DOI are not widely used in SDI. (2) Persistent identifiers are not widely used in SDI.
F2. Data are described with rich metadata	Metadata in Digital Repositories (usually following DataCite) are not sufficiently detailed to describe EO data
F3. Metadata clearly and explicitly include the identifier of the data they describe	SDI and Digital Data repositories provide this capability
F4. (Meta)data are registered or indexed in a searchable resource	SDI and Digital Data repositories provide this capability
Accessible	
A1. (Meta)data are retrievable by their identifier using a standardised communications protocol	(1) In SDI, CSW is the de facto standard (2) In Digital Repositories OAI (3) For data, Digital Repositories do not support OGC standards
A1.1 The protocol is open, free, and universally implementable	SDI and Digital Data repositories provide this capability
A1.2 The protocol allows for an authentication and authorisation procedure, where necessary	SDI and Digital Data repositories provide this capability
A2. Metadata are accessible, even when the data are no longer available	SDI and Digital Data repositories provide this capability
Interoperable	
I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.	SDI and Digital Data repositories provide this capability using XML and JSON encodings
I2. (Meta)data use vocabularies that follow FAIR principles	This is not yet a widely adopted capability in SDI even if efforts exists such as in INSPIRE (Patroumpas et al., 2015)
I3. (Meta)data include qualified references to other (meta)data	SDI and Digital Data repositories provide this capability
Reusable	
R1. Meta(data) are richly described with a plurality of accurate and relevant attributes	ISO19xxx standards are more suited than those used in Digital Repositories (e.g., DataCite)
R1.1. (Meta)data are released with a clear and accessible data usage license	SDI and Digital Data repositories provide this capability
R1.2. (Meta)data are associated with detailed provenance	SDI and Digital Data repositories provide this capability
R1.3. (Meta)data meet domain-relevant community standards	SDI and Digital Data repositories provide this capability

Based on the identified issues, the proposed approach is aiming to enhance traditional SDI with a long-term preservation digital repository ensuring full compliance with FAIR principles while at the same time benefiting from geospatial services capabilities (*Figure 1*).

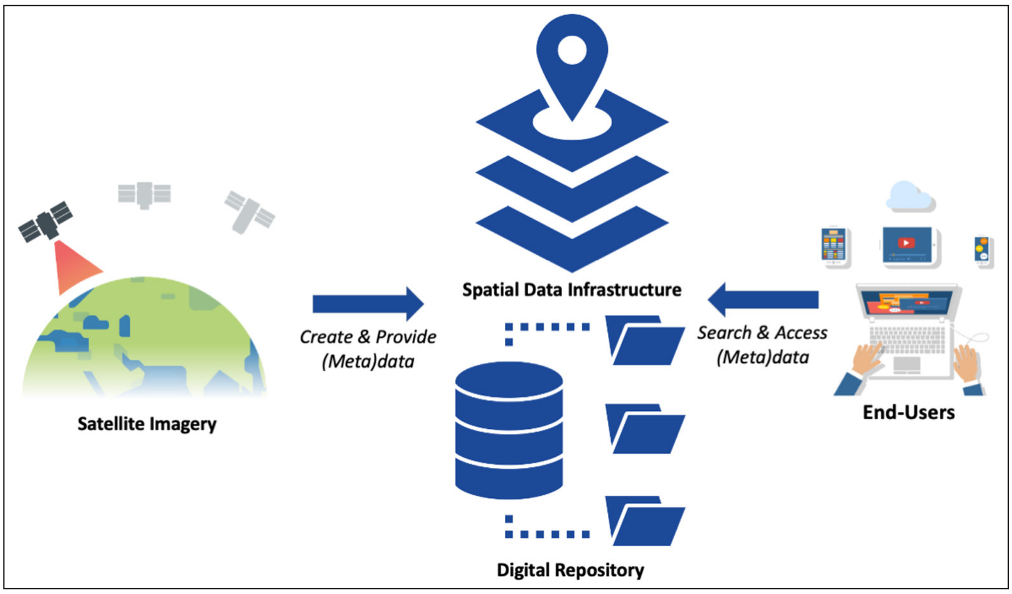

Figure 1 General architecture of a FAIR EO data repository.

2.2 IMPLEMENTATION

To facilitate discovery, access and use of data and information products generated with the Swiss Data Cube, it has been decided to separate upstream and downstream services (Denis et al., 2017). Upstream services are those interacting with the infrastructure (e.g., pre-processing, download of raw images) whereas downstream services are meant to access decision-ready/value-added products (Gregory Giuliani, Masó, et al., 2019). These downstream services are provided by the Swiss Environmental Earth Observations (SwissEnvEO) data repository following the schema presented in *Figure 2*.

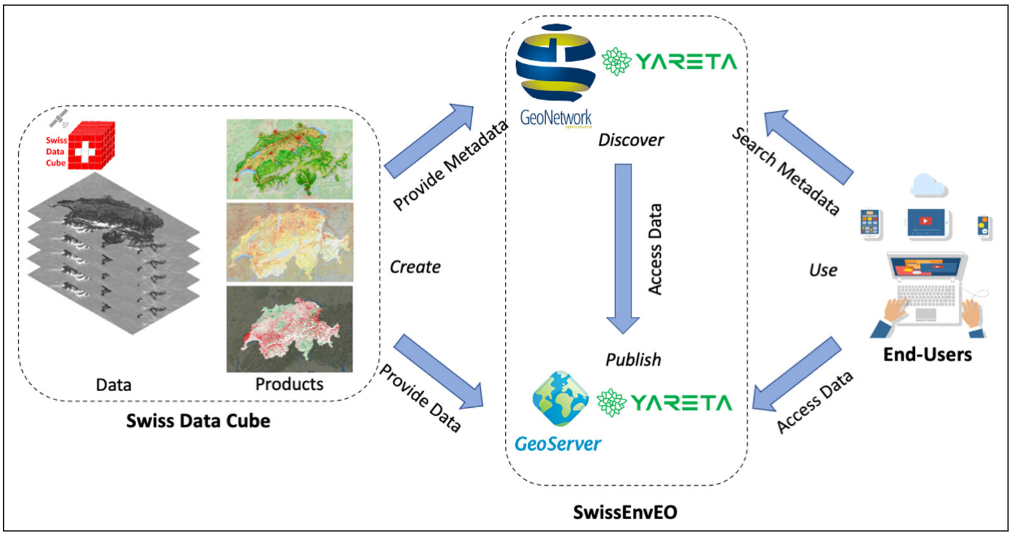

Figure 2 SwissEnvEO Architecture and implemented software components.

SwissEnvEO is an extension of a previous system that was purely based on SDI concepts and technologies. It adopts SDI components, namely GeoNetwork and GeoServer, for creating, managing and publishing respectively metadata and data according to widely adopted geospatial information standards promoted by the OGC and ISO/TC211 (Coetzee et al., 2020; Kotsev et al., 2020). These components ensure compatibility with the geospatial domain. To align with FAIR and expend the usage of data products generated with the Swiss Data Cube, the SDI component is complemented with Yareta, the University of Geneva digital solution for archiving and preserving research data for long term. It has been developed to facilitate the work of researchers, including the management, and sharing of research data. Designed in strict accordance with the FAIR principles that govern research data management best practices, Yareta fulfills the Data Repository requirements of funders. Therefore, the proposed solution for the downstream tier of the Swiss Data Cube strengthens the publication and sharing of final value-added/decision-ready products (e.g., validated analysis results) while at the same time separating the usage of the SDC between scientific/data analysts end-users and more general end-users. This then helps to further broaden the usage of these products. It took a couple of days to connect GeoNetwork with the digital repository while the longest process (i.e., a few weeks) was the production of all archives in compliance with the guidelines for research data management. ***Table 2*** summarizes the implemented solutions and main capabilities of SwissEnvEO (launched in December 2020) in accordance with FAIR principles.

