AGILE: GIScience Series, 5, 11, 2024. https://doi.org/10.5194/agile-giss-5-11-2024
Proceedings of the 27th AGILE Conference on Geographic Information Science, 4–7 Sept. 2024.
Editors: Alison Heppenstall, Mingshu Wang, Urska Demsar, Rob Lemmens, and Jing Yao.
This contribution underwent peer review based on a full paper submission.

Towards semantic enrichment of Earth Observation data: The LEODS framework

Daniela F. Milon-Flores[1], Camille Bernard[1], Jérôme Gensel[1], Gregory Giuliani[2], Bruno Chatenoux[3], and Hy Dao[4]

[1]Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, Grenoble, France
[2]University of Geneva, Institute for Environmental Sciences, Geneva, Switzerland
[3]GRID-Geneva, United Nations Environment Program, Geneva, Switzerland
[4]University of Geneva, Department of Geography and Environment, Geneva, Switzerland

Correspondence: Daniela F. Milon-Flores (daniela.milon-flores@univ-grenoble-alpes.fr)

Abstract. The Earth's ecosystem is facing serious threats due to the depletion of natural resources and environmental pollution. To promote sustainable practices and formulate effective policies that address these issues, both experts and non-expert stakeholders require access to meaningful Open Data. Current Earth monitoring programs provide a large volume of open Earth Observation (EO) data typically organized and managed in EO Data Cubes (EODCs). From these datasets, satellite-derived indices can be calculated for assessing various environmental aspects in areas of interest over time. However, current EOs lack semantics and are isolated from significant Web resources, greatly hindering their comprehension and limiting their use to specialized users. To enhance EO data with semantic richness and ensure their understanding by a wider audience, it is pertinent to adopt a Linked Open Data (LOD) approach. In this paper, we present the Linked Earth Observation Data Series (LEODS) framework designed to publish aggregated EO data in the LOD Cloud. LEODS provides a processing chain that converts EO data into EO-RDF data cubes based on a spatio-temporal modeling approach that ensures integration and future semantic enrichment of EO data while preserving the advantages of traditional EODCs and following the FAIR principles (*i.e.,* findable, accessible, interoperable, and reusable). To highlight the advantages of our proposal, we explore through SPARQL queries and visualizations, the results of implementing LEODS with study areas located in Switzerland and France.

Keywords. Spatiotemporal data, Knowledge Graphs, Semantic Enrichment, Spatial Aggregation, Data Cube, Earth Observations

1 Introduction

The Earth's ecosystem faces imminent dangers from the depletion of natural resources and pollution of air, soil, and water. Leading organizations such as the International Union for Conservation of Nature (IUCN)[1] and the Intergovernmental Panel on Climate Change (IPCC)[2], along with numerous studies (Schuldt et al. (2020); Salomón et al. (2022)), constantly highlight the serious consequences of human activities on the environment. In response, stakeholders — ranging from policymakers, citizens, associations, and analysts — are increasingly focused on monitoring and understanding environmental changes at the local level. This effort, allows them to be better informed and acquire knowledge for promoting sustainable environmental practices and policies. Therefore, access to meaningful and comprehensible Open Data is indispensable for enhancing the understanding and interpretation of climate change trends over time.

Monitoring programs such as US Landsat[3] and European Copernicus Sentinel[4] provide a free and open collection of satellite data depicting the Earth, also known as Earth Observation (EO) data. These datasets enable experts to assess various environmental characteristics in areas of interest, such as monitoring the health of vegetation through satellite-derived indices as the well-known Normalized Difference Vegetation Index (NDVI) (Appel and Pebesma, 2019). Due to the enormous amount of EO data, most state-of-the-art works (Lewis et al., 2017; Giuliani et al., 2017; Appel and Pebesma, 2019) propose to adopt Earth Observation Data Cubes (EODCs) (Baumann, 2017) as a technological solution for storing, managing, accessing,

[1]iucn.org
[2]ipcc.ch
[3] landsat.gsfc.nasa.gov
[4] copernicus.eu/en

and analyzing large EOs (Giuliani et al., 2017). However, traditional EODCs have two major limitations: (1) their interpretation is usually reserved for specialists even when metadata is available to explain the satellite-derived indices (Augustin et al., 2019); (2) their relative isolation from available Web resources supplied by the Open Data Initiative (Ubaldi, 2013) complicates the proper contextualization and investigation of a study area (Nativi et al., 2017; Giuliani et al., 2019). Consequently, the expert-oriented and non-contextualized data available in EODCs are insufficient to adequately understand environmental changes over time and their underlying causes.

