data

MDPI

Article

Paving the Way to Increased Interoperability of Earth Observations Data Cubes

Gregory Giuliani [1,2,*], Joan Masó [3], Paolo Mazzetti [4], Stefano Nativi [5] and Alaitz Zabala [6]

1   Institute for Environmental Sciences, University of Geneva, enviroSPACE, Bd Carl-Vogt 66,
    CH-1205 Geneva, Switzerland
2   Institute for Environmental Sciences, University of Geneva, GRID-Geneva, Bd Carl-Vogt 66,
    CH-1211 Geneva, Switzerland
3   Center for Ecological Research and Forestry Applications (CREAF), Universitat Autònoma de
    Barcelona (UAB), 08193 Bellaterra, Barcelona, Spain
4   National Research Council of Italy (CNR)—Institute of Atmospheric Pollution Research,
    Via Madonna del Piano 10, 50019 Sesto Fiorentino, Italy
5   European Commission Joint Research Center (JRC), Via E. Fermi, 2749, 21027 Ispra, Italy
6   Geography Department, Universitat Autònoma de Barcelona (UAB), 08193 Bellaterra, Barcelona, Spain
*   Correspondence: gregory.giuliani@unige.ch; Tel.: +41-(0)22-379-07-09

check for
updates

Abstract: Earth observations data cubes (EODCs) are a paradigm transforming the way users interact with large spatio-temporal Earth observation (EO) data. It enhances connections between data, applications and users facilitating management, access and use of analysis ready data (ARD). The ambition is allowing users to harness big EO data at a minimum cost and effort. This significant interest is illustrated by various implementations that exist. The novelty of the approach results in different innovative solutions and the lack of commonly agreed definition of EODC. Consequently, their interoperability has been recognized as a major challenge for the global change and Earth system science domains. The objective of this paper is preventing EODC from becoming silos of information; to present how interoperability can be enabled using widely-adopted geospatial standards; and to contribute to the debate of enhanced interoperability of EODC. We demonstrate how standards can be used, profiled and enriched to pave the way to increased interoperability of EODC and can help delivering and leveraging the power of EO data building, efficient discovery, access and processing services.

Keywords: Open Data Cube; remote sensing; geospatial standards; landsat; sentinel; analysis ready data

1. Introduction

The planet Earth is currently on an unsustainable pathway. Increasing pressures on natural resources induced by human activities are affecting the global environment. Regular and continuous monitoring is necessary to assess, understand, and mitigate these environmental changes [1–3]. Consequently, timely and reliable access to data describing physical, chemical, biological and socio-economic conditions can provide the basis for reliable and accountable scientific understanding and knowledge supporting informed decisions and evidence-based policies [4,5]. This can be done by applying the data-information-knowledge-wisdom (DIKW) paradigm [6,7]. In DIKW, information is an added-value product resulting from the comprehension of available data and their relations with physical and/or social phenomena. In turn, knowledge is generated by understanding information and elaborating on valuable patterns.

Earth Observations (EO) data, acquired remotely by satellite or in-situ by sensors, is a valid and globally consistent source of information and knowledge for monitoring the state of the planet and increasing our understanding of Earth processes [8]. EO data are essential to allow long-term global coverage and to monitor land cover changes over large areas through time [9]. With the increased number of spaceborne sensors, the planet is virtually under continuous monitoring, with satellites providing global coverage at medium-to-high spatial and spectral resolutions on a daily basis [10–12]. Furthermore, open data policies have greatly facilitated the access to satellite data, such as the United States Geological Survey (USGS) Landsat, National Aeronautics and Space Administration (NASA)'s Moderate Resolution Imaging Spectroradiometer (MODIS), the Japan Aerospace Exploration Agency (JAXA), Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), or European Sapce Agency (ESA) Sentinels [13,14]. However, handling such large volumes (e.g., tera to petabytes), variety (e.g., radar, optical), and velocity (e.g., new data available daily), as well as the efforts and costs required to transform EO data into meaningful information have restricted systematic analysis to monitor environmental changes [15]. Consequently, the development of large-scale analytical tools allowing effective and efficient information retrieval based on scientific questions, as well as generating decision-ready products remains a major challenge for the EO community [16].

Earth Observations Data Cubes (EODC) have recently emerged as a paradigm transforming the way users interact with large spatio-temporal EO data [17,18]. It enhances connections between data, applications, and users facilitating management, access and use of analysis ready data (ARD) [19,20]. The ambition is to allow scientists, researchers, and different businesses to harness big EO data at a minimum cost and effort [21]. This significant interest is exemplified by various implementations of platforms capable of analyzing EO data that exist, such as the Open Data Cube (ODC) [19], the EarthServer [22], the e-sensing platform [23], the JRC Earth Observation Data and Processing Platform (JEODPP) [24], the Copernicus Data and Information Access Services (DIAS) [25] or the Google Earth Engine (GEE) [26]. The novelty of the approach results in different innovative solutions and, among them, some can be considered some sort of data cube (even if there is a lack of commonly agreed definition of the EODC term), leading to interoperability issues among them precluding effective discovery, common data access and sharing processes on data stored in EODC [20,27]. Consequently, EODC interoperability has been recognized as a major challenge for the Global Change and Earth System science domains [27].

Therefore, the objectives of this paper are: (1) To better characterize EODC (e.g., differentiate between data cube and cloud-based processing facilities, such as DIAS or Google Earth Engine); (2) issue recommendations to prevent EODC from becoming silos of information; and (3) present/demonstrate how existing geospatial standards can be profiled and enriched to pave the way to increased syntactic and semantic interoperability, as well as addressing use and orchestration of EODC and can help the delivering and leveraging the power of EO data in building efficient discovery, access and processing services.

2. Earth Observation Data Cube and Analysis Ready Data Infrastructures

To better characterize EODC, six different aspects have been identified covering respective well-established data science domains, allowing to describe EODC into meaningful and manageable parts with the ultimate objective to ensure compatibility and consistency for efficient data discovery, view, access and processing [20,27].

The "faces" concept was then further elaborated, leading to the definition of six viewpoints [27] characterizing a data cube infrastructure: (1) The semantic view, covering the information stored in the content and their semantics; (2) the geometric view, covering the content in geometrical representation, in particular their discretization and digital structuring aspects; (3) the encoding view, dealing with the multi-dimension aspects, including pre-processing and analytical processing aspects; (4) the interaction/interface view, dealing with the analytical functionalities provided by the infrastructure and their accessibility via web-based Application Programming Interfaces (APIs).

(5) Interconnection/platform view, dealing with the software components and services necessary to realize the cybernetics framework; (6) the composition/ecosystem view, concerning the infrastructure composability with analogous systems and governance aspects.

To enable and facilitate full interoperability of EODC, as well as leveraging the rich legacy of Business Intelligence, it is important to make sure all the views are adequately addressed and kept technology-neutral [27]. To achieve this, a crucial action consists of the identification of existing and mature models and patterns promoting the adoption of standard approaches.

To better characterize EODC, it is important to differentiate them from cloud-based processing facilities, such as DIAS, GEE or Earth on Amazon Web Services (AWS). Cloud-based EO platforms commonly provide (free and open) access to global EO datasets (available datasets are growing daily) along with powerful space and time analysis tools supporting different programming languages (e.g., JavaScript, Python and R). Recently, these online platforms have transformed the user community working with satellite EO data. They removed most of the burden for data preparation, yielding rapid results and fostering a community of contributors, which is growing fast. However, they lock users into a platform (sort of commercial) dependency, with well-known challenges. Potential identified concerns are: (1) Users do not know whether a given platform will be sustained and/or evolved in the future; (2) the provision of limited time and spatial scale for analyses; (3) the provision of cloud-based computing only (i.e., no options for hubs or local computing solutions); (4) users are requested to upload their analytical processing and even local data, while data download is discouraged or not even allowed; (5) platform providers require the right to "own" all the data utilized on the platform; (6) users get only those datasets that providers offer, limiting data interoperability (e.g., Landsat 8 or Sentinel 1 data can be missing); (7) data are often not ready to be analyzed (e.g., top-of-the-atmosphere—TOA—reflectance data).

Most of these potential drawbacks can be tackled by utilizing EODC. For example, users can install on their own computing infrastructure an open source software solution (such as the ODC), that allows for storing different type of data (e.g., Landsat, Sentinel, SPOT, MODIS, aerial and/or drone imagery, etc.). This solution provides improved control, more flexibility and scalability, both in terms of usage, and a further sense of ownership. EODC support an efficient and joint use of multiple datasets, enhancing their interoperability and complementarity. This facilitates not only data sharing but also the sharing of code, tools and algorithms. Finally, it grants the possibility to develop local and/or regional solutions that avoid commercial and internet dependence. For these reasons, more and more cloud-based EO data infrastructures are considering offering EODC services to their potential users.

3. Current Interoperability Levels of EODC

3.1. Software Systems Interoperability

Interoperability was first defined by the Institute of Electrical and Electronics Engineers (IEEE) as "the ability of two or more systems or components to exchange information and to use the information that has been exchanged" [28]. In the present digital transformation era, interoperability is a critical software system attribute, since it enables different systems interaction to support the society daily activities. The emerging technologies composing systems-of-systems have increased their importance and scope. Interoperability can be thought as the ability of software systems to interact for a specific purpose, once their differences (development platforms, data formats, culture, and legal issues) have been overcome [29]. Interoperability is not a clear-cut characteristic of a system; there exist different levels (or types) of interoperability, spanning from system integrability (including technical and syntactic interoperability levels) to system composability (including semantic, dynamic and conceptual interoperability levels). Commonly, system interoperability is achieved by pushing open standards—either de-facto or de-jure.

For the scope of this manuscript, interoperability may be defined as the ability of different data cube infrastructures to connect and communicate in a coordinated way, providing a rich experience to

users. For example, interoperable data cubes should provide the necessary functions to allow users to access and analyze data with (virtually) no effort, regardless of data origin, structure and encoding.

3.2. Interoperability Contexts for EODC Software Systems

Earth Observation Data Cubes are an infrastructure managing (long) time series of observations referring to the Earth; i.e., characterized by a spatial reference system. To provide users with a rich experience, EODC interoperability must be considered in respect of the context where their use is planned. For a given EODC, it is possible to consider three different and increasing (from the composability point of view) levels of interoperability:

1. Interoperability among EODCs.
2. Interoperability of EODCs with other types of geospatial data cubes.
3. Interoperability of EODCs with general-purpose data cube infrastructures.