FAIR PRINCIPLES	IMPLEMENTED SOLUTIONS
Findable	
F1. (Meta)data are assigned a globally unique and persistent identifier	DOI [Yareta] & UUID [GeoNetwork]
F2. Data are described with rich metadata	ISO19115-2/ISO19139-2 [GeoNetwork] & DataCite [Yareta]
F3. Metadata clearly and explicitly include the identifier of the data they describe	Embedded in ISO & DataCite
F4. (Meta)data are registered or indexed in a searchable resource	GeoNetwork and Yareta both provide this functionality
Accessible	
A1. (Meta)data are retrievable by their identifier using a standardised communications protocol	*Metadata:* OGC CSW, OAI-PMH, Z39.50, THREDDS, Webdav, WAF, OpenSearch *Data:* OGC WMS, WCS
A1.1 The protocol is open, free, and universally implementable	Based on OGC standards
A1.2 The protocol allows for an authentication and authorisation procedure, where necessary	True, already supported in OGC standards as well as in the Yareta API
A2. Metadata are accessible, even when the data are no longer available	True, using GeoNetwork
Interoperable	
I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.	OGC/ISO standards for geospatial data rely on XML and JSON and provide clear usage guidelines
I2. (Meta)data use vocabularies that follow FAIR principles	True with the use of Yareta. GeoNetwork has a capability to add vocabularies
I3. (Meta)data include qualified references to other (meta)data	True, it feasible both in GeoNetwork and Yareta
Reusable	
R1. Meta(data) are richly described with a plurality of accurate and relevant attributes	True with the use of ISO 19xxx standards
R1.1. (Meta)data are released with a clear and accessible data usage license	In ISO you can easily provide licenses
R1.2. (Meta)data are associated with detailed provenance	In ISO, provenance should be provided
R1.3. (Meta)data meet domain-relevant community standards	True with OGC and ISO standards

Table 2 FAIR principles and implemented solutions.

3. RESULTS

To validate the technical feasibility, identify possible issues and determine the potential of the proposed solution, the implementation has been tested on the SDC Analysis Ready Data collections for Landsat 5-7-8, Sentinel-1 and Sentinel-2 (Chatenoux et al., n.d.) as well as two time-series products (annual and seasonal) from environmental indices: the Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1974) and the Normalized Difference Water Index (NDWI) (Gao, 1996). These two indices are widely used in Earth Observations and are considered as Essential Variables for environmental monitoring (Gregory Giuliani, Egger, et al., 2020). In total, SwissEnvEO currently accounts for nine data products (for a total volume of 7TB): The five ARD collections are described in Chatenoux et al. (2021) and the four spectral indices presented in ***Table 3***.

These data sets are annual and seasonal time-series generated from 34 years of Landsat Analysis Ready Data (ARD). Normalized Difference Vegetation Index (NDVI) quantifies vegetation by measuring the difference between near-infrared (NIR) (which vegetation strongly reflects) and red light (R) (which vegetation absorbs) using this generic formula (1):

PRODUCTS	LINKS
Normalized Difference Vegetation Index (NDVI) – Annual Mean – Switzerland	*Metadata: https://geonetwork.swissdatacube.org/geonetwork/srv/eng/catalog.search#/metadata/ddd5e734-1f1a-4e06-9402-7041ec625119* *DOI: https://doi.org/10.26037/yareta:kpmscrogqbdhvjeuev2ydrzk7y* *DataCite: https://search.datacite.org/works/10.26037/yareta:kpmscrogqbdhvjeuev2ydrzk7y*
Normalized Difference Vegetation Index (NDVI) – Seasonal Mean – Switzerland	*Metadata: https://geonetwork.swissdatacube.org/geonetwork/srv/eng/catalog.search#/metadata/8e3ae49a-fb2e-44cb-b83d-729c66f7738b* *DOI: https://doi.org/10.26037/yareta:voy277qzczgzbgeiyrcldztuti* *DataCite: https://search.datacite.org/works/10.26037/yareta:voy277qzczgzbgeiyrcldztuti*
Normalized Difference Water Index (NDWI) – Annual Mean – Switzerland	*Metadata: https://geonetwork.swissdatacube.org/geonetwork/srv/eng/catalog.search#/metadata/1008ba03-a57d-42d0-b7d7-3a861d91c4be* *DOI: https://doi.org/10.26037/yareta:xzczpcai2nbp5l4na7rx2oelse* *DataCite: https://search.datacite.org/works/10.26037/yareta:xzczpcai2nbp5l4na7rx2oelse*
Normalized Difference Water Index (NDWI) – Seasonal Mean – Switzerland	*Metadata: https://geonetwork.swissdatacube.org/geonetwork/srv/eng/catalog.search#/metadata/af697b57-cf0a-4dc3-aa46-b394cf9f8c72* *DOI: https://doi.org/10.26037/yareta:bwtgg2z5cbhf3e47rugqhmhgui* *DataCite: https://search.datacite.org/works/10.26037/yareta:bwtgg2z5cbhf3e47rugqhmhgui*

Table 3 Products with their respective metadata, DOI and DataCite search linkgs.

$$\mathrm{NDVI} = (\mathrm{NIR} - \mathrm{R}) / (\mathrm{NIR} + \mathrm{R}) \quad (1)$$

NDVI values ranges from –1 to +1. NDVI is used to estimate vegetation greenness and is useful in understanding vegetation density and assessing changes in plant health.

NDWI quantifies plant water content by measuring the difference between Near-Infrared (NIR) and Short-Wave Infrared (SWIR) (or Green) channels using this generic formula (2):

$$\mathrm{NDWI} = (\mathrm{NIR} - \mathrm{SWIR}) / (\mathrm{NIR} + \mathrm{SWIR}) \quad (2)$$

NDWI values ranges from –1 to +1. NDWI is a good proxy for plant water stress and therefore useful for drought monitoring and early warning. NDWI is also sometimes referred as Normalized Difference Moisture Index (NDMI).

In SwissEnvEO, the entry point for searching data products is represented by GeoNetwork: *https://geonetwork.swissdatacube.org/*. Data are stored once in the Network Attached Storage (NAS) of the University of Geneva and are further accessed by both GeoServer and Yareta (***Figure 2***). GeoServer provides interfaces for geospatial services (e.g., WMS, WCS) whereas Yareta provide long-term preservation capabilities and access to a static copy of the data sets.

The current available services from SwissEnvEO are listed below:

Discovery services

- CSW 2.0.2: *https://geonetwork.swissdatacube.org/geonetwork/srv/eng/csw?service=CSW&version=2.0.2&request=GetCapabilities*
- OpeanSearch EO 1.0: *https://geoserver.swissdatacube.org/geoserver/oseo/description*
- OAI: *https://yareta.unige.ch/oai*

View services

- Web Map Service (WMS) with EO extension 1.1.1/1.3.0: *https://geoserver.swissdatacube.org/geoserver/ows?service=wms&version=1.1.1&request=GetCapabilities*
- Web Map Tile Service (WMTS) 1.0.0: *https://geoserver.swissdatacube.org/geoserver/gwc/service/wmts?REQUEST=GetCapabilities*
- Tile Map Service (TMS) 1.0.0: *https://geoserver.swissdatacube.org/geoserver/gwc/service/tms/1.0.0*
- Web Map Tile Cached (WMS-C) 1.1.1: *https://geoserver.swissdatacube.org/geoserver/gwc/service/wms?request=GetCapabilities&version=1.1.1&tiled=true*

Download services

- Web Coverage Service (WCS) with EO extension 1.0.0 / 1.1.0 / 1.1 / 1.1.1 / 2.0.1: *https://geoserver.swissdatacube.org/geoserver/ows?service=wcs&version=1.0.0&request=GetCapabilities*
- Web Feature Service (WFS) 1.0.0 / 1.1.0 / 2.0.0: *https://geoserver.swissdatacube.org/geoserver/ows?service=wfs&version=1.0.0&request=GetCapabilities*

In addition to these services, GeoNetwork and Yareta provide Application Programming Interfaces (API) to support developers in integrating resources into their applications:

- GeoNetwork: *https://geonetwork.swissdatacube.org/geonetwork/doc/api/index.html*
- Yareta: *https://yareta.unige.ch/docs/DLCM-APIs.html*

With this suite of services, it is possible to make EO data products generated with the Swiss Data Cube fully FAIR-compliant. It allows to make resources *Findable* using the different discovery services. Resources can be directly searched in the GeoNetwork catalog or in Yareta (***Table 3***).