In order to address these challenges, a paradigm that has gained popularity in recent decades, known as Linked Open Data (LOD), offers a comprehensive framework for describing and connecting Web-based information (Patel and Jain, 2021; Hogan, 2020; de Sousa, 2023). By leveraging Semantic Web (SW) technologies, including standards like RDF (Resource Description Framework) and OWL (Web Ontology Language) for ontology modeling, data can be effectively described and published as Knowledge Graphs (KG) within the LOD Cloud. This approach not only enhances data accessibility for both humans and machines (Cyganiak et al., 2014), but also provides robust support for data integration (Tran et al., 2020b), all while adhering to the FAIR principles (*i.e.,* findable, accessible, interoperable, and reusable). Therefore, to enhance EO data with semantic richness, guarantee its understanding by a wider audience, and break down data silos while introducing causal links between different datasets, it is pertinent to adopt a LOD approach.

Precisely, the RDF Data Cube (QB) is a SW technology specifically designed for publishing data in the form of open-linked data cubes (Richard Cyganiak and Tennison, 2014). This World Wide Web Consortium (W3C) standard ontology enables the integration of heterogeneous data sharing common dimensions such as time and space. Moreover, it facilitates the semantic enrichment of EO data by linking it to various resources within the LOD Cloud. Due to its compatibility with the Statistical Data and Metadata eXchange (SDMX) standard model [5], the QB vocabulary is extensively utilized by national and international statistical agencies, such as those in Scotland[6] and France[7], for publishing and exchanging socioeconomic data. However, despite its potential, it is not widely employed for the publication of EO data, with notable exceptions such as the collaboration between the Open Geospatial Consortium (OGC) and the W3C. This initiative focuses on representing dense raster-level EO data in the LOD Cloud using the QB vocabulary (Brizhinev et al., 2017). Nevertheless, it is important to note that this approach may result in high data storage costs, as each pixel is stored in the cube.

[5] sdmx.org/

[6] statistics.gov.scot/home

[7] rdf.insee.fr/def/index.html

Our research aims to effectively leverage the QB vocabulary in conjunction with several SW technologies to explore their advantages within the EO domain. Therefore, we introduce the Linked Earth Observation Data Series (LEODS) framework, which is specifically designed to integrate and publish aggregated EO data on the LOD Cloud. LEODS provides and implements a processing chain focused on transforming EO data into what we call EO-RDF data cubes. Primarily, our approach aggregates satellite-derived indices at the lowest administrative division level to be as close as possible to stakeholders and ensure contextualization with local resources. This strategy also helps to mitigate the high storage costs associated with pixel-level data. Besides, note that for uniformity in this study, the term “municipality” is hereafter adopted as the denomination of the lowest administrative division level [8]. Subsequently, the framework adopts SW technologies for the spatiotemporal modeling of the cubes, ensuring their semantic enrichment and integration with various Web resources such as socio-economic, urban, climatic, and legislative texts Finally, to illustrate the benefits of our proposal, we utilize SPARQL queries and visualizations to explore the results of implementing the LEODS framework in real study areas located in the geographical extension of Switzerland and part of France. As a result, three EO-RDF data cubes with hierarchical dimensions: spatial (municipalities, departments, countries), temporal (daily, monthly, seasonal, annual), and indices (vegetation, water, snow, and urbanization) are now available in a SPARQL endpoint.

Furthermore, LEODS is the first step towards the main objective of the French-Swiss collaborative project called TRACES[9], that is, to build a Knowledge Graph (KG) integrating EO data with various data sources (socio-economic, urban, legislative texts, etc.) to monitor the environmental evolution of areas of interest.