All the three interoperability contexts are facilitated by the past and present activities on the harmonization and mediation of EO information; i.e., the standardization process.

3.2.1. Interoperability among EODCs

For the geometric view, when characterizing an EODC [27], spatio-temporal coverages (ISO 19123, 2005) are largely recognized as the referential representation for observation of physical phenomena. Therefore, it is the general consensus on building cubes with a spatio-temporal domain. Indeed, this is an effective cube geometry for fast generation of a time-series, which is one of the most commonly used cases. However, it is worth noting that there is still heterogeneity in a number of domain dimensions (2D, 2D+T, 3D, 3D+T, and 3D+T). Besides, some commonly used cases, such as simulations, would actually need more than one temporal dimension as part of the domain. Therefore, at this level, a major challenge in EODCs interoperability concerns the harmonization of domain dimensions among different data cube implementations. Another relevant geometrical aspect is related to the metrics that is superimposed to a given data cube. In particular, this includes the coordinate reference system adopted. Harmonization of data from data cubes with significantly different spatio-temporal reference systems would require a lot of computation. This might void the processing assets stemming from the use of data cube infrastructures in respect to other (more traditional) data services.

Concerning the EODC semantic view, interoperability can leverage the on-going activities by the communities of practice, in the EO and Earth science domain, to define a set of essential variables [30–32] and variable name conventions [33]. However, semantic interoperability must be seen as ancillary to the more important pragmatic interoperability, which is the real requirement from users. data cubes are designed for efficient processing in support of specific cases of use, thus, pragmatic aspects (e.g., data resolution and fitness-for-purpose) should be considered as relevant as semantic ones. The (long) ingestion time required for efficient computation of time series may be frustrated by a time-consuming pre-processing to make data usable for a specific use-case. To work out pragmatic interoperability, the aspects related to moving from a data to an ARD system should be considered; e.g., pixel alignment, atmospheric correction.

3.2.2. Interoperability of EODCs with Other Types of Geospatial Data Cubes

Moving interoperability to the more general level of geospatial data cubes, interoperability issues increase. With the term "geospatial data cube," we consider data cubes that encode information, characterized by a spatio-temporal content, which may be represented as not making use of the coverage model. Actually, this is a common situation, in particular when socio-economic information is provided as aggregated at local, regional and national level—or with reference to any administrative boundary, in general. For example, a data cube may report a set of parameters (e.g., GDP, school enrollment, and life expectancy) by country. In this case, at least one of the dimensions (i.e., the country) has a spatial content, but it is expressed as a geographical feature and not as a coverage function.

There are also more complex cases where the geographical feature is not a dimension, such as features changing over time; e.g. the area affected by a flood, the set of protected areas in Europe, or the set of countries that are members of the United Nations. In these cases, interoperability from the geometry point of view can become very complex. Of course, the semantic viewpoint also highlights a higher complexity, since providing common semantics for different domains is still an open issue—one of multi-disciplinary interoperability.

3.2.3. Interoperability of EODCs with General-Purpose Data Cubes

Interoperability with general-purpose data cubes, where information has no explicit geographical content is even more complex. This is also a common situation in socio-economic contexts, where data (in particular statistical data) are aggregated according to non-spatio-temporal dimensions–e.g., life expectancy by job category, wealth and income by age category, etc.

In principle, most of the interoperability issues that interest data cubes have been already recognized and largely addressed by the science studying the interoperability of geospatial information systems. For that reason, some of these issues can be solved by adopting the existing standards or mediation tools.

However, from an engineering point of view, it useful to reflect on the peculiarity aspects of ARD and data cube systems: Their diversity in respect to a traditional data/information system. In particular, it is important to consider that making data cubes interoperable does not mean building a virtual data cube—like we commonly do implementing data systems federation. By simply making data cubes interoperable, it would build an information system that accesses data cubes, but that it is not necessarily a data cube itself. Data cubes are intended as systems tailored to (optimized for) specific-use cases. They were conceived to implement ARD systems. Therefore, they are required to implement interoperability at the pragmatic level. Different data cubes may be "ready" for different uses, and putting them together would likely result in a system that is not necessarily ready for a commonly defined purpose.

4. Enhancing Interoperability Using Standards

4.1. Stakeholders and Patterns

To cover all the six interoperability views, defined by Nativi et al. [27], different stakeholders must be engaged, including disciplinary experts (e.g., experts on Earth system, geospatial information, multidimensional data management, online analytical processing, HPC, and ecosystems), standardization organizations, and the users (e.g., business intelligence association, and policy makers) who must provide the use cases to be addressed by data cubes.

In developing interoperability solutions, well-accepted and innovative patterns must be considered. For example:

- Semantic interoperability:

  ○ Data and information typing specifications.
  ○ Semantic and ontological languages to be used along to enrich and disambiguate content metadata.
  ○ Co-design patterns.

- Geometry interoperability:

  ○ Geospatial information models.
  ○ Business intelligence and the online analytical processing multidimensional modeling.

- Encoding interoperability:

  ○ Well accepted file systems and formats patterns.

- ○ Multidimensional DB.
- ○ Big data tiling strategies.

- Interaction/interface interoperability:

  - ○ Web APIs.
  - ○ Online analytical processing (OLAP) APIs.
  - ○ Web Notebook tools.
  - ○ Well-adopted interoperability protocols (e.g. OGC, W3C, IET).

- Interconnection/platform interoperability:

  - ○ System-of-Systems (SoS) patterns.
  - ○ Software design patterns.
  - ○ Cloud computing interoperability patterns.

- Composition/ecosystems interoperability:

  - ○ Software Ecosystem (SECO) patterns.
  - ○ SoS virtual/collaborative architectures.
  - ○ SoS governance styles (e.g., directed, collaborative, acknowledged, virtual).

4.2. Documenting Data and Data Discovery

In the DIKW paradigm, the first step in the data value chain (e.g., a set of actions from data capture to decision-ready products) is known as data discovery [34]. It allows users to search, find, and evaluate suitable data that will be further used in models or other analytical workflows. Data discovery is realized through catalogs containing relevant information describing datasets (e.g., spatial resolution, spatial extent, temporal resolution) [35]. These detailed descriptions are commonly referred to as metadata [36–38]. To contribute to initiatives such as the Global Earth Observation System of Systems (GEOSS), it is required to use data description specifications (i.e., metadata standards) to document datasets and store metadata in interoperable catalogs to facilitate exchange and use by various systems [39].

Different open standard schemas have been developed to describe geospatial data [40]. The most widely used standards are developed by the International Organization for Standardization (ISO)/Technical Committee (TC), 211 Geographic Information/Geomatics, and the Open Geospatial Consortium (OGC). They concern data and service description (ISO19115-1 and ISO19119), their respective schema implementation (ISO19139-1), and the Catalog Service for the Web Interface (OGC CSW) [41,42]. With this suite of standards, users can adequately document data and provide standardized search, discovery, and query capabilities over the internet [43].

Currently, the vast majority of metadata catalogs relate to geographical data (e.g., map agency's products) and only a few of them concern EO data [44]. EO products are normally distributed by the data producers as scenes or granules (a spatial fragment of a satellite path) of data with a metadata document for each scene. In this partition of a product in space and time, most of the metadata content is identical and is repeated in each scene. That is the reason why a catalogue interface holding these metadata records will generate hundreds or thousands of hits for a thematic query. A hierarchical structure of metadata describing a product as a single unit that has multiple scenes needs to be adopted to make the catalogues useful [45]. Moreover, among the various data cube implementations, the Open Data Cube and RasDaMan/EarthServer are the most widely adopted solutions [19,22]. They arrange data in a hierarchical way, and expose data at the product level, making it visible as a single entity, but they lack metadata description and catalog interfaces impeding efficient and effective discovery mechanism.

To tackle this issue and store relevant metadata information about satellite data (e.g., acquisition, sensing, and bands) in a online metadata catalog, the XML schema ISO 19139-2 extends the original metadata schema to support additional aspects more significant for the gridded and imagery information defined in ISO19115-2, and offers an interesting possible solution [44]. Additionally, the SpatioTemporal Asset Catalog (STAC, https://stacspec.org) is under substantial development. It is an emerging metadata standard, primarily designed for remote sensing imagery, aiming at standardizing the way geospatial assets are exposed online and queried. Interestingly, there are preliminary efforts to extend ODC to use STAC files as a source of information to index data (https://github.com/radiantearth/stac-spec/tree/master/extensions/datacube).

4.3. Data Quality and Uncertainty

Nowadays, data cubes barely have data quality information in their metadata records. A couple of complimentary approaches to populate this lack of information have to be considered.

First of all, the uncertainty associated to each image can be estimated by several means. On one hand, several papers (such as [46–49]) assess the general accuracy for certain instruments, based mainly on Calibration and Validation (Cal/Val) campaigns or invariant areas. These approaches give a general uncertainty value associated to each sensor and band so that can be applied to a single product for the product-level metadata. Another refinement that can be done in this direction is to consider that this error is modulated according to the incidence angle, as the bigger this angle, the more specular effects in reflectance, and thus the higher the errors expected. Considering that effect, a different uncertainty value can be associated to each pixel of the scene; i.e. obtaining an image describing the uncertainty for each pixel of a certain image. Once this information is available, the uncertainty in any analytical operation using the imagery can be computed by propagating the original uncertainties to the final product using classical error propagation formulas and map algebra. For example, as the Normalized Difference Vegetation Index (NDVI) is computed as:

$$NDVI = \frac{IR + R}{IR - R} = \frac{M}{N}$$

(added and subtracted variables, and then divided). The uncertainty of each pixel can be computed using the two following formulas:

$$S_{NDVI} = \sqrt{(S_M^2 \times N^2 + S_N^2 \times M^2)}$$

$$S_{M \text{ and } N \text{ } (IR \pm R)} = \sqrt{(S_{IR}^2 + S_R^2)}$$

As this uncertainty propagation can be computed with map algebra, Open Data Cubes should be able to include these calculations in their routines, and thus, automatically generate the final products as well as their uncertainty.

The second approach is to assess the quality of the final product not by error propagation but by comparing the product to a known ground truth, thus validating its thematic accuracy. This is the most common approach for final products where uncertainty propagation is not directly applicable, such as land use land cover, or leaf area index (LAI) area, in which the product is validated against known values obtained by field work or other means. In this approach, the obtained quality assessment is generally documented at the dataset level (not at pixel level).