Metadata have unique persistent identifiers (i.e., DOI and UUID) and described following ISO19115-2/19139-2 and DataCite schemas (***Table 3***). The ISO schema for metadata has been chosen because it is the most commonly used metadata standard in the geospatial community and adopted by the major geospatial data sharing initiatives such as the Global Earth Observation System of Systems (GEOSS) or the Infrastructure for Spatial Information in the European Community (INSPIRE) (European Commission, 2007; GEO secretariat, 2008).

The provided discovery interfaces allow also to directly search resources in dedicated desktop applications such as Geographical Information Systems (GIS). Figure 6 shows an example of a search with the "NDVI" term in MetaSearch, a QGIS (*https://www.qgis.org*) plugin to search metadata catalogs with the CSW interface. Once a record has been selected, then users can obtain details and can find relevant links enabling them to access data either as dynamic WMS service or download a static copy with the DOI link (***Figure 3***).

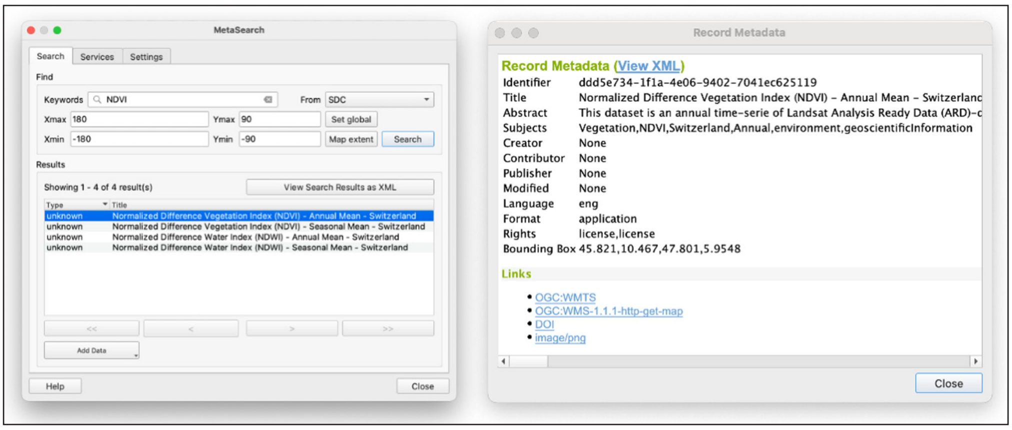

Figure 3 Results of search in MetaSearch using the CSW interface (left) and details on the metadata record with the relevant links for accessing data.

The use of OGC services and DOI broaden the usage of the EO products stored in SwissEnvEO making all resources easily and seamlessly *Accessible*. Metadata are retrievable using standardized interfaces such as the OGC CSW, OAI-PMH and data using OGC WMS and WCS. ***Figure 4*** illustrates that the same dataset discovered before (i.e., NDVI Annual Mean) can be visualized and accessed in different client applications such as a dedicated application from the Swiss Data Cube; the mapping platform of the Swiss government; a desktop GIS application; or in Google Earth.

Consequently, all resources available and accessible in SwissEnvEO are *Interoperable* and *Reusable*, being published with the suite of standardized services mentioned previously.

Besides the visualization, data can be also downloaded using either the WCS service or a static copy from Yareta by clicking on the metadata record DOI link (*https://doi.org/10.26037/yareta:kpmscrogqbdhvjeuev2ydrzk7y*) and developers can further benefit from APIs and services to integrate data products into their own applications. The provided API (*https://admin.yareta.unige.ch/administration/docs/DLCM-IntegrationGuide.html*) follows the Open Archival Information System

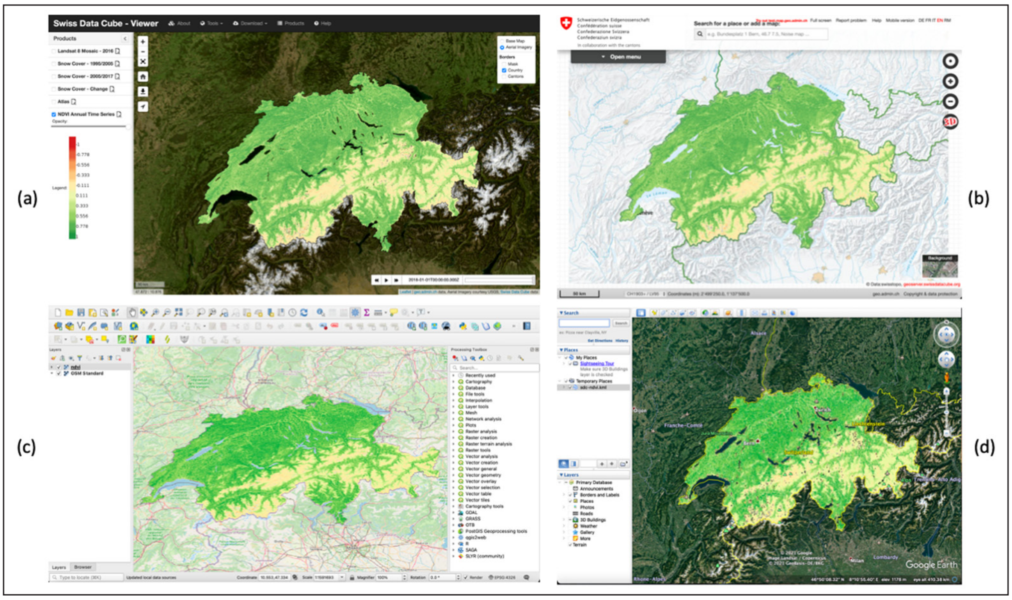

Figure 4 NDVI Annual mean data set served as WMS as seen in **(a)** the Swiss Data Cube Viewer; **(b)** the mapping platform of the Swiss Confederation; **(c)** in a GIS Desktop client and **(d)** in Google Earth Pro.

(OAIS – ISO14721) model and best practices of preservation (Peng et al., 2021). It allows to create a deposit and add data files to it as well as general management functionalities such as approval, update, delete, authenticate. It also allows to search an archive (by archive ID or DOI), download it or export its metadata description. More details on the API and all functionalities are available at: *https://admin.yareta.unige.ch/administration/docs/DLCM-APIs.html*.

4. DISCUSSION

The proposed approach demonstrates that it is feasible to make EO data products fully compliant with FAIR principles while at the same time benefiting for the full suite of standards widely used in the geospatial community. Such solution increases the fairness of geospatial data making them "geoFAIR". It should be noted that even in this paper examples are related to land EO, the proposed approach is sufficiently generic to be applied on any type of EO data products (e.g., water).