The article is structured as follows: Section 2 outlines the related work. Section 3 introduces the LEODS framework and the entire processing chain it supports. In Section 4, the study areas and the results obtained are presented. Section 5 offers discussions on the paper. Finally, Section 6 concludes the article and presents future work.

2 Related work

As noted by (Augustin et al., 2019; Van Der Meer et al., 2022), traditional EODCs store sensor data lacking semantic information. To accurately interpret EO data, a process of semantic enrichment is essential, which involves augmenting EOs with additional contextual information. Two approaches for semantic enrichment are commonly

[8] The lowest administrative division level term varies among regions, *e.g.,* “commune” in France and Switzerland and “municipality” in Great Britain.

[9] TRACES project PRCI Franco-Suisse. Funded by ANR and FNS. Website: traces-anr-fns.imag.fr/

employed in literature: content-based and ontology-based methods. Content-based methods specialize in extracting meaningful features directly from the image content, including pixel color information (Augustin et al., 2019), land cover types (Simoes et al., 2021; Datcu et al., 2003), and pixel quality (Sudmanns et al., 2021), thus enhancing the sensor data. On the other hand, ontology-based approaches (Brizhinev et al., 2017; Lefort et al., 2012; Ayadi et al., 2022) leverage SW technologies to represent and augment EO data with Linked Data resources. These latter methods are described below.

In the LOD paradigm, ontology models, also named as vocabularies, are used to define domain-specific concepts and their relationships. During our research, we have identified several ontologies centered on EO data; the most relevant ones are:

1. The joint W3C and OGC Spatial Data on the Web (SDW) Working Group developed a set of ontologies to describe sensors, actuators, observations, procedures, etc. A first module called SOSA (Sensor, Observation, Sampler, and Actuator)[10] and an extension module called SSN (Semantic Sensor Network)[11]. Although both ontologies can be easily integrated to support in-depth descriptions related to sensors, the ontologies may be too complex if a simple sensor description is needed.

2. The W3C RDF Data Cube Vocabulary (QB)[12] is a W3C recommendation for the publishing of multidimensional data beyond the EO domain in the form of open-linked data cubes. The QB vocabulary is aligned with OLAP (Online Analytical Processing) concepts, allowing efficient storage, management, and accessibility of the data. Although the vocabulary is widely used in socio-economic domains, a mis-modeling of its components (*e.g.,* dimensions, measures, and attributes) can lead to silos of cubes that cannot be reused in LOD Cloud resulting in major isolation problems.

3. The Agricultural Information Model (AIM) is an ontology developed in the framework of the Horizon 2020 DEMETER project Palma et al. (2022). The goal of AIM is to enable information interoperability between data domains such as agricultural data, EO data, and meteorological data by using state-of-the-art ontologies such as SNN/SOSA, GeoSPARQL, and OWL-Time. At the moment AIM remains under development and is expected to be approved as an OGC standard in the future.

Later, with or without relying on the ontologies above, several initiatives to publish EO data in the LOD Cloud were proposed, including those of TELEIOS and LEO projects (Koubarakis et al., 2014). In both studies several tools, such as Geotriples and Sextant (Nikolaou et al., 2015; Kyzirakos et al., 2018), and ontologies such as stRDF (Koubarakis and Kyzirakos, 2010) were developed to manage big EO data. Specifically in (Koubarakis et al., 2016), the authors propose to use the stRDF ontology to publish satellite image data in the LOD Cloud. However, their approach is limited to using SW technologies to publish only satellite image metadata (*e.g.,* time of acquisition and geographical coverage) as LOD, while ignoring the publication of other relevant information such as environmental indices that can be calculated from the imagery. In the work of (Tran et al., 2020a, b), the authors argue that raster data is not human-readable and that SW ontologies can be used for the integration of data calculated from raster, *e.g.,* land cover indices, change indicators, etc. Thus, as a first contribution, they propose a network of ontologies that allows the representation of such data in the LOD Cloud. Among the ontologies, they reuse the SOSA vocabulary to describe raster observations, and the TSN (Territorial Statistical Nomenclature Ontology)[13] and OWL-Time ontologies to associate spatio-temporal concepts. Subsequently, a methodology is introduced to extract features from the pixel data and map them to the semantic model. Although a network of ontologies seems to be the most appropriate approach to comply the LOD vision, this specific work does not use the QB ontology despite being a standard recommendation suitable for modeling and integrating heterogeneous spatio-temporal data. QB has already been used to organize data of various natures in the LOD Cloud, such as medical data (Leroux and Lefort, 2012; Rodriguez and Hogan, 2021; Casey et al., 2022), historical data (Bayerl and Granitzer, 2015) and especially socio-economic data[14] due to its compatibility with the SDMX model. However, despite the benefits of the QB vocabulary, it is rarely used in the field of EO data. In the following, we describe the limited works found in the literature that adopt an ontology network approach involving QB to publish and enrich EO data in the LOD Cloud.