Regarding how to include data quality (both product and pixel level) in metadata, the widely selected standard is the ISO19157. ISO19157 identifies the conceptual model for describing quality elements in a geographic dataset and defines several quality elements describing different quality aspects in a dataset: Completeness, logical consistency, positional accuracy, thematic accuracy, temporal quality, usability and metaquality. Besides ISO19157, the Quality Mark-up Language (http://www.qualityml.org) is both a vocabulary and an encoding for data quality that was originally

developed in GeoViQua FP7 project and extended in OGC TestBed12. This vocabulary proposes a clear encoding of quality elements (using standardized quality measures) in XML metadata documents, and can be used for describing the quality of the original products, as well as to define further quality evaluations carried out over the datasets [50].

4.4. Data Visualization and Download

Data cubes are essentially analytical frameworks. The multidimensional nature of the data cube makes it difficult to visualize, but a standard solution can be found, allowing time series to be stored in a data cube for efficient analysis, while at the same time being able to be easily visualized. Moreover, the benefit of using the data cube as the origin to create data visualizations, is that it allows creating visualizations with any combination of the data cube dimensions; for example, it can be applied to extract 2D imagery at low resolution for WMS visualization of a time series evolution, as well as 1D time profile diagrams, or x/t time slices with WCS ready to visualize. Most data cubes have data organized in ways that are optimal for these kind of operations, allowing on the fly fast visualization.

The most common way of doing optimizing as such, is by reducing the number of dimensions to less than three. Standard data visualization is generally achieved using geospatial web services, as they are particularly suitable for this purpose [51]. In that sense, OGC Web Map Service (WMS) and Web Map Tiled Service (WMTS) are common interoperable solutions to show maps from different origins together in a single image [52]. WMS and WMTS are particularly fit for accommodating multidimensional data cubes, due to their capability to define extra dimensions (on top of a two-dimensional CRS). A time parameter is defined in the WMS standards, and can be added as an extra dimension in WMTS. An additional, band parameter, can also be defined by the server to select among several bands of a EO product. This could be very convenient for an optical product (e.g., Sentinel 2), since it can be represented as a single layer and the band parameter is used to select the spectral band, while the time parameter can be used to extract a time slice of the time series. In both solutions, the server executes an internal algorithm that reduces the dimensions to the two spatial ones (on a certain CRS) and creates a 2D portrayal by applying a color palette, or by combining bands into RGB combinations. Both are characterized by requesting a simplification of the original data at a particular scale, time slice, and in a format that is easy to display. Depending on the scale requested (also known as zoom level), a small portion of the data are requested at near full resolution, or a large piece of information is requested at low resolution. Normally the client is not receiving the actual values of the datasets (e.g., EO data), only naive visualization where only colors are encoded in JPEG or PNG formats. In WMS, in trying to make the time parameter flexible, the authors introduced some variants (such as the possibility of indicating time intervals, or the nearest values) that are a bit vague and complex to implement. In that sense, the precision of the time values may introduce uncertainties in requests and responses. To prevent that, we recommend that the service enumerates the time values available in the capabilities document, and the client use them as literals and only indicated time instants in the time parameter (this recommendation coincides with the only possible use for a time extra dimension in WMTS). Another fundamental problem in WMS is that the semantics of the time parameter is unclear and could refer to the acquisition time, processing time, publication time, etc. The OGC WMS Earth Observation profile recommends using the time parameter only for the acquisition time [53]. It is always possible to create other time dimensions with other semantics if needed. We believe that by following these recommendations, the WMS time dimension is usable and modern clients are capable to overcome these restrictions by offering smarter user interfaces that, e.g., present time arrows or combine individual WMS requests in animations. A couple of previous attempts to improve the situation have been developed. On one side, ncWMS introduced several extensions in GetFeatureInfo, symbolization and vertical and temporal dimensions [54]. On the other side, a simple extension of WMS adding a binary raw data format and the move of some of the portrayal capabilities from the server to the client has also been proposed [55]. The latter option is used in the Catalan Data Cube, adding to the client much more than simple data visualization (including time profiles that take advantage of the having the data

cube as the origin of the visualization), such as simple statistical calculations or pixel-based operations among layers from different sources [56].

Is it possible to build a middleware that interprets a WMS or a WMTS request, extracts the extra dimensions and translates it into specific data cube operations for cutting a re-sampling. If the response needs to be produced on the fly, it shall be fast in both extreme circumstances. Generally, fast extracting fragments from data cubes are achieved by saving the data, pre-sampled on different scales.

Regarding data download, the Web Coverage Service (WCS) standard is the right way to go. WCS 2.0 is based on the Coverage Implementation Schema (CIS) that describes the data dimensions as well as the thematic bands offered by the service in a standard language. With WCS, we can easily filter the data and extract a fragment of it, in the same number of dimensions (sub-setting) or in a reduced number (slicing). The response can be one or more common data formats (e.g., a GeoTIFF or a NetCDF). Actually, the approach could be used as standard way of exchanging products among data cubes. A data cube could act as a WCS client and progressively request the data from another data cube that implements WCS services. Moreover, a data cube that has been updated with new data could use WCS-T (transactional WCS) to automatically update other data cubes that have the same product with new time slices or completely new products.

The Web Coverage Service (WCS) standard can also be used for visualization. With the help of a modern JavaScript library (e.g., https://geotiffjs.github.io/) GeoTIFF images can also be seen in the map browser directly. Another interesting application of WCS is the extraction of a temporal profile of a point by reducing the number dimensions to one—the time. By doing so, a GetCoverage request is capable of responding to a time series (e.g., timeseriesML) that can be presented in a map client as a diagram.

4.5. Data Processing

Data cubes being analytical frameworks, computing and IT infrastructures are significant elements to enhance data flow, data transformation to information, as well as analysis and processing of the ever-increasing volume of EO data that exceed the current capacities of existing computers [12,16,57,58].

To deal with these issues of processing algorithms, sharing, and computing power, the OGC Web Processing Service (WPS) [59] and Web Coverage Processing Service (WCPS) [60], together with the high-performance and distributed computing paradigm, can be beneficial [61–63].

The WPS standard specification defines how to invoke input and output data as a web-based processing service. It provided rules on how a user/client can ask for a process execution, how a provider should publish a given processing algorithm as a service, and how inputs and outputs are managed [64]. This standard facilitates algorithms sharing in an interoperable fashion. However, it is not well adapted for raster analysis [60]. Consequently, the Web Coverage Processing Service (WCPS) specification has emerged defining a query language for processing multi-dimensional raster coverages [65]. Those two standards provide effective foundations to enhance interoperability of EO data cubes and ensure that when a user send the same request and processing algorithm can be executed on different data cubes.

Finally, to efficiently turn EO data into information and knowledge, effective processing solutions are necessary. Distributed and high-performance computing (HPC) infrastructure like clusters, grids, or clouds are adequate solutions [62,66,67]. It is now possible to benefit from the computing power of these infrastructures while using interoperable processing services in a transparent manner, hiding the complexity of these infrastructure to users [68–70]. Such integration can help leveraging the capabilities of these infrastructures and support model-as-a-Service approaches, such as the GEO Model Web [71] or data cubes [18].

4.6. Data Reproducibility

Provenance also includes the description of the algorithms used, their inputs and outputs, the computing environment where the process runs, the organization/person responsible for the

product, etc. [72]. A provenance record consists of a format list of processes and data sources used to create a derived product. By documenting provenance in the metadata, traceability of scientific results is possible, and the same result can be reproduced in the same or in another environment. The new revision of 19115-2 proposes a model to record all inputs necessary to execute an analytic processes, as well as to describe the process itself. This approach is fully compatible with the WPS standard proposed in the previous section [73]. At the moment, to execute any process in the data cube, the executing environment knows everything about the job requested, and it is in the perfect position to save this information in the metadata. The data cube environment should facilitate this recording task and integrate it in the processing operations without any user intervention. If data cubes record provenance information in the same way, it could be possible use the provenance information of a result produced in one environment and reproduce the same result in another environment.

4.7. Data Integration, Semantics and Value Chains

The analysis ready data updated in real time can be the basis for elaborating Essential Climate Variables (ECV) [74]. The generalization of the essential variable concept to other areas, such us biodiversity (EBV), offers new opportunities to monitor the biodiversity, the ecosystems, and other sectors; and opens the door to generate policy related indicators (e.g., Aichi targets or the Sustainable Development Goals (SDG)) [4,32]. Many of the essential variables that can be extracted from remote sensing as high processing level products are indeed describing ECVs [75] or EBVs [76]. A semantic view is necessary for an effective usability and interoperability of data cube products [27]. Connecting data cube high level products to structured keyword dictionaries and formal ontologies is necessary. Tagging the data cube products with the essential variables concept provides a degree of formal semantics to a well-defined and accepted set of measurable class names. The adoption of a formal and common vocabulary of essential variables (EVs) can facilitate the discovery of the relevant data for a particular application. Additional metadata can make the data usable by providing concrete information on spatial resolution, periodicity, and units of measure. This information should be included in the product description of the data cube in the form of ISO 19115 keywords, or by linking to formal ontologies encoded in RDF or OWL in the Internet.

4.8. Data Ingestion from Data Cubes and from Data Providers

Currently, each remote sensing data provider is serving data in a different way. Assuming that the product we want to download is made available for free, commonly, we face an ingestion process that has two phases: Discovery of new scenes by formulating a spatiotemporal query and retrieving the individual scenes. Most data providers offer a visual interface to find the relevant data, which is good for retrieving some samples, but it is not useful to make the data ingestion process in a data cube fully automatic. Some providers offer a Web API to discover and retrieve the data. In the discovery phase, an HTTP GET or a POST request containing a spatiotemporal query and some extra parameters (e.g., the maximum cloud coverage allowed) is submitted to the server. Often the result is a file with a list of hits that includes the names of the scenes that comply with the requested constraints. The client needs to explore this file and formulate more requests; one for each scene to finally get the wanted product subset. To make the situation more complicated, the number of hits of the discovery response might be limited to a maximum number and download will only be possible once authenticated in the system. An ad-hoc script will be necessary to make the requests automatic. Assuming that we were able to get the required data, that is only the start, because there are significant differences among the composition of the scenes that a single data provider offers in terms of metadata content and data formats. There is a need for a more standardized way to document the structure of a remote sensing product and to agree on a standardized format for scene distribution that can be used to ingest the data automatically. Since this is not the case today, every product will require metadata transpositions and format transformations.