The implementation allows storing data in a single place and then publish them according to different standardized service interfaces. Interoperability of EODC has been recognized has a major challenge for the global change and earth system science communities and therefore such a solution can contribute to overcome this important issue and preventing that EODC become silos of information (Gregory Giuliani, Masó, et al., 2019). It helps to move towards more open and reproducible EO science by providing access to Analysis Ready Data and products compliant with FAIR principles, enabling national-scale studies and making results available in a standardized way. The Swiss Data Cube is fully committed to Open Science and is covering most of the Open Science facets:[1] supplying *Open Data* as Analysis Ready Data; algorithms as *Open Notebooks* and *Open Source* code; scientific publications in Open Access; and Open Education Resources such as the Bringing Open Data Cube into Practice material[2] (Asmaryan et al., 2019). And now all data and products are compliant with FAIR principles, are stored in Long Term Preservation repository, have DOI, and data and metadata are published according to widely adopted standards. Besides the OGC and ISO standards, the proposed approach supports the Open Archival Information System (OAIS – ISO14721[3]) to ensure long term preservation of data; the Open Archive Initiative Protocol for Metadata Harvesting (OAI-PMH[4]) to facilitate interoperability and exchange of metadata; implements the DataCite Metadata Schema[5] and the Digital Object Identifier (DOI – ISO26324[6]) for identifying resources;

1 *https://www.fosteropenscience.eu/content/what-open-science-introduction*.

2 *https://www.swissdatacube.org/index.php/edsproject/bringing-open-data-cube-into-practice/*.

3 *http://www.oais.info/*.

4 *https://www.openarchives.org/pmh/*.

5 *https://schema.datacite.org/*.

6 *https://www.iso.org/standard/43506.html*.

and finally integrates the Open Researcher and Contributor ID (ORCID[7]) to value research data sets through their visibility and attribution. Ultimately, such an approach allows to benefit from standards all along the data-life cycle from data collection and documentation to curation and preservation facilitating exploitation and sharing of generated products.

Such approach can also reduce the amount of data processing for remote sensing data users having access to ready-to-use products such as land surface reflectance, cloud-free mosaics and time-series estimates of different environmental variables providing consistent, standardized and multi-temporal data to support robust environmental change monitoring at national scale (He et al., 2018).

By enabling interoperable discovery and access to the generated data and making them available to national services (e.g., federal geoportal) contribute to major data sharing initiatives both nationally, like the Swiss public administration's central portal for open government data (*https://opendata.swiss/*), and globally, like the Global Earth Observation System of Systems (GEOSS) (Nativi et al., 2015).

Even if this approach contributes to tackle the challenge of enhanced FAIRness of satellite EO resources made available through EODC, FAIR principles are still not widely adopted in the geoscience community (Stall et al., 2019). There is an undeniable need for development of infrastructures, services, and best practices to enable a systemic change of science practices towards Open and Reproducible Science. Such cultural changes take time but progresses are encouraging, technical solutions exist but the biggest challenges concern organizational and institutional issues (Stall et al., 2019). To increase adoption and endorsement to FAIR principles, and similarly the benefits of EO data from end-users, capacity development and education are essential (Coetzee et al., 2020; Gregory Giuliani et al., 2016; Kganyago & Mhangara, 2019). It will strengthen the reliability and efficiency of scientific research while at the same time enhance the credibility of scientific literature and further stimulate discovery and innovation (Munafò et al., 2017).

With the rapid diffusion, adoption and implementation of EODC concepts and technologies (Killough, 2018), and despite the currently lack of methodology to apply FAIR principles on generated information products, the proposed approach offers a great potential of replicability serving as a guidance/best practices, which can enhance the relevance of EO data to support official statistics. Indeed, EO data have significant potential to provide more timely statistical outputs, to reduce the frequency of surveys, to reduce respondent burden and other costs, to provide data at a more disaggregated level for informed decision making and to provide new statistics and statistical insights. For example, EO data can efficiently and effectively support the monitoring of the Sustainable Development Goals (SDGs) by improving timeliness and relevance of indicators without compromising their impartiality and methodological soundness. (Anderson et al., 2017; United Nations, 2017; Whitcraft et al., 2019).

In term of lessons learned, and besides the benefits listed above, the main challenge was to identify the most suitable software components to build SwissEnvEO. Since the beginning, it has been decided to rely on freely and openly available solutions and to make the design of the architecture as flexible as possible so that users who want to implement a similar approach can easily replicate it. Basically, what is needed is (1) a geospatial metadata catalogue and editor; (2) a geospatial data publishing server, and (3) a digital repository. They only requirement is that they rely on the use of some well-known standards (such as OGC and ISO) to enable effective communication between the components. From the geospatial data management side, GeoNetwork and GeoServer are widely adopted solutions and therefore the choice was somehow easy. Nevertheless, these components can be replaced by either open (e.g., pycsw, CatMDEdit, Mapserver) or proprietary (e.g., ArcGIS) solutions. Regarding the digital repository, solution like Zenodo, Pangaea, or Dryad can also be used instead of the institutional solution used in SwissEnvEO. Another challenge that appeared, once the system was operational, was the need to have proper guidelines to help data producers in publishing and documenting their products. To that, we are currently developing documentation to support not only producers but also users of SwissEnvEO. Consequently, capacity development is an important aspect to take into consideration (Gregory Giuliani et al., 2016).

7 *https://orcid.org/*.

Soon, we plan to generate and make available national ready-to-use products on SDG, Essential Variables, and other data sets of national significance. Indeed, remotely sensed Earth Observations data correspond to unique observations and can be used for other environmental/territorial analyses. Therefore, there is a wide range of reuse opportunities while at the same time getting recognition for its release.

We also plan to implement in SwissEnvEO additional service interfaces and standards to further comply and enhance the dissemination of products generated from the Swiss Data Cube. Of interest, we are considering the SpatioTemporal Asset Catalog (STAC – *https://stacspec.org*), an emerging standard to describe large spatio-temporal EO data sets (Ferreira et al., 2020), as well as implementing the emerging OGC API (*http://ogcapi.ogc.org/*) and supporting the JavaScript Object Notation for Linked Data (JSON-LD) for encoding linked data using JSON (*https://json-ld.org*) to make data discoverable in the Google Dataset Search (*https://datasetsearch.research.google.com*).

Ultimately, the proposed approach, can be seen as a significant contribution to the Open Science Charter (*https://www.unige.ch/openscience/en/open-science/*) of the University of Geneva, the host institution of the Swiss Data Cube. This charter marks the commitment of the University and the academic community to the sharing of scientific knowledge. The Charter affirms the adherence to the principles of open science, which aims to ensure free, streamlined access to scientific publications, to research data, and to the methodologies used to generate that data. The Charter also encourages management and sharing of data by researchers with respect to FAIR principles and access to scientific publications in Open Access. It advocates that efforts by researchers to make their data and publications accessible be considered for research evaluation. The Charter also encourages the university to train and support the academic community in the sharing of scientific knowledge.

5. CONCLUSIONS

Open Science is increasingly advocated for making scientific research more accessible and usable by different categories of users. However, Earth Observations Open Science is still undervalued, and different challenges remain (e.g., socio-cultural, technological, political, organizational, economic, and legal) to achieve the vision of transforming EO data into actionable knowledge by lowering the entry barrier to massive-use Big Earth Data analysis and derived information products. Currently, FAIR-compliant digital repositories cannot fully satisfy the needs of geospatial users while Spatial Data Infrastructures are not yet fully FAIR-compliant and have difficulties in handling Big Earth Data. Moreover, FAIR-compliance is not sufficient for open science. It makes data it easy to find, access and reuse, but free and open data and information policies are essential for open science.