In (Brizhinev et al., 2017), the W3C and the OGC consortia introduced a method for publishing EO raster data using ontologies such as QB, SNN, and GeoSPARQL. The proposal describes how dense geospatial raster data can be organized in the dimensions of the cube: latitude, longitude, and time, and also explains how pixel metadata and their provenance can be attached to their components. However, within this approach, data is published pixel by pixel, which is very costly in terms of data storage. Furthermore, it would be more appropriate to publish aggregated data at the lowest administration level, *e.g.,* municipalities, which are meaningful to stakeholders and enable an adequate contextualization of a given area with complementary resources such as socio-economic ones. The works of Lefort et al. (2012); Ayadi et al. (2022) describe

[10] w3.org/ns/sosa/

[11] w3.org/ns/ssn/

[12] w3.org/TR/vocab-data-cube/

[13] lig-tdcge.imag.fr/tsn/index.html

[14] Scotland: statistics.gov.scot/home and France: rdf.insee.fr/def/index.html

a similar approach to represent meteorological data such as temperature and humidity as LOD to monitor climate variability and capture its behavior. Among the reused ontologies, the QB vocabulary is used to create spatiotemporal slices of meteorological observations enriched with statistical attributes. In addition, both methods leverage the SSN ontology to perform a more tailored description of the observations, such as sensor-related features and data collection methodology. Although this procedure granularly and technically contextualizes the EOs, the result is a data cube restricted to represent only meteorological data. We believe that this part of their approach diverges from our objective, since non-expert stakeholders are not concerned with such a low level of granularity description of the observations and the lack of such data does not affect the interpretation of the environmental evolution.

In our work, we are interested in opening up EO data for wide audiences by reusing standard semantic models including the RDF data cube. Unlike most of the related work described above, our focus is neither on the simple representation of metadata as LOD nor on the treatment of EO data at the raster level. Instead, we focus on publishing satellite-derived indices aggregated at the lowest administrative division level, to be as close as possible to expert and non-expert stakeholders interested in the environmental aspects of their municipalities. Therefore, a specific framework is needed to describe how to preprocess, model, publish, and explore aggregated EO spatiotemporal data on the LOD Cloud. The approach should adopt a multidimensional modeling generic enough (a) to be applicable to any globally measured EO data and (b) to ensure the integration, through shared spatiotemporal dimensions, with data cubes published in domains other than EO.

3 The LEODS framework

In this section, we present the LEODS framework, which supports the transformation of EO data into RDF data cubes; while assuring its future integration and enrichment with several Web resources. As illustrated in Figure 1, LEODS covers a processing chain that: (1) aggregates EO data, initially at the pixel level, to the municipality level, (2) designs a multidimensional model that gives structure to the data cubes, (3) instantiates the model and publishes the data as EO-RDF data cubes on the LOD Cloud, and (4) provides users with predefined SPARQL queries to explore the open-linked data cubes. In the following subsections, each step of the LEODS framework is explained.