The combination of WCS and the CIS standards could open the door to define a standard profile to discover and retrieve the necessary scenes. CIS 2.0 incorporates an extended model that allows for a partition data structure that complies with the requirements of a classical remote sensing distribution in scenes or granules. WCS can incorporate security to authenticate users and will allow for formulating a single request to discover the necessary partitions that will be downloaded in a second phase.

Such approaches could be adopted not only by the remote sensing data providers but also by the data cubes, opening the door for having a protocol that a data cube could apply to harvest any other data cube has implemented the standard approach, something that is not currently possible forcing people to having to develop adaptors.

5. Examples of Enhanced Interoperability from the Swiss Data Cube and the Catalan Data Cube

5.1. Swiss Data Cube Discovery, View, Download and Process Services

To fully benefit from freely and openly available Landsat and Copernicus data archives for national environmental monitoring purposes, the Swiss Federal Office for the Environment (FOEN) is supporting the development of the Swiss Data Cube (SDC—http://www.swissdatacube.ch). The SDC is currently being developed, implemented and operated by the United Nations 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). The objectives of the SDC are twofold. First, to support the Swiss government for environmental monitoring and reporting; and second, to enable Swiss scientific institutions to take advantage of satellite EO data for research and innovation.

The SDC is built on the Open Data Cube software suite [19], and currently holds 35 years of Landsat 5, 7, and 8 (1984–2019), 5 years of Sentinel-1 (2014–2019), and 4 years of Sentinel-2 (2015–2019) of ARD over Switzerland [77,78]. This archive is updated on a daily basis with the most recent data and contains approximately 10,000 scenes, accounting for a total volume of 6 TB, and more than 200 billion observations nationwide.

Currently, one of the key challenges that SDC has to tackle to ensure its scalability, is enhancing the interoperability. Indeed, making data, metadata, and algorithms interoperable will: (1) Facilitate the interaction with the SDC from an increasing number of users; (2) allow connecting results of analysis with other datasets; (3) enhance the data value chain; and (4) ease contributions to major regional and/or international data sharing efforts, such as GEOSS.

Initial interoperability arrangements are currently under development. In the SDC, we decided to distinguish between upstream and downstream services [79,80]. The upstream tier relates to services to interact with the infrastructure (e.g., processing, view, download) while the downstream tier allows users interacting with decision-ready/value-added products. Both tiers are implementing widely adopted open standards for modeling and implementing geospatial information interoperability advanced by the OGC and ISO/TC211.

Regarding the upstream tier, the following strategy for implementing standards is being adopted:

- *Discovery:* SDC description is being done using the ISO19115-2 and ISO19139-2 standards to support gridded and imagery information. The XML schema has been deployed and exposed using the GeoNetwork metadata catalog to store all relevant information to adequately describe the SDC content (e.g., sensors, spatial resolution, temporal resolution, spectral bands). The schema plugin has been downloaded from the GeoNetwork GitHub repository: https://github.com/geonetwork/schema-plugins. Moreover, the GeoNetwork catalog allows exposing an OGC CSW interface for publishing metadata records and allowing users to query the catalog content.
- *View and Download:* To leverage the content of the SDC for visualization and download, respectively OGC WMS and WCS are under implementation. The datacube-ows component (available at: https://github.com/opendatacube/datacube-ows) implements the WMS and WCS standards allowing an interoperable access to Landsat and Sentinels data.

- *Process:* To expose analytical functionalities (e.g., algorithms) developed in the SDC using the ODC Python Application Programming Interface (API), it has been decided as using a PyWPS implementation (https://pywps.org). The main advantage is that it is also written in Python, and allows easy to expose, dedicated Python scripts as interoperable WPS services. That approach is currently under implementation and testing.

Concerning the downstream tier, it has been separated from the SDC for the reason that only final products (e.g., validated analysis results) are concerned. That facilitates the publication and sharing of good quality results through value-added/decision-ready products, while at the same time separating the usage of the Swiss Data Cube between scientific/data analysts end-users and more general end-users.

To that, a specific GeoServer instance with dedicated EO extensions and time support has been implemented. It allows users to efficiently interact with multi-dimensional (e.g., space and time), gridded, and image products generated with the SDC. It currently supports:

Discovery services

- CSW 2.0.2.
- OpeanSearch EO 1.0.

View services

- Web Map Service (WMS) with EO extension 1.1.1/1.3.0.
- Web Map Tile Service (WMTS) 1.0.0.
- Tile Map Service (TMS) 1.0.0.
- Web Map Tile Cached (WMS-C) 1.1.1.

Download services

- Web Coverage Service (WCS) with EO extension 1.0.0/1.1.0/1.1/1.1.1/2.0.1.
- Web Feature Service (WFS) 1.0.0/1.1.0/2.0.0.

To further ease user's interaction with SDC products, a web-based application called the Swiss Data Cube Viewer (http://www.swissdatacube.org/viewer/) has been developed (Figure 1). It allows visualizing, querying, and downloading time-series data generated with the SDC. This JavaScript application provides a simple, responsive template for building web mapping applications with Bootstrap, Leaflet, and typeahead.js. It provides the following functionalities:

- Visualizing and Downloading single raster product layers;
- Visualizing and Downloading time-series raster product layers;
- Generating a graph for a given pixel of a time-series raster product layer;
- Access data products in users 'client with WMS and WCS standards;
- Metadata support.

The application is entirely open-source, and the code can be download at: https://github.com/GRIDgva/SwissDataCube/tree/master/viewer

The use of OGC and ISO standards can enhance syntactic interoperability of the SDC and can help delivering and leveraging the power of EO data building efficient discovery, access and processing services provided by the SDC.
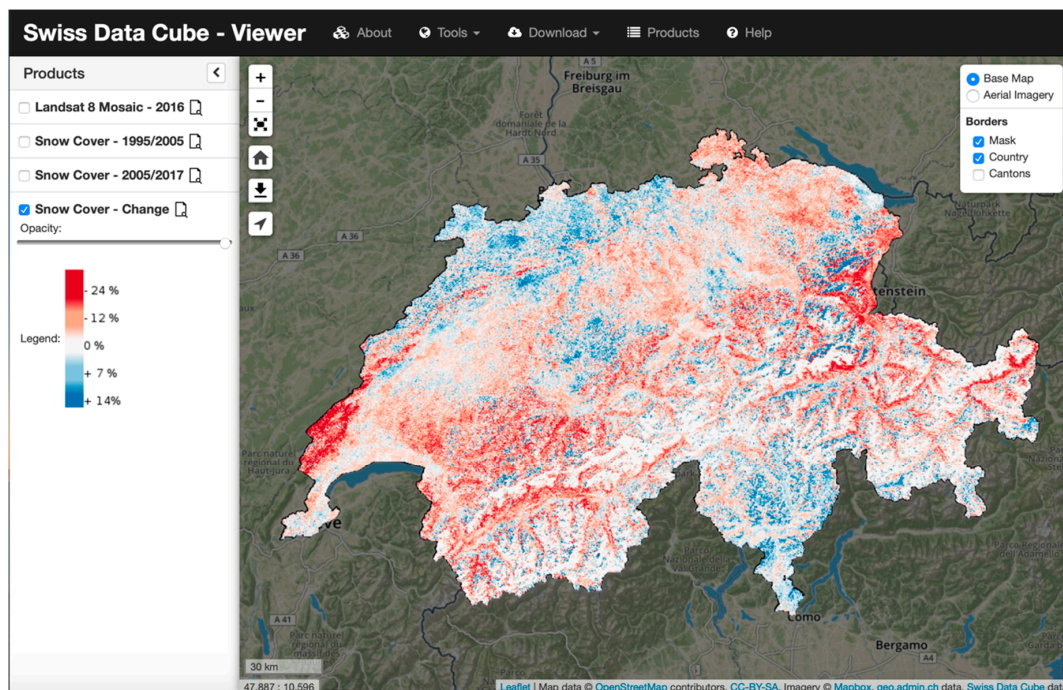
Figure 1. The Swiss Data Cube; viewer showing snow cover change over Switzerland between 1995 and 2017.

5.2. The Catalan Data Cube View Service with Analytical Features

The Department of Environment of the Catalan Government and Centre de Recerca Ecològica i Aplicacions Forestal (CREAF) created the SatCat data portal (http://www.opengis.uab.cat/wms/satcat) that organized the historical Landsat archive (from years 1972 to 2017) over Catalonia in a single portal, providing visualization and download functionalities based on OGC international standards [81]. The initiative is still up and running, plus continuously and manually updated; but a considerable amount of processing work is needed to keep the portal up-to-date, and to incorporate the increasing flow of the new imagery. The use of an instance of the Open Data Cube can help to automate some of the processes, thus the Catalan Data Cube (CDC) was created as a regional data cube with easily managed, modest resource requirements in mind. The CDC is currently being implemented and operated by GRUMETS research group (http://www.grumets.uab.cat/index_eng.htm—mainly composed by CREAF and Autonomous University of Barcelona), as well as the SatCat.

The CDC (http://datacube.uab.cat) is being developed, collecting the same kind of optical imagery aimed by the SatCat, but only if it is available under the ARD paradigm over Catalonia (Spain), thus it is limited to Sentinel-2 level 2A data flows coming from ESA at the moment. The CDC is built on the Open Data Cube software suite [19], with some additional Python scripts and currently holds more than 1 year of Sentinel-2 (March 2018, April 2019). The archive is updated on a monthly basis with the most recent data, and contains at the time of writing these lines, 1562 granules, forming 132 daily slides, accounting for a total volume of 1.18 TB over Catalonia.

Following the tier separation introduced in the Swiss Data Cube section, in the Catalan Data Cube we also distinguish between "upstream tier" and the "downstream tier". The upstream tier is composed of the Sentinel 2 imagery Level 2A and the downstream tier; they are elaborated on-the-fly by the web client. Regarding the upstream tier, the following strategy for implementing standards is being adopted:

- *Discovery:* Since the number of products is limited, we are using the WMS GetCapabilities response as the document that acts as a catalogue, and provides links to ISO19115-2 and ISO19139-2 metadata documents.