SwissEnvEO aims at contributing to tackle these issues by enhancing traditional SDI with digital repository capabilities for ensuring full compliance with FAIR principles while at the same time benefiting from geospatial services capabilities. It facilitates the publication of EO-based products in accordance with FAIR principles making them "ready-to-use" by many different profiles of users. Indeed, making data, metadata, and algorithms interoperable and as open as possible allows: (1) facilitating the interaction with the Swiss Data Cube from an increasing number of users; (2) connecting results of analysis with other data sets (e.g., in-situ measurements, sensors, socio-economic data); (3) enhancing the data-value chain of generated products; and (4) easing contributions to leading national, regional and/or international data sharing efforts.

Being based on open solutions, the proposed approach is entirely reproducible, helping to efficiently disseminate key environmental information at the national scale. It can facilitate research, foster new collaborations, and contribute to national digital strategies.

To conclude, SwissEnvEO is a step towards more Open and Reproducible EO Science, providing an effective mean to build socially robust, replicable, and reusable knowledge, to generate ready-to-use products supporting evidence-based decisions.

ACKNOWLEDGEMENTS

The views expressed in the paper are those of the authors and do not necessarily reflect the views of the institutions they belong to.

This research was funded by European Commission "Horizon 2020 Program" ERA-PLANET/GEOEssential project, grant number 689443.

Results of this publication are partly or fully relying on the Swiss Data Cube (*http://www.swissdatacube.org*), operated and maintained by UN Environment/GRID-Geneva, the University of Geneva, the University of Zurich and the Swiss Federal Institute for Forest, Snow and Landscape Research WSL.

Finally, we would like to thank Mathieu Vonlanthen and Nicolas Rod from the University of Geneva, Yareta/DLCM Digital Repository for their support in archiving data.

FUNDING INFORMATION

This research was funded by European Commission "Horizon 2020 Program" ERA-PLANET/GEOEssential project, grant number 689443.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Gregory Giuliani *orcid.org/0000-0002-1825-8865*
University of Geneva, Institute for Environmental Sciences, GRID-Geneva, Bd Carl-Vogt 66, CH-1211 Geneva, CH; University of Geneva, Institute for Environmental Sciences, EnviroSPACE lab., 66 Boulevard Carl-Vogt, 1205 Geneva, CH

Hugues Cazeaux *orcid.org/0000-0002-5618-2670*
University of Geneva, Information Systems, 24 rue du Général Dufour, 1211 Geneva 4, CH

Pierre-Yves Burgi *orcid.org/0000-0002-4956-9279*
University of Geneva, Information Systems, 24 rue du Général Dufour, 1211 Geneva 4, CH

Charlotte Poussin
University of Geneva, Institute for Environmental Sciences, GRID-Geneva, Bd Carl-Vogt 66, CH-1211 Geneva, CH

Jean-Philippe Richard *orcid.org/0000-0002-7476-392X*
University of Geneva, Institute for Environmental Sciences, GRID-Geneva, Bd Carl-Vogt 66, CH-1211 Geneva, CH

Bruno Chatenoux *orcid.org/0000-0001-9947-2945*
University of Geneva, Institute for Environmental Sciences, GRID-Geneva, Bd Carl-Vogt 66, CH-1211 Geneva, CH

REFERENCES

Amorim, RC, Castro, JA, Rocha da Silva, J and **Ribeiro, C.** 2017. A comparison of research data management platforms: Architecture, flexible metadata and interoperability. *Universal Access in the Information Society*, 16(4): 851–862. DOI: *https://doi.org/10.1007/s10209-016-0475-y*

Anderson, K, Ryan, B, Sonntag, W, Kavvada, A and **Friedl, L.** 2017. Earth observation in service of the 2030 Agenda for Sustainable Development. *Geo-Spatial Information Science*, 1–20. DOI: *https://doi.org/10.1080/10095020.2017.1333230*

Andries, A, Morse, S, Murphy, RJ, Lynch, J and **Woolliams, ER.** 2019. Seeing Sustainability from Space: Using Earth Observation Data to Populate the UN Sustainable Development Goal Indicators. *Sustainability*, 11(18): 5062. DOI: *https://doi.org/10.3390/su11185062*

Andries, A, Morse, S, Murphy, R, Lynch, J, Woolliams, E and **Fonweban, J.** 2019. Translation of Earth observation data into sustainable development indicators: An analytical framework. *Sustainable Development*, 27(3): 366–376. DOI: *https://doi.org/10.1002/sd.1908*

Appel, M and **Pebesma, E.** 2019. On-Demand Processing of Data Cubes from Satellite Image Collections with the gdalcubes Library. *Data*, 4(3): 92. DOI: *https://doi.org/10.3390/data4030092*

Ariza-Porras, C, Bravo, G, Villamizar, M, Moreno, A, Castro, H, Galindo, G, Cabera, E, Valbuena, S and **Lozano, P.** 2017. *CDCol: A geoscience data cube that meets Colombian needs*, 87–99. DOI: *https://doi.org/10.1007/978-3-319-66562-7_7*

Asmaryan, S, Muradyan, V, Tepanosyan, G, Hovsepyan, A, Saghatelyan, A, Astsatryan, H, Grigoryan, H, Abrahamyan, R, Guigoz, Y and **Giuliani, G.** 2019. Paving the Way towards an Armenian Data Cube. *Data*, 4(3): 117. DOI: *https://doi.org/10.3390/data4030117*

Baumann, P. 2018. Datacube Standards and their Contribution to Analysis-Ready Data. *IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium*, 2051–2053. DOI: *https://doi.org/10.1109/IGARSS.2018.8518994*

Baumann, P. 2019. From Sensor-Centric to User-Centric—When are data Analysis-Ready? *IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium*, 4779–4782. DOI: *https://doi.org/10.1109/IGARSS.2019.8899773*

Baumann, P, Furtado, P, Ritsch, R and **Widmann, N.** 1997. *The RasDaMan approach to multidimensional database management*, 166–173. DOI: *https://doi.org/10.1145/331697.331732*

Baumann, P, Mazzetti, P, Ungar, J, Barbera, R, Barboni, D, Beccati, A, Bigagli, L, Boldrini, E, Bruno, R, Calanducci, A, Campalani, P, Clements, O, Dumitru, A, Grant, M, Herzig, P, Kakaletris, G, Laxton, J, Koltsida, P, Lipskoch, K, Wagner, S, et al. 2016. Big Data Analytics for Earth Sciences: The EarthServer approach. *International Journal of Digital Earth*, 9(1): 3–29. DOI: *https://doi.org/10.1080/17538947.2014.1003106*

Baumann, P, Misev, D, Merticariu, V and **Huu, BP.** 2019. Datacubes: Towards Space/Time Analysis-Ready Data. In *Service-Oriented Mapping*, 269–299. Cham: Springer. DOI: *https://doi.org/10.1007/978-3-319-72434-8_14*

Bermudez, L. 2017. New frontiers on open standards for geo-spatial science. *Geo-Spatial Information Science*, 20(2): 126–133. DOI: *https://doi.org/10.1080/10095020.2017.1325613*

Breuer, A, Janetschek, H and **Malerba, D.** 2019. Translating Sustainable Development Goal (SDG) Interdependencies into Policy Advice. *Sustainability*, 11(7): 2092. DOI: *https://doi.org/10.3390/su11072092*

Camara, G, Queiroz, G, Vinhas, L, Ferreira, K, Cartaxo, R, Simoes, R, Llapa, E, Assis, L and **Sanchez, A.** 2017. *The E-Sensing Architecture for Big Earth Observation Data Analysis*, 48–51. DOI: *https://doi.org/10.2760/383579*

Chatenoux, B, Richard, J-P, Small, D, Roeoesli, C, Windgate, V, Poussin, C, Rodila, D, Peduzzi, P, Steinmeier, C, Ginzler, C, Psomas, A, Schaepman, M and **Giuliani, G.** n.d. The Swiss Data Cube: Analysis Ready Data archive using Earth Observations of Switzerland. *Scientific Data*.