3.1 Step 1: EO data preparation

The initial phase of our framework focuses on aggregating EO data from pixel to municipality level, aiming to (1) be as close as possible to the stakeholders, (2) ensure municipalities contextualization with available local Web resources, and (3) manipulate multidimensional structured data. This process begins with satellite image acquisition for a specified study area and observation period, followed by essential preprocessing steps like geometric correction and denoising. The spatial aggregation occurs at the lowest administrative division level, *i.e.,* the municipality level, where derived environmental indices, such as NDVI, are calculated with various temporal aggregations like seasonal and annual. Zonal statistical methods, such as mean and standard deviation (std) calculations, are then applied to summarize the data for each index. Finally, the preprocessed data are delivered as vector time-series stored in tabular files. This phase is optional if the user is directly working with time series of EO vectors.

3.2 Step 2: Modeling EO semantic data cubes

The LEODS framework's second step focuses on modeling the Data Structure Definition (DSD) of the QB ontology to encapsulate the essential characteristics of EO data into RDF data cubes. To achieve this, our first objective is to link the EO vector time-series values, obtained previously, with the components of the DSD. These components include: (1) *Observations*: Representing the observed values in the dataset. (2) *Dimensions*: Observations can be organized along two or more dimensions. Given the nature of EO data, standard dimensions like space and time are essential. Additionally, domain-specific dimensions, such as environmental indices derived from satellite imagery, should be incorporated. It is advisable to refer to the list of standard dimensions[15] provided by the SDMX model to determine if tabular values can be mapped as dimensions. We also suggest defining the dimensions with a hierarchical structure to allow essential SPARQL queries when exploring the data. (3) *Attributes*: Provide context to observations. For instance, the value 50 becomes meaningful when associated with attributes like hectares (HA). The QUDT vocabulary[16] offers a range of attributes that can be readily reused. (4) *Measures*: Refer to the specific phenomenon observed and are closely related to attributes. In the EO data domain, typical measures include zonal statistics (*e.g.,* mean, std), land cover surface, and land surface temperature. SDMX provides a list of standard measures[17], simplifying the identification of this component.

Once the DSD components have been identified, it is necessary to describe the data itself using standard ontologies that specify the dimensions of the data. Examples of relevant ontologies include: (1) The OWL-Time ontology[18], developed by the W3C, as a standard for representing temporal information in the LOD Cloud. It defines classes and properties for describing time-related concepts such as instants, intervals, and durations. (2) The GeoSPARQL on-

[15] purl.org/linked-data/sdmx/2009/dimension
[16] qudt.org/vocab/unit/
[17] purl.org/linked-data/sdmx/2009/measure
[18] w3.org/TR/owl-time/

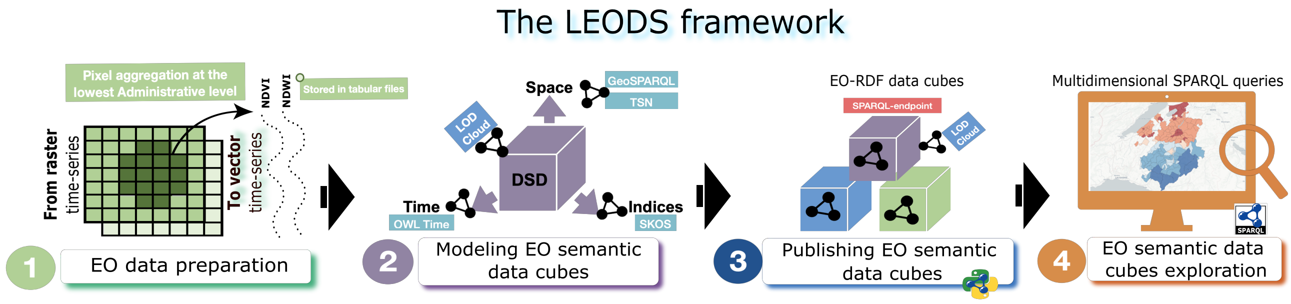


Figure 1. Pipeline of the LEODS framework.