- *View and Download:* To leverage the content of the CDC for visualization, OGC WMS has been implemented, and for download a Web Coverage Service (WCS) is under consideration.
- *Process:* The CDC relies on the ODC processing API, but currently it is not exposing processing services at the moment. Instead, it relies on what is possible to do in the MiraMon Map Browser client side.

These services are possible as part of the MiraMon Map Server CGI suite that is encoded in C and developed on top of libraries coming from the MiraMon software.

To further ease a user's interaction with CDC products, a web-based application called the "SatCat 2.0: Catalan Data Cube" (http://datacube.uab.cat/cdc) has been developed. It allows visualizing imagery and time-series data generated with the CDC. As briefly mentioned before, this web-based application is based in the MiraMon Map Browser (https://github.com/joanma747/MiraMonMapBrowser, open source web map client), which uses an extension of OGC WMS that allows querying/retrieving data in a raw binary array format. This solution allows the client to save in memory, the actual values of each band, and then use JavaScript code to operate with the data, providing some analytical tools to the user. The SatCat 2.0 provides the following functionalities to the user (some of them are only possible by the binary arrays approach):

- Visualizing single raster product layers.
- Generating histograms or pie charts for single raster product layers.
- Modifying raster visualization by describing enhancing contrast parameters or by changing colour palettes.
- Visualizing time-series raster product layers as animations.
- Generating a graph for a given point in space of a time-series raster product layer.
- Applying spatial filters (by setting a condition in another layer; e.g., representing normalized difference vegetation index (NDVI) values only if the elevation is lower a certain value, or only for a certain land use category).
- Creating new dynamic layers by complex calculations among the bands of the different available datasets.
- Accessing products in your own client server with WMS standards.
- Accessing metadata and reading or contributing geospatial user feedback.

Figure 2 shows the list of bands made available by the Sentinel 2 sensor followed by a list of colour combinations and band indices dynamically computed by the client side. As an example, a normalized difference vegetation index (NDVI) that is dynamically computed by the client while rendering it in the view (using the values of the necessary bands of each Sentinel-2 image). Moreover, a histogram showing the frequency of the NDVI values on the image is obtained. The dynamic calculation of layers can be as simple as this vegetation index, or a complex model using several bands and layers.

The generation of animations is possible in classic WMS services (TIME parameter), but thanks to the binary arrays approach, it is possible to present plots of the temporal evolution of one or more points in the animated area and eventually detect anomalies by comparing them with the mean and variance of the visible values in the bounding box (Figure 3). For both histograms and temporal profiles, data can be copied and pasted into a spreadsheet for further analysis.
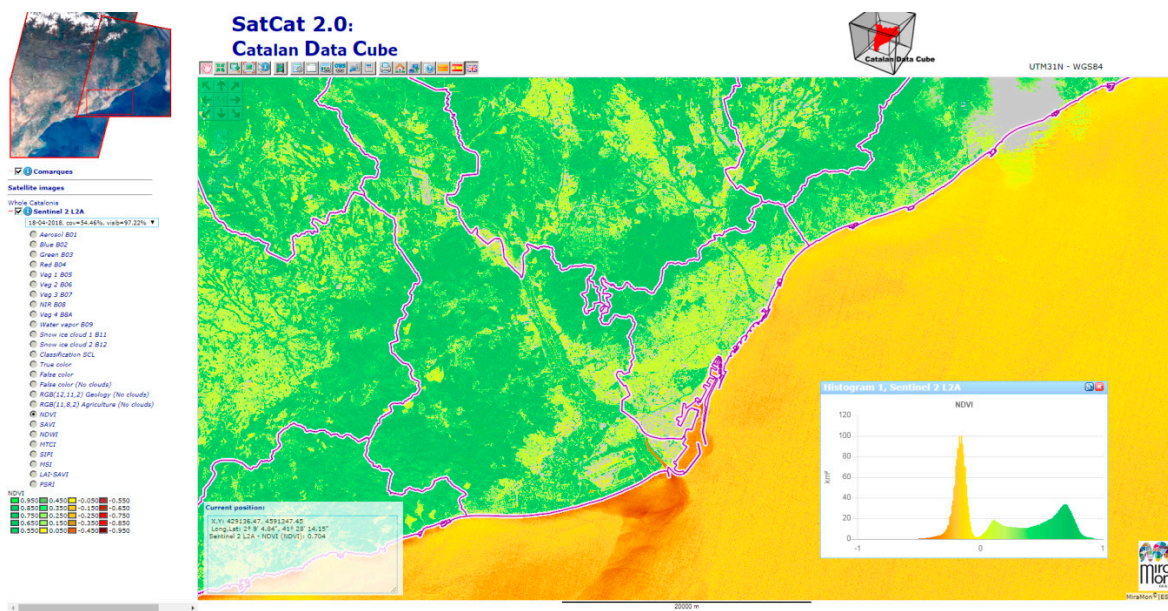
Figure 2. The Catalan Data Cube. Dynamic normalized difference vegetation index (NDVI) values and a histogram are over the view area, computed by the client using original Sentinel-2A red and infrared bands retrieved as binary arrays (centered in Barcelona and surroundings).
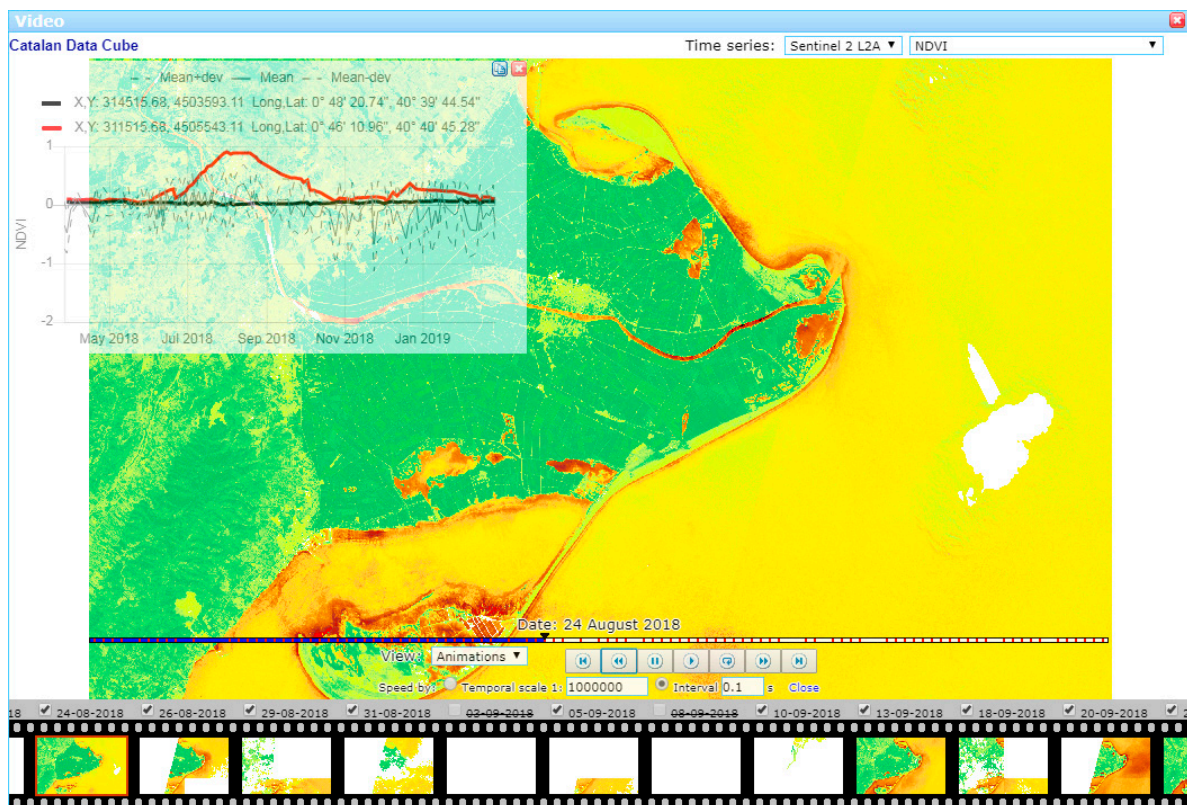
Figure 3. The Catalan Data Cube. Dynamic NDVI layer animation including, a temporal profile for sand (black) and crop (red) areas (centered in the Ebre river delta area).

6. Discussion and Perspectives

EODCs are becoming increasingly important to support Earth system science studies, enabling new global change applications. They can provide the long baseline required to determine trends,

define the present, and inform the future; they can deliver a unique capability to process, interrogate, and present EO satellite data in response to environmental issues; and allow tracking changes across a given area in unprecedented details.

However, the wider use and future success of EODC will largely depend on features such as: Usability and flexibility to address various users' needs and interoperability for contributing to the digital industry revolution [27]. Consequently, interoperability can be considered as a key factor for successful and wide adoption of EODC technology and is an absolute necessity to develop and implement regional and global data cube ecosystems.

Based on the present work, we think that the following three elements need to be carefully considered:

(1) Understanding the differences between traditional data system's interoperability (e.g., for discovery and download) and data cube system's interoperability for data analytics.
(2) Investigation of the significance of analysis ready data for specific applications.
(3) Explore the non-technological interoperability dimensions, such as governance and policy.

6.1. Interoperability Paradigms

In the EO domain, data interoperability has traditionally applied the "Discovery and Access" paradigm, which consists of discovering a remote dataset, downloading it to a local server, and using it locally; e.g., visualizing it or processing it to generate new data or information. In extreme synthesis, datasets have been moved though the network to be ingested in local data management platforms that support independent and monolithic applications.

With the advent of large datasets (i.e., data collected over a long-term or massive spatial data series) and the raise of virtual computing capacities, a new (and more efficient) IT paradigm emerged: The "Distributed Application" paradigm. The new approach aims at using the web as the analytics platform for building (distributed) applications, and makes use of microservices and container-based technologies. This time, datasets are not moved around, but application algorithms are deployed around (using the containerization technology) to be run where datasets are, working out a virtual collection of independent services that work together. For the EO Community, data cubes may be an important instrument to implement this paradigm—by facilitating remote data analytics. To achieve that, it is important to understand the best level of interoperability to be pursued by data cubes and which are the most effective instruments.