Coetzee, S, Ivánová, I, Mitasova, H and **Brovelli, MA.** 2020. Open Geospatial Software and Data: A Review of the Current State and A Perspective into the Future. *ISPRS International Journal of Geo-Information*, 9(2): 90. DOI: *https://doi.org/10.3390/ijgi9020090*

Curry, PA and **Moosdorf, N.** 2019. An open source web application for distributed geospatial data exploration. *Scientific Data*, 6(1): 190014. DOI: *https://doi.org/10.1038/sdata.2019.14*

Denis, G, Claverie, A, Pasco, X, Darnis, J-P, de Maupeou, B, Lafaye, M and **Morel, E.** 2017. Towards disruptions in Earth observation? New Earth Observation systems and markets evolution: Possible scenarios and impacts. *Acta Astronautica*, 137: 415–433. DOI: *https://doi.org/10.1016/j.actaastro.2017.04.034*

Dhu, T, Dunn, B, Lewis, B, Lymburner, L, Mueller, N, Telfer, E, Lewis, A, McIntyre, A, Minchin, S and **Phillips, C.** 2017. Digital earth Australia – unlocking new value from earth observation data. *Big Earth Data*, 1(1–2): 64–74. DOI: *https://doi.org/10.1080/20964471.2017.1402490*

Dhu, T, Giuliani, G, Juárez, J, Kavvada, A, Killough, B, Merodio, P, Minchin, S and **Ramage, S.** 2019. National Open Data Cubes and Their Contribution to Country-Level Development Policies and Practices. *Data*, 4(4): 144. DOI: *https://doi.org/10.3390/data4040144*

Drusch, M, Del Bello, U, Carlier, S, Colin, O, Fernandez, V, Gascon, F, Hoersch, B, Isola, C, Laberinti, P, Martimort, P, Meygret, A, Spoto, F, Sy, O, Marchese, F and **Bargellini, P.** 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sensing of Environment*, 120: 25–36. DOI: *https://doi.org/10.1016/j.rse.2011.11.026*

Dwyer, J, Roy, D, Sauer, B, Jenkerson, C, Zhang, H and **Lymburner, L.** 2018. *Analysis Ready Data: Enabling Analysis of the Landsat Archive*. DOI: *https://doi.org/10.20944/preprints201808.0029.v1*

Estoque, RC. 2020. A Review of the Sustainability Concept and the State of SDG Monitoring Using Remote Sensing. *Remote Sensing*, 12(11): 1770. DOI: *https://doi.org/10.3390/rs12111770*

European Commission. 2007. *Directive 2007/2/EC of the European Parliament and the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)*, 14.

European Commission. 2018. *The DIAS: User-friendly Access to Copernicus Data and Information*.

Ferrari, T, Scardaci, D and **Andreozzi, S.** 2018. The Open Science Commons for the European Research Area. In *Earth Observation Open Science and Innovation*, 43–67. Cham: Springer. DOI: *https://doi.org/10.1007/978-3-319-65633-5_3*

Ferreira, KR, Queiroz, GR, Vinhas, L, Marujo, RFB, Simoes, REO, Picoli, MCA, Camara, G, Cartaxo, R, Gomes, VCF, Santos, LA, Sanchez, AH, Arcanjo, JS, Fronza, JG, Noronha, CA, Costa, RW, Zaglia, MC, Zioti, F, Korting, TS, Soares, AR, Fonseca, LMG, et al. 2020. Earth Observation Data Cubes for Brazil: Requirements, Methodology and Products. *Remote Sensing*, 12(24): 4033. DOI: *https://doi.org/10.3390/rs12244033*

Fritz, S, See, L, Carlson, T, Haklay, M. (Muki), Oliver, JL, Fraisl, D, Mondardini, R, Brocklehurst, M, Shanley, LA, Schade, S, Wehn, U, Abrate, T, Anstee, J, Arnold, S, Billot, M, Campbell, J, Espey,

J, Gold, M, Hager, G, West, S, et al. 2019. Citizen science and the United Nations Sustainable Development Goals. *Nature Sustainability*, 2(10): 922–930. DOI: *https://doi.org/10.1038/s41893-019-0390-3*

Gao, B. 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3): 257–266. DOI: *https://doi.org/10.1016/S0034-4257(96)00067-3*

GEO secretariat. 2008. *Building a Global Earth Observation System of Systems*, 7.

Giuliani, G, Camara, G, Killough, B and **Minchin, S.** 2019. Earth Observation Open Science: Enhancing Reproducible Science Using Data Cubes. *Data*, 4(4): 147. DOI: *https://doi.org/10.3390/data4040147*

Giuliani, G, Chatenoux, B, Benvenuti, A, Lacroix, P, Santoro, M and **Mazzetti, P.** 2020. Monitoring land degradation at national level using satellite Earth Observation time-series data to support SDG15 – exploring the potential of data cube. *Big Earth Data*, 4(1): 1–20. DOI: *https://doi.org/10.1080/20964471.2020.1711633*

Giuliani, G, Chatenoux, B, Bono, AD, Rodila, D, Richard, J-P, Allenbach, K, Dao, H and **Peduzzi, P.** 2017. Building an Earth Observations Data Cube: Lessons learned from the Swiss Data Cube (SDC) on generating Analysis Ready Data (ARD). *Big Earth Data*, 1(1): 1–18. DOI: *https://doi.org/10.1080/20964471.2017.1398903*

Giuliani, G, Chatenoux, B, Honeck, E and **Richard, J.** 2018. Towards Sentinel-2 Analysis Ready Data: A Swiss Data Cube Perspective. *IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium*, 8659–8662. DOI: *https://doi.org/10.1109/IGARSS.2018.8517954*

Giuliani, G, Dao, H, De Bono, A, Chatenoux, B, Allenbach, K, De Laborie, P, Rodila, D, Alexandris, N and **Peduzzi, P.** 2017. Live Monitoring of Earth Surface (LiMES): A framework for monitoring environmental changes from Earth Observations. *Remote Sensing of Environment*. DOI: *https://doi.org/10.1016/j.rse.2017.05.040*

Giuliani, G, Dubois, A and **Lacroix, P.** 2014. Testing OGC Web Feature and Coverage Service performance: Towards efficient delivery of geospatial data. *Journal of Spatial Information Science*, 7: 1–23. DOI: *https://doi.org/10.5311/JOSIS.2013.7.112*

Giuliani, G, Egger, E, Italiano, J, Poussin, C, Richard, J-P and **Chatenoux, B.** 2020. Essential Variables for Environmental Monitoring: What are the Possible Contributions of Earth Observation Data Cubes? *Data*, 5(4): 100. DOI: *https://doi.org/10.3390/data5040100*

Giuliani, G, Guigoz, Y, Lacroix, P, Ray, N and **Lehmann, A.** 2016. Facilitating the production of ISO-compliant metadata of geospatial datasets. *International Journal of Applied Earth Observation and Geoinformation*, 44: 239–243. DOI: *https://doi.org/10.1016/j.jag.2015.08.010*

Giuliani, G, Lacroix, P, Guigoz, Y, Roncella, R, Bigagli, L, Santoro, M, Mazzetti, P, Nativi, S, Ray, N and **Lehmann, A.** 2016. Bringing GEOSS Services into Practice: A Capacity Building Resource on Spatial Data Infrastructures (SDI). *Transactions in GIS*, 21(4): 811–824. DOI: *https://doi.org/10.1111/tgis.12209*