tology[19], developed by the OGC, as a standard for representing and querying geospatial data in the Semantic Web. GeoSPARQL allows the integration of geospatial data with other linked data sources, facilitating spatial analysis and reasoning. (3) The Territorial Statistical Nomenclature (TSN) ontology is the outcome of our prior research (Bernard et al., 2018). It focuses on the semantic representation of geographic divisions, including administrative units such as regions, districts, and municipalities, as well as their temporal evolution. (4) The Simple Knowledge Organization System (SKOS)[20] is a widely used ontology for representing hierarchies systems such as thesauri and taxonomies. It provides a set of classes and properties to describe concepts, labels, and relationships between concepts.
Furthermore, it is recommended to enhance the cube components with basic semantics. For example, connecting the spatial dimension to official geographic KGs available in the LOD Cloud can provide a more detailed characterization of municipalities. Additionally, including metadata that describes and provides context about the components, such as their provenance and meaning, improves interpretability and reusability

DSD modeling is a crucial step and some criteria should be considered during the process: (a) the model must be generic enough to ensure integration with heterogeneous data sharing standard dimensions such as time and space and to be applicable to any globally measured EO data beyond environmental indices, and (b) the cube components should be properly defined and reusable as possible to mitigate the silos of RDF data cubes.

3.3 Step 3: Publishing EO semantic data cubes

After modeling the structure of the RDF data cubes, the next step consists of the semi-automatic transformation of the EO vector time-series into EO-RDF data cubes following the modeled DSD. This process can be facilitated using tools such as the RDFlib library[21], the Open Refine tool[22], the RDF mapping language[23], and the command-line tool Tarql[24]. Open Refine supports the RDF language as an extension, allowing the instance of basic components of the model. However, we recommend using it for an initial interaction process with QB components and RDF syntax, as the tool can be restrictive during implementation. For greater flexibility, we suggest using the RDFlib library to instantiate the full conceptual model of the cubes (the DSD) in RDF triples using Python; which simplifies the RDF syntax. Then, ad-hoc Tarql scripts can be implemented to automatically convert the observations, in tabular files, to RDF files following the DSD cube model. Subsequently, both the instantiated model and the observations files can be concatenated into a single file containing EO-RDF data cubes. Then, to fully comply with the vision of the LOD paradigm, the produced linked cubes must be published and made accessible to interested stakeholders through a SPARQL endpoint, a Web service that enables the querying of RDF data.

3.4 Step 4: EO semantic data cubes exploration

Finally, once the EO-RDF data cubes are accessible, it is crucial to initiate the exploration phase, demonstrating how to extract meaningful information from these cubes. For this aim, we propose to use the GraphDB tool[25], which, in addition to storing triples, includes a LOD graph visualization module. This visualization is essential for discovering and navigating through the cube components. GraphDB also supports SPARQL and GeoSPARQL queries, essential for manipulating and retrieving information from EO-RDF data cubes. Thus, to facilitate analysis and interpretation, we suggest providing users with a set of predefined queries that exploit the advantages of the implemented cube model. Greater diversity in the queries enhances the chances for interested users, experts or non-experts, to comprehend the breadth of information that can be extracted from the EO-RDF data cubes.

[19] opengeospatial.org/standards/geosparql
[20] w3.org/2004/02/skos/
[21] rdflib.readthedocs.io/en/stable/
[22] openrefine.org/
[23] rml.io/specs/rml/
[24] tarql.github.io/
[25] graphdb.ontotext.com/