Advanced data cubes can be seen as databases able to organize and retrieve data on demand, and present it in the form of data structures, but also as processing facilities on top of data structures. In a data cube, data and processing are close together. Depending on the emphasis, data cubes can be used as traditional interoperability instruments (e.g., standard protocols for data discovery and access) to discover and download selections of remote sensing data, or the data cube systems can go beyond the "traditional" data systems—moving away of the "Discovery and Access" paradigm by becoming processing facilities, allowing data processing algorithms to be sent to the data cube engine and executed where the data is. Therefore, for EO data cubes, advanced interoperability tools were recently developed, such as coverage processing and query languages. Indeed, they can be used to implement system integrability. However, they are not sufficient to achieve semantic and pragmatic interoperability, which are important to achieve the "Distributed Application" paradigm that can process and use resources of more than one data cube. This interoperability shortcoming may push data cube systems to become monolithic platforms that operate in isolation only as clients of other interoperable services, but not offering interoperable interfaces by themselves.

6.2. Analysis Ready Data and Data Cubes

In general, data cubes can be seen as an analytical technological solution for taking advantage of ARD, defined by CEOS as "data that have been processed to a minimum set of requirements and

organized into a form that allows immediate analysis, with a minimum of additional user effort and interoperability, both through time and with other datasets" [82]. Despite this generic definition, CEOS came up with three specifications for surface reflectance, surface temperature, and radar backscatter (http://ceos.org/ard/). Focusing on the surface reflectance products, the general concept of data readiness entails common content pre-processing, such as: Atmospheric correction, and cloud coverage masking—this may be called radiometric readiness. There is no way to create a three-dimensional (i.e., x, y, t) data product if the radiometry of the values are not homogeneous over the time dimension. The most obvious advantage for the users of remote sensing data is that when they wish to undertake large area and long-time series analyses, they no longer need to invest in computationally expensive atmospheric correction processing chains to pre-process data. ARD saves time by providing a standard solution for data preparation that should be valid for most common applications. ARD products help with saving money, because expensive processing time is only executed once by the data producer. Another kind of data readiness (at least its lower level) is the geometrical readiness, by implementing data resampling and re-projection on a common grid environment. It is not possible to build a multidimensional data cube with scenes that are not organized in a way that they become geometrically co-registered. In principle, ARD does not assume a particular application; e.g., land cover studies. For example, ARD does not assume that cloud's shadows and snow shall be removed. Instead, producers apply state of the art algorithms to detect clouds and the shadows that they create, haze, and snow; and provide an extra band that tells which pixels are clouds, which are hazy, which are shadowed, and which contain snow. With these masks, the users can filter the values that they consider unusable by themselves in the data cube. For all these reasons, it can be considered that ARD allows working at the "pragmatic-level" of interoperability. Some specific usages of remote sensing data might require a more careful consideration for preparing the product for analysis. In the H2020 ECOPOTENTIAL project, we have experienced the need for applying shadow compensation techniques for steep mountain areas that provide better results than the ones coming from ARD directly generated by the data provider. However, there is another important benefit on using ARD: Users who wish to share and compare scientific and application results can still prefer to use ARD to reduce the potential discrepancies in results, due to differences in pre-processing, incrementing interoperability and comparability of higher-level products [78]. Nevertheless, ARD procedures can also create silos: For example, Landsat collection 1 and Sentinel 2 Level 2A were created as two independent ARD products, and they cannot be used together directly. The creation of a harmonized virtual product required to define a different ARD protocol and that forces, preserving the unique features of each data source, and some compromises must be made, such as adjusting Sentinel-2 (S2)/MSI radiometry to replicate the spectral bandpasses of Landsat 8/OLI, adopt a common 30 m resolution or adopt the Sentinel 2 UTM projection and tiling system [83].

Indeed, when populating data cubes, we are forced to satisfy a set of requirements that takes into account the actual use of the data, managed by them. The choice of an EO data cube array of dimensions, coordinate reference system, or data resolution are largely optimized for a limited set of relevant uses (e.g., time series analysis or changes detection using optical data). The data cube might still be useful for other kinds of applications, but performances would be suboptimal. As a consequence, application-driven optimization affects the different aspects and levels of data cubes' interoperability. In building distributed applications, the use of heterogeneous data cubes (i.e., differing for coordinate reference system, resolutions, etc.) would provide minimal or no benefit in comparison with using general purpose (or traditional) data systems. While a set of integrable data systems is still a data system, a set of integrable (i.e., technically and syntactically interoperable) data cubes is not necessarily a data cube itself. That defining of a data Cube is such by the virtue of relevant semantic/pragmatic decisions.

6.3. Non-Technological Interoperability Dimensions

Besides all technological aspects required enabling effective interoperability of EODC, we need also to consider non-technological aspects such as governance and policy. These elements mostly relate on human and organizational aspects that are equally important from an interoperability perspective [84]. From our point of view, the following three are essential for enhancing EODC interoperability:

(1) Currently, a commonly agreed definition and taxonomy of EODC is lacking. To our knowledge, the Data Cube Manifesto [85] is the only attempt to give a general and holistic definition, defining six requirements that must be met, in order to be considered as an EODC. This manifesto can be a good starting point to be further refined, looking at the various existing implementations, and embedding the effort in standardization processes, such as those supported by the OGC and ISO.

(2) Efforts should be persued to support Open Data and Sharing policies. Indeed, since 2008 the entire Landsat archive has been made freely and publicly available, followed by a tremendous increase in usage, investigations, and applications [86,87]. The Landsat Open Data policy is an excellent example of how to maximize the return on large investments in satellite missions [13,86]. Withtout such a policy, develing EODC technology would not have been possible. Togerther with FAIR (Findable, Accessible, Interoperable, Re-usable) data principles [88,89], EODC can enable moving towards effective and efficient EO Open Science.

(3) Finally, a fundamental aspect that needs to be considered is the governance. Without effective governance mechanisms and structures, it will prevent a successful implementation of EODC at national levels. Further, that will be even more important when one thinks about federated data cubes at regional and/or global levels. Governance will be the first challenge to tackle in this context. For example, in the case of the Swiss Data Cube, an incremental strategy has been developed. During the initial phase of the SDC, only one organization was involved taking care to test the data cube technology, deploy the software, ingest data, and developed initial demonstration applications. That helps fast movement and agility to closely collaborate under the mandate of the Swiss government. Now that the SDC is reaching some mature levels, new key partners in the field of EO in the country have been added to project bringing their respective expertise, and allowing consolidating the network across the country. This resulted in the signature of a Memorandum of Understanding (MoU) in June 2019 between UNEP/GRID-Geneva, the University of Geneva (UNIGE), the University of Zurich (UZH), and the Federal Institute for Forest, Snow and Landscape research (WSL). This cooperation agreement aims at fostering the use of Earth observation data for environmental monitoring on a national scale. The MoU is a pivotal instrument to clarify and implement a suitable governance structure commonly agreed by the different parties.

7. Conclusions

Addressing the interoperability challenge of EODC is essential to prevent the various EODC implementations becoming silos of information. Currently, not many efforts have been made to enhance interoperability of data cubes.

In this paper, we discuss and demonstrate how interoperability can be enabled using widely-adopted OGC and ISO geospatial standards, and how these standards can help delivering and leveraging the power of EO data building efficient discovery, access, and processing services. These standards are applied in different ways in current data cube implementations, such as the Swiss Data Cube and the Catalan Data Cube were, we have identified that OGC services mainly improve the "Discovery and Access" paradigm. An opposite paradigm of moving the processing code close to the data is facilitated by current containerization technology and rich query languages such as WCPS. The real challenge is to realize the "Distributed Application" paradigm, wherein data cubes can work together to produce analytical results.

Realizing the objective of providing EO-based information services and decision-ready products responding to users' needs requires effective and efficient mechanisms along the data value chain. EO data are essential to monitor and understand environmental changes. Consequently, it is necessary to make data and information products not in the form that it is collected, but in the form that is being used by the largest number of users possible. One step in this direction is the implantation of analysis ready data products by data providers. Being able to easily and efficiently combine EO-based data with other data sources is a crucial prerequisite to enable multi-disciplinary scientific analysis on our changing environment. Interoperable data cube services can significantly contribute to effective knowledge generation towards a more sustainable world, supporting decision and policy-makers making decisions based on evidence, and the best scientific knowledge.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rockström, J.; Bai, X.; deVries, B. Global sustainability: The challenge ahead. *Glob. Sustain.* **2018**, *1*, e6. [CrossRef]
2. Steffen, W.; Richardson, K.; Rockström, J.; Cornell, S.E.; Fetzer, I.; Bennett, E.M.; Biggs, R.; Carpenter, S.R.; Vries, W.D.; de Wit, C.A.; et al. Planetary boundaries: Guiding human development on a changing planet. *Science* **2015**, *347*, 1259855. [CrossRef] [PubMed]
3. Biermann, F.; Bai, X.; Bondre, N.; Broadgate, W.; Chen, C.T.A.; Dube, O.P.; Erisman, W.; Glaser, M.; van der Hel, S.; Lemos, M.C.; et al. Down to Earth: Contextualizing the Anthropocene. *Glob. Environ. Chang.* **2006**, *39*, 341–350. [CrossRef]
4. Giuliani, G.; Nativi, S.; Obregon, A.; Beniston, M.; Lehmann, A. Spatially enabling the Global Framework for Climate Services: Reviewing geospatial solutions to efficiently share and integrate climate data & information. *Clim. Serv.* **2007**, *8*, 44–58. [CrossRef]
5. Lehmann, A.; Chaplin-Kramer, R.; Lacayo, M.; Giuliani, G.; Thau, D.; Koy, M.; Goldberg, G.; Sharp, R., Jr. Lifting the Information Barriers to Address Sustainability Challenges with Data from Physical Geography and Earth Observation. *Sustainability* **2017**, *9*, 858. [CrossRef]
6. Ackoff, R.L. From Data to Wisdom. In *Ackoff's Best*; John Wiley & Sons: New York, NY, USA, 1999; pp. 170–172.
7. Rowley, J. The wisdom hierarchy: Representations of the DIKW hierarchy. *J. Inf. Sci.* **2007**, *33*, 163–180. [CrossRef]
8. Giuliani, G.; Dao, H.; Bono, A.D.; Chatenoux, B.; Allenbach, K.; Laborie, P.D.; Rodila, D.; Alexandris, N.; Peduzzi, P. Live Monitoring of Earth Surface (LiMES): A framework for monitoring environmental changes from Earth Observations. *Remote Sens. Environ.* **2017**, *202*, 222–233. [CrossRef]
9. Skidmore, A.K.; Pettorelli, N.; Coops, N.C.; Geller, G.N.; Hansen, M.; Lucas, R.; Mücher, C.A.; O'Connor, B.; Paganini, M.; Pereira, H.M.; et al. Environmental science: Agree on biodiversity metrics to track from space. *Nat. News* **2015**, *523*, 403–405. [CrossRef] [PubMed]
10. Anderson, K.; Ryan, B.; Sonntag, W.; Kavvada, A.; Friedl, L. Earth observation in service of the 2030 Agenda for Sustainable Development. *Geo-Spat. Inf. Sci.* **2017**, *20*, 1–20. [CrossRef]
11. Kavvada, A.; Held, A. Analysis-Ready Earth Observation Data and the United Nations Sustainable Development Goals. In Proceedings of the IGARSS 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 434–436. [CrossRef]