Giuliani, G, Masó, J, Mazzetti, P, Nativi, S and **Zabala, A.** 2019. Paving the Way to Increased Interoperability of Earth Observations Data Cubes. *Data*, 4(3): 113. DOI: *https://doi.org/10.3390/data4030113*

Giuliani, G, Nativi, S, Obregon, A, Beniston, M and **Lehmann, A.** 2017. Spatially enabling the Global Framework for Climate Services: Reviewing geospatial solutions to efficiently share and integrate climate data & information. *Climate Services*, 8: 44–58. DOI: *https://doi.org/10.1016/j.cliser.2017.08.003*

Giuliani, G and **Peduzzi, P.** 2011. The PREVIEW Global Risk Data Platform: A geoportal to serve and share global data on risk to natural hazards. *Nat. Hazards Earth Syst. Sci.*, 11(1): 53–66. DOI: *https://doi.org/10.5194/nhess-11-53-2011*

Giuliani, G, Ray, N, Schwarzer, S, De Bono, A, Peduzzi, P, Dao, H, Van Woerden, J, Witt, R, Beniston, M and **Lehmann, A.** 2011. Sharing Environmental Data through GEOSS. *International Journal of Applied Geospatial Research*, 2(1): 1–17. DOI: *https://doi.org/10.4018/jagr.2011010101*

Gomes, VCF, Queiroz, GR and **Ferreira, KR.** 2020. An Overview of Platforms for Big Earth Observation Data Management and Analysis. *Remote Sensing*, 12(8): 1253. DOI: *https://doi.org/10.3390/rs12081253*

Gorelick, N, Hancher, M, Dixon, M, Ilyushchenko, S, Thau, D and **Moore, R.** 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202: 18–27. DOI: *https://doi.org/10.1016/j.rse.2017.06.031*

Guo, H. 2017. Big Earth data: A new frontier in Earth and information sciences. *Big Earth Data*, 1(1–2): 4–20. DOI: *https://doi.org/10.1080/20964471.2017.1403062*

Guo, H, Nativi, S, Liang, D, Craglia, M, Wang, L, Schade, S, Corban, C, He, G, Pesaresi, M, Li, J, Shirazi, Z, Liu, J and **Annoni, A.** 2020. Big Earth Data science: An information framework for a sustainable planet. *International Journal of Digital Earth*, 1–25. DOI: *https://doi.org/10.1080/17538947.2020.1743785*

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

He, G, Zhang, Z, Jiao, W, Long, T, Peng, Y, Wang, G, Yin, R, Wang, W, Zhang, X, Liu, H, Cheng, B and **Xiang, B.** 2018. Generation of ready to use (RTU) products over China based on Landsat series data. *Big Earth Data*, 1–9. DOI: *https://doi.org/10.1080/20964471.2018.1433370*

Henricksen, B. 2007. *UNSDI Compendium: A UNSDI Vision, Implementation Strategy, and Reference Architecture*, 150.

Honeck, E, Castello, R, Chatenoux, B, Richard, J-P, Lehmann, A and **Giuliani, G.** 2018. From a Vegetation Index to a Sustainable Development Goal Indicator: Forest Trend Monitoring Using Three Decades of Earth Observations across Switzerland. *ISPRS International Journal of Geo-Information*, 7(12): 455. DOI: *https://doi.org/10.3390/ijgi7120455*

Innerebner, M, Costa, A, Chuprikova, E, Monsorno, R and **Ventura, B.** 2017. Organizing earth observation data inside a spatial data infrastructure. *Earth Science Informatics*, 10(1): 55–68. DOI: *https://doi.org/10.1007/s12145-016-0276-0*

IPBES. 2018. *Summary for policymakers of the thematic assessment of land degradation and restoration*. IPBES secretariat. *https://www.ipbes.net*.

ISO. 2020. *ISO 19165-2:2020 Geographic information—Preservation of digital data and metadata—Part 2: Content specifications for Earth observation data and derived digital products*, 30. ISO. *https://www.iso.org/standard/73810.html*.

Kavvada, A, Metternicht, G, Kerblat, F, Mudau, N, Haldorson, M, Laldaparsad, S, Friedl, L, Held, A and **Chuvieco, E.** 2020. Towards delivering on the sustainable development goals using earth observations. *Remote Sensing of Environment*, 247: 111930. DOI: *https://doi.org/10.1016/j.rse.2020.111930*

Kganyago, M and **Mhangara, P.** 2019. The Role of African Emerging Space Agencies in Earth Observation Capacity Building for Facilitating the Implementation and Monitoring of the African Development Agenda: The Case of African Earth Observation Program. *ISPRS International Journal of Geo-Information*, 8(7): 292. DOI: *https://doi.org/10.3390/ijgi8070292*

Killough, B. 2018. Overview of the Open Data Cube Initiative. *IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium*, 8629–8632. DOI: *https://doi.org/10.1109/IGARSS.2018.8517694*

Kotsev, A, Minghini, M, Tomas, R, Cetl, V and **Lutz, M.** 2020. From Spatial Data Infrastructures to Data Spaces—A Technological Perspective on the Evolution of European SDIs. *ISPRS International Journal of Geo-Information*, 9(3): 176. DOI: *https://doi.org/10.3390/ijgi9030176*

Lewis, A, Lacey, J, Mecklenburg, S, Ross, J, Siqueira, A, Killough, B, Szantoi, Z, Tadono, T, Rosenavist, A, Goryl, P, Miranda, N and **Hosford, S.** 2018. CEOS Analysis Ready Data for Land (CARD4L) Overview. *IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium*, 7407–7410. DOI: *https://doi.org/10.1109/IGARSS.2018.8519255*

Li, J and **Roy, DP.** 2017. A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. *Remote Sensing*, 9(9): 902. DOI: *https://doi.org/10.3390/rs9090902*

Maso, J, Zabala, A, Serral, I and **Pons, X.** 2019. A Portal Offering Standard Visualization and Analysis on top of an Open Data Cube for Sub-National Regions: The Catalan Data Cube Example. *Data*, 4(3): 96. DOI: *https://doi.org/10.3390/data4030096*

Maus, V, Câmara, G, Appel, M and **Pebesma, E.** 2019. dtwSat: Time-Weighted Dynamic Time Warping for Satellite Image Time Series Analysis in R. *Journal of Statistical Software*, 88(1): 1–31. DOI: *https://doi.org/10.18637/jss.v088.i05*

Merodio Gómez, P, Pérez García, M, García Seco, G, Ramírez Santiago, A and **Tapia Johnson, C.** 2019. The Americas' Spatial Data Infrastructure. *ISPRS International Journal of Geo-Information*, 8(10): 432. DOI: *https://doi.org/10.3390/ijgi8100432*

Munafò, MR, Nosek, BA, Bishop, DVM, Button, KS, Chambers, CD, Percie du Sert, N, Simonsohn, U, Wagenmakers, E-J, Ware, JJ and **Ioannidis, JPA.** 2017. A manifesto for reproducible science. *Nature Human Behaviour*, 1(1): 0021. DOI: *https://doi.org/10.1038/s41562-016-0021*

Nativi, S, Mazzetti, P and **Craglia, M.** 2017. A view-based model of data-cube to support big earth data systems interoperability. *Big Earth Data*, 1–25. DOI: *https://doi.org/10.1080/20964471.2017.1404232*

Nativi, S, Mazzetti, P, Santoro, M, Papeschi, F, Craglia, M and **Ochiai, O.** 2015. Big Data challenges in building the Global Earth Observation System of Systems. *Environmental Modelling & Software*, 68: 1–26. DOI: *https://doi.org/10.1016/j.envsoft.2015.01.017*