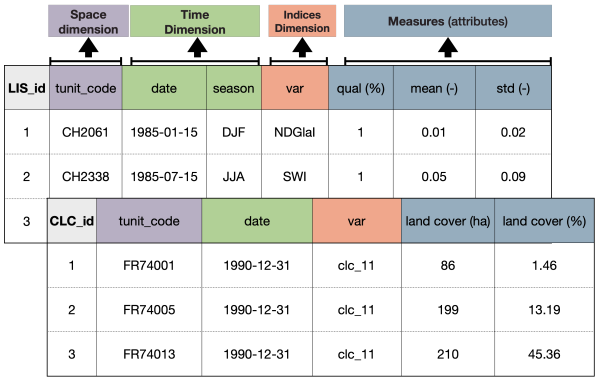
4 Implementing the LEODS framework

4.1 Study area selection process

The TRACES project partners are the creators of the Swiss Data Cube (SDC), a well-known EODC providing analysis-ready data on the geographical extent of Switzerland and part of France since 1984 (Giuliani et al., 2017). Therefore, by following the LEODS framework's processing chain, we initially identified three significant study areas within the SDC: Evian in France, Fribourg in Switzerland, and Grand-Geneva, situated along the border of both countries. Through preprocessing and spatial aggregation of the selected images, we obtained a total of 365 municipalities across the three study areas: 37 in Evian, 127 in Fribourg, and 201 in Grand-Geneva. Subsequently, we calculated three families of environmental indices: (a) Landsat land cover characterization indices (LIS), (b) Landsat land surface temperature indices (LST), and (c) Land cover indices based on the standard Corine nomenclature (CLC) (See Table 1). LIS and LST indices are accessible as seasonal data from 1985 to 2022, providing approximately 152 observations per index (4 observations per year over 38 years). Zonal statistics, including mean, std, and data quality features, were computed for both families. Contrarily, the CLC dataset only provides 5 observations for the years 1990, 2000, 2006, 2012, and 2018. Although lacking zonal statistics, it does offer calculations of land cover area occupation in hectares (HA) and percentages (%). After completing the processing steps the data was delivered as EO vector time-series in CSV files (Refer Figure 2). Further information on the indices can be found in our repository (Milon-Flores et al., 2024).

4.2 Conceptual model of EO-RDF data cubes

The second phase of our framework involves modeling the structure of EO-RDF data cubes. As outlined in Section 3, the Time and Space dimensions are inherent to evolving EO data. Hence, we aim to identify the tabular values that best correspond to the cube components. In Figure 2, we readily identify the `tunit_code` as the values aligning with the Spatial dimension. Specifically, this column denotes the codes assigned to the selected municipalities, where "FR" indicates French municipalities and "CH" indicates Swiss municipalities. To ensure seamless integration and enrichment within the cubes, we find it appropriate to associate the spatial dimension with the TSN ontology. Consequently, each municipality inherits properties such as `tsn:name` and `tsn:code`, enhancing its description. Continuing, we identify the concepts of `date` and `season` as time-related. However, since the standard OWL-Time ontology lacks seasonal concepts such as spring and summer, we create and extend these concepts in the ontology to fully represent the temporal dimension of the cube. Subsequently, we identify the `var` column as our domain-specific Indices dimension. During our research,



LIS_id	tunit_code	date	season	var	qual (%)	mean (-)	std (-)
1	CH2061	1985-01-15	DJF	NDGlal	1	0.01	0.02
2	CH2338	1985-07-15	JJA	SWI	1	0.05	0.09
3							

CLC_id	tunit_code	date	var	land cover (ha)	land cover (%)
1	FR74001	1990-12-31	clc_11	86	1.46
2	FR74005	1990-12-31	clc_11	199	13.19
3	FR74013	1990-12-31	clc_11	210	45.36

Figure 2. Identifying the QB components in the LIS and CLC tabular files. The column `tunit_code` denotes the municipality codes and is aligned with the spatial dimension. The columns `date` and `season` pertain to the temporal dimension. The columns `var` describe the indices obtained for each dataset. Correspondingly, `qual`, `mean` and `std` represent statistical measures derived from the LIS satellite indices aggregated at the municipality level. In contrast, `land cover (HA)` and `land cover (%)` denote, respectively, the number of hectares and the percentage of hectares within municipalities for the class `clc` in the CLC dataset.