12. Boulton, G. The challenges of a Big Data Earth. *Big Earth Data* **2018**, *2*, 1–7. [CrossRef]

13. Ryan, B. The benefits from open data are immense. *Geospat. World* **2016**, 72–73.

14. Zhu, Z.; Wulder, M.A.; Roy, D.P.; Woodcock, C.E.; Hansen, M.C.; Radeloff, V.C.; Healey, S.P.; Schaaf, C.; Hostert, P.; Strobl, P.; et al. Benefits of the free and open Landsat data policy. *Remote Sens. Environ.* **2019**, *224*, 382–385. [CrossRef]

15. Lewis, A.; Lymburner, L.; Purss, M.B.J.; Brooke, B.; Evans, B.; Oliver, S.; Dekker, A.G.; Irons, J.R.; Minchin, S.; Mueller, N.; et al. Rapid, high-resolution detection of environmental change over continental scales from satellite data—The Earth Observation Data Cube. *Int. J. Digit. Earth* **2016**, *9*, 106–111. [CrossRef]

16. Sudmanns, S.; Lang, S.; Tiede, D. Big Earth Data: From Data to Information. *GI Forum* **2018**, *1*, 184–193. [CrossRef]

17. Baumann, P.; Misev, P.; Merticariu, V.; Huu, B.P.; Bell, B.B. Datacubes: A Technology Survey. In Proceedings of the IGARSS 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 430–433. [CrossRef]

18. Baumann, P.; Rossi, A.P.; Bell, B.; Clements, O.; Evans, B.; Hoenig, H.; Hogan, P.; Kakaletris, G.; Koltsida, P.; Mantovani, S.; et al. Fostering Cross-Disciplinary Earth Science Through Datacube Analytics. In *Earth Observation Open Science and Innovation*; Springer: Cham, Switzerland, 24 January 2018; pp. 91–119. [CrossRef]

19. Killough, B. Overview of the Open Data Cube Initiative. In Proceedings of the IGARSS 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 8629–8632. [CrossRef]

20. Strobl, P.; Marchetti, P.G. The Six Faces of the Data Cube. In Proceedings of the 2017 Conference on Big Data from Space, Toulouse, France, 28–30 November 2017; pp. 32–35. [CrossRef]

21. Rizvi, S.R.; Killough, B.; Cherry, A.; Gowda, S. The Ceos Data Cube Portal: A User-Friendly, Open Source Software Solution for the Distribution, Exploration, Analysis, and Visualization of Analysis Ready Data. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 8639–8642. [CrossRef]

22. Baumann, P.; Mazzetti, P.; Ungar, J.; Barbera, R.; Barboni, D.; Beccati, A.; Bigagli, L.; Boldrini, E.; Bruno, R.; Calanducci, A.; et al. Big Data Analytics for Earth Sciences: The EarthServer approach. *Int. J. Digit. Earth* **2016**, *9*, 3–29. [CrossRef]

23. Camara, G.; Assis, L.F.; Ribeiro, G.; Ferreira, K.R.; Llapa, E.; Vinhas, L. Big Earth Observation Data Analytics: Matching Requirements to System Architectures. In Proceedings of the 5th ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data, Burlingame, CA, USA, 31 October 2016; ACM: New York, NY, USA; pp. 1–6. [CrossRef]

24. Soille, P.; Burger, A.; de Marchi, D.; Kempeneers, P.; Rodriguez, A.R.D.; Syrris, V.; Vasilev, V. A versatile data-intensive computing platform for information retrieval from big geospatial data. *Future Gener. Comput. Syst.* **2018**, *81*, 30–40. [CrossRef]

25. European Commission. The DIAS: User-friendly Access to Copernicus Data and Information. Available online: https://www.copernicus.eu/sites/default/files/Copernicus_DIAS_Factsheet_June2018.pdf (accessed on 6 July 2018).

26. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]

27. Nativi, S.; Mazzetti, P.; Craglia, M. A view-based model of data-cube to support big earth data systems interoperability. *Big Earth Data* **2017**, *1*, 75–99. [CrossRef]

28. Geraci, A. *IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries*; IEEE Std 610; IEEE Press Piscataway: Piscataway, NJ, USA, 1991; pp. 1–217. [CrossRef]

29. Motta, R.C.; Oliveira, K.M.D.; Travassos, G.H. Rethinking Interoperability in Contemporary Software Systems. In Proceedings of the 2017 IEEE/ACM Joint 5th International Workshop on Software Engineering for Systems-of-Systems and 11th Workshop on Distributed Software Development, Software Ecosystems and Systems-of-Systems (JSOS), Buenos Aires, Argentina, 23 May 2017; pp. 9–15. [CrossRef]

30. Nativi, S.; Santoro, M.; Giuliani, G.; Mazzetti, P. Towards a knowledge base to support global change policy goals. *Int. J. Digit. Earth* **2019**, 1–29. [CrossRef]

31. Lehmann, A.; Nativi, S.; Mazzetti, P.; Maso, J.; Serral, I.; Spengler, D.; Niamir, A.; McCallum, I.; Lacroix, P.; Patias, P.; et al. GEOEssential–mainstreaming workflows from data sources to environment policy indicators with essential variables. *Int. J. Digit. Earth* **2019**, 1–17. [CrossRef]

32. Masó, J.; Serral, I.; Domingo-Marimon, C.; Zabala, A. Earth observations for sustainable development goals monitoring based on essential variables and driver-pressure-state-impact-response indicators. *Int. J. Digit. Earth* **2019**, 1–19. [CrossRef]

33. Domenico, B.; Nativi, S. OGC CF-netCDF 3.0 encoding using GML Coverage Application Schema Netcdf SWG, 2 November 2015. Available online: https://docs.opengeospatial.org/is/14-100r2/14-100r2.html (accessed on 7 June 2019).

34. Miller, H.G.; Mork, P. From Data to Decisions: A Value Chain for Big Data. *It Prof.* **2013**, *15*, 57–59. [CrossRef]

35. Giuliani, G.; Guigoz, Y.; Lacroix, P.; Ray, N.; Lehmann, A. Facilitating the production of ISO-compliant metadata of geospatial datasets. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *44*, 239–243. [CrossRef]

36. Bone, C.; Ager, A.; Bunzel, K.; Tierney, L. A geospatial search engine for discovering multi-format geospatial data across the web. *Int. J. Digit. Earth* **2016**, *9*, 47–62. [CrossRef]

37. Corti, P.; Kralidis, A.T.; Lewis, B. Enhancing discovery in spatial data infrastructures using a search engine. *PeerJ. Comput. Sci.* **2018**, *4*, e152. [CrossRef]

38. Lehmann, A.; Giuliani, G.; Ray, N.; Rahman, K.; Abbaspour, K.; Nativi, S.; Craglia, M.; Cripe, D.; Quevauviller, P.; Beniston, M. Reviewing innovative Earth observation solutions for filling science-policy gaps in hydrology. *J. Hydrol.* **2014**, *518*, 267–277. [CrossRef]

39. Díaz, P.; Masó, J.; Sevillano, E.; Ninyerola, M.; Zabala, A.; Serral, I.; Pons, X. Analysis of quality metadata in the GEOSS Clearinghouse. *Int. J. Spat. Data Infrastruct. Res.* **2012**, *7*, 352–377. [CrossRef]

40. Trilles, S.; Díaz, L.; Huerta, J. Approach to Facilitating Geospatial Data and Metadata Publication Using a Standard Geoservice. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 126. [CrossRef]

41. European Commission. *INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119*; European Commission Joint Research Centre: Abingdon, UK, 2010.

42. Senkler, K.; Voges, U.; Remke, A. An ISO 19115/19119 profile for OGC catalogue services CSW 2.0. Presented at the 10th EC GI & GIS Workshop, Warsaw, Poland, 23–25 June 2004; p. 9.

43. Bruha, L. Large geospatial images discovery: Metadata model and technological framework. *Geoinf. FCE CTU* **2015**, *14*, 21–36. [CrossRef]

44. Innerebner, M.; Costa, A.; Chuprikova, E.; Monsorno, R.; Ventura, B. Organizing earth observation data inside a spatial data infrastructure. *Earth Sci. Inf.* **2017**, *10*, 55–68. [CrossRef]

45. Zabala Torres, A.; Masó, J. Integrated hierarchical metadata proposal: Series, layer, entity and attribute metadata. Presented at the XXII International Cartographic Conference (ed.) Mapping Approaches into a Changing World, A Coruña, Spain, 9–15 July 2005.

46. Thome, K.J.; Helder, D.L.; Aaron, D.; Dewald, J.D. Landsat-5 TM and Landsat-7 ETM+ absolute radiometric calibration using the reflectance-based method. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 2777–2785. [CrossRef]

47. Markham, B.L.; Thome, K.J.; Barsi, J.A.; Kaita, E.; Helder, D.L.; Barker, J.L.; Scaramuzza, P.L. Landsat-7 ETM+ on-orbit reflective-band radiometric stability and absolute calibration. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 2810–2820. [CrossRef]

48. Barsi, J.A.; Markham, B.L.; Helder, D.L. Continued monitoring of Landsat reflective band calibration using pseudo-invariant calibration sites. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 7007–7010. [CrossRef]

49. Mishra, N.; Helder, D.; Angal, A.; Choi, J.; Xiong, X. Absolute Calibration of Optical Satellite Sensors Using Libya 4 Pseudo Invariant Calibration Site. *Remote Sens.* **2014**, *6*, 1327–1346. [CrossRef]

50. Zabala, A.; Maso, J. *Testbed-12 Imagery Quality and Accuracy Engineering Report*; OGC 16-050; OGC: Wayland, MA, USA, 2017.