Nativi, S, Santoro, M, Giuliani, G and **Mazzetti, P.** 2019. Towards a knowledge base to support global change policy goals. *International Journal of Digital Earth*, 1–29. DOI: *https://doi.org/10.1080/17538947.2018.1559367*

Patroumpas, K, Georgomanolis, N, Stratiotis, T, Alexakis, M and **Athanasiou, S.** 2015. Exposing INSPIRE on the Semantic Web. *Journal of Web Semantics*, 35: 53–62. DOI: *https://doi.org/10.1016/j.websem.2015.09.003*

Peng, G, Lacagnina, C, Downs, RR, Ramapriyan, H, Ivánová, I, Ganske, A, Jones, D, Bastin, L, Wyborn, L and **Bastrakova, I.** 2021. *International Community Guidelines for Sharing and Reusing Quality Information of Individual Earth Science Datasets*, 60. DOI: *https://doi.org/10.31219/osf.io/xsu4p*

Poussin, C, Guigoz, Y, Palazzi, E, Terzago, S, Chatenoux, B and **Giuliani, G.** 2019. Snow Cover Evolution in the Gran Paradiso National Park, Italian Alps, Using the Earth Observation Data Cube. *Data*, 4(4): 138. DOI: *https://doi.org/10.3390/data4040138*

Ramage, S. 2011, July. Geospatial standards, data sharing and interoperability. *PoPositionIT*, 36–37.

Rizvi, SR, Killough, B, Cherry, A and **Gowda, S.** 2018. The Ceos Data Cube Portal: A User-Friendly, Open Source Software Solution for the Distribution, Exploration, Analysis, and Visualization of Analysis Ready Data. *IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium*, 8639–8642. DOI: *https://doi.org/10.1109/IGARSS.2018.8518727*

Rouse, JW, Schell, JA, Wand, DD and **Harlan, JC.** 1974. *Monitoring the vernal advancement of retrogradation of natural vegetation* (NASA/GSFC, Type III, Final Report), 1–371.

Rowley, J. 2007. The wisdom hierarchy: Representations of the DIKW hierarchy. *Journal of Information Science*, 33(2): 163–180. DOI: *https://doi.org/10.1177/0165551506070706*

Saito, O, Managi, S, Kanie, N, Kauffman, J and **Takeuchi, K.** 2017. Sustainability science and implementing the sustainable development goals. *Sustainability Science*, 12(6): 907–910. DOI: *https://doi.org/10.1007/s11625-017-0486-5*

Salzano, R, Salvatori, R, Valt, M, Giuliani, G, Chatenoux, B and **Ioppi, L.** 2019. Automated Classification of Terrestrial Images: The Contribution to the Remote Sensing of Snow Cover. *Geosciences*, 9(2): 97. DOI: *https://doi.org/10.3390/geosciences9020097*

Science Business. 2019. *Why Open Science is the Future*, 16. Science Business.

Simonis, I. 2019. Standardization Efforts Across Space Agencies: Applications and Analysis Ready Data Discovery in the Cloud. *IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium*, 4441–4443. DOI: *https://doi.org/10.1109/IGARSS.2019.8899229*

Soille, P, Burger, A, De Marchi, D, Kempeneers, P, Rodriguez, D, Syrris, V and **Vasilev, V.** 2018. A versatile data-intensive computing platform for information retrieval from big geospatial data. *Future Generation Computer Systems*, 81: 30–40. DOI: *https://doi.org/10.1016/j.future.2017.11.007*

Stall, S, Yarmey, L, Cutcher-Gershenfeld, J, Hanson, B, Lehnert, K, Nosek, B, Parsons, M, Robinson, E and **Wyborn, L.** 2019. Make scientific data FAIR. *Nature*, 570(7759): 27. DOI: *https://doi.org/10.1038/d41586-019-01720-7*

Sudmanns, M, Tiede, D, Augustin, H and **Lang, S.** 2019. Assessing global Sentinel-2 coverage dynamics and data availability for operational Earth observation (EO) applications using the EO-Compass. *International Journal of Digital Earth*, 1–17. DOI: *https://doi.org/10.1080/17538947.2019.1572799*

The future we want. 2012. Nations, United 53.

Tondapu, G, Markert, K, Lindquist, EJ, Wiell, D, Díaz, ASP, Johnson, G, Ashmall, W, Chishtie, F, Ate, P, Tenneson, K, Patterson, MS, Ricci, S, Fontanarosa, R and **Saah, D.** 2018. A SERVIR FAO Open Source Partnership: Co-development of Open Source Web Technologies using Earth Observation for Land Cover Mapping. *AGU Fall Meeting Abstracts*, 21. *http://adsabs.harvard.edu/abs/2018AGUFMIN21B..27T*

Truckenbrodt, J, Freemantle, T, Williams, C, Jones, T, Small, D, Dubois, C, Thiel, C, Rossi, C, Syriou, A and **Giuliani, G.** 2019. Towards Sentinel-1 SAR Analysis-Ready Data: A Best Practices Assessment on Preparing Backscatter Data for the Cube. *Data*, 4(3): 93. DOI: *https://doi.org/10.3390/data4030093*

United Nations. 2017. *Earth Observations for Official Statistics*, 170.

Virapongse, A, Pearlman, F, Pearlman, J, Murambadoro, MD, Kuwayama, Y and **Glasscoe, MT.** 2020. Ten rules to increase the societal value of earth observations. *Earth Science Informatics*. DOI: *https://doi.org/10.1007/s12145-020-00453-w*

Wagner, CS and **Jonkers, K.** 2017. Open countries have strong science. *Nature News*, 550(7674): 32. DOI: *https://doi.org/10.1038/550032a*

Whitcraft, AK, Becker-Reshef, I, Justice, CO, Gifford, L, Kavvada, A and **Jarvis, I.** 2019. No pixel left behind: Toward integrating Earth Observations for agriculture into the United Nations Sustainable Development Goals framework. *Remote Sensing of Environment*, 235: 111470. DOI: *https://doi.org/10.1016/j.rse.2019.111470*

Wilkinson, MD, Dumontier, M, Aalbersberg, IjJ, Appleton, G, Axton, M, Baak, A, Blomberg, N, Boiten, J-W, Santos, LB. da S, Bourne, PE, Bouwman, J, Brookes, AJ, Clark, T, Crosas, M, Dillo, I, Dumon, O, Edmunds, S, Evelo, CT, Finkers, R, Mons, B, et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3: 160018. DOI: *https://doi.org/10.1038/sdata.2016.18*

World Economic Forum. 2020. *The Global Risks Report 2020*, 102.

Wulder, MA, Masek, JG, Cohen, WB, Loveland, TR and **Woodcock, CE.** 2012. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, 122: 2–10. DOI: *https://doi.org/10.1016/j.rse.2012.01.010*

Zhu, Z. 2019. Science of Landsat Analysis Ready Data. *Remote Sensing*, 11(18): 2166. DOI: *https://doi.org/10.3390/rs11182166*

TO CITE THIS ARTICLE:
Giuliani, G, Cazeaux, H, Burgi, P-Y, Poussin, C, Richard, J-P and Chatenoux, B. 2021. SwissEnvEO: A FAIR National Environmental Data Repository for Earth Observation Open Science. *Data Science Journal*, 20: 22, pp. 1–18. DOI: *https://doi.org/10.5334/dsj-2021-022*

Submitted: 04 March 2021
Accepted: 15 May 2021
Published: 31 May 2021

Data Science Journal is a peer-reviewed open access journal published by Ubiquity Press.