51. Vitolo, C.; Elkhatib, Y.; Reusser, D.; Macleod, C.J.A.; Buytaert, W. Web technologies for environmental Big Data. *Environ. Modell. Softw.* **2015**, *63*, 185–198. [CrossRef]

52. Hu, C.; Zhao, Y.; Li, J.; Ma, D.; Li, X. Geospatial Web Service for Remote Sensing Data Visualization. In Proceedings of the 2011 IEEE International Conference on Advanced Information Networking and Applications, Washington, DC, USA, 22–25 March 2011; pp. 594–601. [CrossRef]

53. Lankester, T.H.G. *OpenGIS Web Map Services—Profile for EO Prodcuts*; OGC: Wayland, MA, USA, 2009.

54. Blower, J.D.; Gemmell, A.L.; Griffiths, G.H.; Haines, K.; Santokhee, A.; Yang, X. A Web Map Service implementation for the visualization of multidimensional gridded environmental data. *Environ. Modell. Softw.* **2013**, *47*, 218–224. [CrossRef]

55. Maso, J.; Zabala, A.; Serral, I.; Pons, X. Remote Sensing Analytical Geospatial Operations Directly in the Web Browser. In Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Delft, The Netherlands, 1–5 October 2018; Volume XLII–4, pp. 403–410. [CrossRef]

56. Maso, J.; Zabala, A.; Serral, I.; Pons, X. A Portal Offering Standard Visualization and Analysis on top of an Open Data Cube for Sub-National Regions: The Catalan Data Cube Example. *Data* **2019**, *4*, 96. [CrossRef]

57. Guo, H. Big Earth data: A new frontier in Earth and information sciences. *Big Earth Data* **2017**, *1*, 4–20. [CrossRef]

58. Guo, H.; Liu, Z.; Jiang, H.; Wang, C.; Liu, J.; Liang, D. Big Earth Data: A new challenge and opportunity for Digital Earth's development. *Int. J. Digit. Earth* **2017**, *10*, 1–12. [CrossRef]

59. Open Geospatial Consortium. *OGC OWS-6 WPS Grid Processing Profile Engineering Report*; OGC: Wayland, MA, USA, 2009.

60. Baumann, P.; Misev, D.; Merticariu, V.; Huu, B.P. Datacubes: Towards Space/Time Analysis-Ready Data. In *Service-Oriented Mapping*; Springer: Cham, Switzerland, 2019; pp. 269–299. [CrossRef]

61. Nandra, C.; Bâcu, V.; Gorgan, D. Parallel Earth Data Tasks Processing on a Distributed Cloud Based Computing Architecture. In Proceedings of the 2017 21st International Conference on Control Systems and Computer Science (CSCS), Bucharest, Romania, 29–31 May 2017; pp. 677–684. [CrossRef]

62. Evans, B.; Wyborn, L.; Pugh, T.; Allen, C.; Antony, J.; Gohar, K.; Porter, D.; Smillie, J.; Trenham, C.; Wang, J.B.; et al. The NCI High Performance Computing and High Performance Data Platform to Support the Analysis of Petascale Environmental Data Collections. In *Environmental Software Systems. Infrastructures, Services and Applications, Proceedings of the 11th IFIP WG 5.11 International Symposium, ISESS 2015, Melbourne, VIC, Australia, 25–27 March 2015*; Springer: Cham, Germany, 2015; pp. 569–577. [CrossRef]

63. Purss, M.B.J.; Lewis, A.; Oliver, S.; Ip, A.; Sixsmith, J.; Evans, B.; Edberg, R.; Frankish, G.; Hurst, L.; Chan, T. Unlocking the Australian Landsat Archive—From dark data to High Performance Data infrastructures. *GeoResJ* **2015**, *6*, 135–140. [CrossRef]

64. Castronova, A.M.; Goodall, J.L.; Elag, M.M. Models as web services using the Open Geospatial Consortium (OGC) Web Processing Service (WPS) standard. *Environ. Modell. Softw.* **2013**, *41*, 72–83. [CrossRef]

65. Baumann, P. Datacube Standards and their Contribution to Analysis-Ready Data. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 2051–2053. [CrossRef]

66. Xue, Y.; Palmer-Brown, D.; Guo, H.D. The use of high-performance and high-throughput computing for the fertilization of digital earth and global change studies. *Int. J. Digit. Earth* **2011**, *4*, 185–210. [CrossRef]

67. Olasz, A.; Thai, B.N.; Kristóf, D. Development of a new framework for Distributed Processing of Geospatial Big Data. *Int. J. Spat. Data Infrastruct. Res.* **2016**, *12*, 85–111. [CrossRef]

68. Giuliani, G.; Nativi, S.; Lehmann, A.; Ray, N. WPS mediation: An approach to process geospatial data on different computing backends. *Comput. Geosci.* **2012**, *47*, 20–33. [CrossRef]

69. Mazzetti, P.; Roncella, R.; Mihon, D.; Bacu, V.; Lacroi, P.; Guigoz, Y.; Ray, N.; Giuliani, G.; Gorgan, D.; Nativi, S. Integration of data and computing infrastructures for earth science: An image mosaicking use-case. *Earth Sci. Inf.* **2016**, 1–18. [CrossRef]

70. Rodila, D.; Ray, N.; Gorgan, D. Conceptual model for environmental science applications on parallel and distributed infrastructures. *Environ. Syst. Res.* **2015**, *4*, 1–16. [CrossRef]

71. Nativi, S.; Mazzetti, P.; Geller, G.N. Environmental model access and interoperability: The GEO Model Web initiative. *Environ. Model. Softw.* **2013**, *39*, 214–228. [CrossRef]

72. Di, L.; Yue, P.; Ramapriyan, H.K.; King, R.L. Geoscience Data Provenance: An Overview. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 5065–5072. [CrossRef]

73. Closa, G.; Maso, J.; Zabala, A.; Pesquer, L.; Pons, X. A provenance metadata model integrating ISO Geospatial lineage and the OGC WPS: Conceptual model and implementation. *Trans. GIS* **2019**. [CrossRef]

74. Lewis, A.; Olivera, S.; Lymburner, L.; Evans, B.; Wyborn, L.; Mueller, N.; Raevksi, G.; Hooke, J.; Woodcock, R.; Sixsmith, J.; et al. The Australian Geoscience Data Cube—Foundations and lessons learned. *Remote Sens. Environ.* **2017**, *202*, 276–292. [CrossRef]

75. Wagner, W.; Dorigo, W.; de Jeu, R.; Fernandez, D.; Benveniste, J.; Haas, E.; Ertl, M. Fusion of active and passive microwave observations to create an Essential Climate Variable data record on soil moisture. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *7*, 315–321. [CrossRef]

76. Pettorelli, N.; Wegmann, M.; Skidmore, A.; Lucas, R.; Rocchini, D.; Fernandez, N.; Turak, E.; Reyers, B.; Geller, G.N.; Belward, A.; et al. Framing the concept of satellite remote sensing essential biodiversity variables: Challenges and future directions. *Remote Sens. Ecol. Conserv.* **2016**, *2*, 276–292. [CrossRef]
77. Giuliani, G.; Chatenoux, B.; De Bono, A.; Rodila, D.; Richard, J.-P.; Allenbach, K.; Dao, H.; Peduzzi, P. Building an Earth Observations Data Cube: Lessons learned from the Swiss Data Cube (SDC) on generating Analysis Ready Data (ARD). *Big Earth Data* **2017**, *1*, 1–18. [CrossRef]
78. Giuliani, G.; Chatenoux, B.; Honeck, E.; Richard, J. Towards Sentinel-2 Analysis Ready Data: A Swiss Data Cube Perspective. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 8659–8662. [CrossRef]
79. Denis, G.; Claverie, A.; Pasco, X.; Darnis, J.P.; de Maupeou, B.; Lafaye, M.; Morel, E. Towards disruptions in Earth observation? New Earth Observation systems and markets evolution: Possible scenarios and impacts. *Acta Astronaut.* **2017**, *137*, 415–433. [CrossRef]
80. European Commission. *Big Data in Earth Observation*; European Commission: Brussels, Belgium, 2017.
81. Maso, J.; Pons, X.; Zabala, A. Tuning the second-generation SDI: Theoretical aspects and real use cases. *Int. J. Geogr. Inf. Sci.* **2012**, *26*, 983–1014. [CrossRef]
82. Dwyer, J.; Roy, D.; Sauer, B.; Jenkerson, C.; Zhang, H.; Lymburner, L. Analysis Ready Data: Enabling Analysis of the Landsat Archive. *Remote Sens.* **2018**, *10*, 1363. [CrossRef]
83. Claverie, M.; Ju, J.; Masek, J.G.; Dungan, J.L.; Vermote, E.F.; Roger, J.C.; Justice, C.; Skakun, S.V. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sens. Environ.* **2018**, *219*, 145–161. [CrossRef]
84. Giuliani, G.; Lacroix, P.; Guigoz, Y.; Roncella, R.; Bigagli, L.; Santoro, M.; Lehmann, A. Bringing GEOSS Services into Practice: A Capacity Building Resource on Spatial Data Infrastructures (SDI). *Trans. GIS* **2016**, *21*, 811–824. [CrossRef]
85. Baumann, P. *The Datacube Manifesto*; European Commission: Brussels, Belgium, 2017.
86. Wulder, M.A.; Masek, J.G.; Cohen, W.B.; Loveland, T.R.; Woodcock, C.E. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sens. Environ.* **2018**, *122*, 2–10. [CrossRef]
87. Wulder, M.A.; Loveland, T.R.; Roy, D.P.; Crawford, C.J.; Masek, J.G.; Woodcock, C.E.; Dwyer, J.; Hermosilla, T.; Hipple, J.D.; Hostert, P.; et al. Current status of Landsat program, science, and applications. *Remote Sens. Environ.* **2019**, *225*, 127–147. [CrossRef]
88. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Bouwman, J.; Schultes, E.; Roos, M.; Grethe, J.S.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **2016**, *3*. [CrossRef]
89. Stall, S.; Yarmey, L.; Cutcher-Gershenfeld, J.; Hanson, B.; Lehnert, K.; Nosek, B.; Wyborn, L. Make scientific data FAIR. *Nature* **2019**, *570*. [CrossRef] [PubMed]