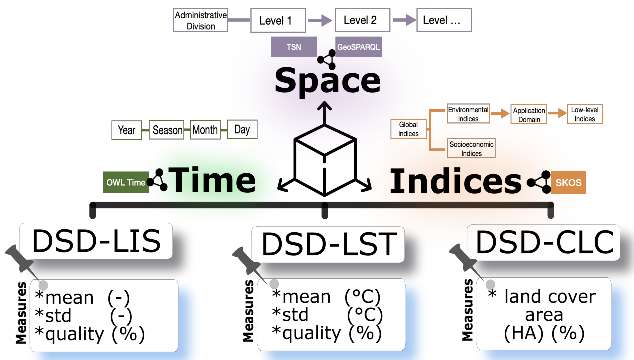


Figure 3. Final DSD design of the three EO-RDF data cubes.

we discovered the lack of an ontology dedicated to the description of satellite indices. Therefore, leveraging the SKOS vocabulary, we have developed and enriched a thesaurus covering the full range of indices. Later, we tailor each dimension to be hierarchical, meaning the Spatial dimension is organized as "municipalities, departments, and countries", the Temporal dimension is organized as "daily, monthly, seasonal, and annual", and the Thematic indices as "low-level indices, application domain, environmental indices, and overall indices". Note that special attention is dedicated to the Indices dimension to ensure its generality, allowing for the integration of indices beyond the scope of EO (*e.g.,* population, sensing instruments, etc.) into the cube. In identifying measures and attributes for CLC, we easily categorize the `Land Cover Area` measure with two units (HA and %). Whereas, for LIS and LST, we identify three common measures: mean, std, and data qual-

Environmental indices				
Indice family	**Application domain**	**Indice**	**Name**	**Total number of indices**
LIS	Vegetation	NDVI	Normalized Difference Vegetation Index	20
		...		
	Water	NDWI	Normalized Difference Water Index	
		...		
	Urbanization	NDBI	Normalized Difference Built-Up Index	
		...		
	Snow	NDSI	Normalized Difference Snow Index	
		...		
LST	Temperature	st	Surface Temperature	1
CLC	Land cover	clc-11	Urban fabric	15
		...		

Table 1. Three families integrate the selected environmental indices for our case study: LIS, LST and CLC.

ity, each associated with various attributes such as temperature in degrees Celsius (°C), percentage (%), and unitless values (-)[26]. For this reason, we instantiate three DSDs to preserve the diverse units within the same data cube. Figure 3 illustrates the final conceptual model of our EO-RDF data cubes, encompassing hierarchical dimensions, measures, attributes, and reused ontologies. Finally, metadata and LOD connections are incorporated to enrich and contextualize the cube components. A significant contribution lies in the provided metadata for the indices. Since there is limited information available on expert-oriented indices like NDBI on the LOD Cloud, descriptions, calculation formulas, and bibliographic citations have been added to semantically enrich each index. A similar process is conducted for the municipalities by establishing connections to LOD resources available in official statistical agencies. Specifically, Swiss municipalities are linked to the Federal Statistical Office (FSO)[27] SPARQL endpoint, while French municipalities are associated with data from both the National Institute of Statistics and Economic Studies (INSEE)[28] and the National Institute of Geographic Information (IGN), the latter being semantically described and hosted in the GeoChange repository[29]. As a result, besides enriching the EO data, the linking process also enables the acquisition of valuable information absent in the original dataset.

4.3 Production of EO-RDF data cubes

To instantiate the cubes semi-automatically, Python and Tarql scripts have been implemented and executed. The scripts specifically map values from tabular files to RDF files following the modeled structure. As a result, three EO-RDF data cubes, in addition to preserving all the advantages of traditional EODCs, now organize multidimensional Linked Data that integrates basic semantics, such as metadata and connections to LOD resources. This significantly contributes to their enrichment and reuse. The cubes have been stored in a GraphDB triple store, where users can access them through a SPARQL endpoint[30].

4.4 Spatio-temporal SPARQL queries for linked data exploration

The last step of the framework consists in exploring the produced EO-RDF data cubes. SPARQL and GeoSPARQL queries can be used to retrieve relevant information from them. As part of our methodology, we aim to provide stakeholders with predefined SPARQL queries that facilitate the exploration of the open link cubes. Although our current focus in this study is spatio-temporal queries, we have implemented different types of queries currently accessible in a GitHub repository (Milon-Flores et al., 2024). The repository hosts basic queries to retrieve cube components such as dimensions, measures and attributes, as well as more advanced queries that exploit the hierarchical components of the cubes to perform OLAP operations such as Drill-down and Roll-up. As an initial query, the Listing 1 code retrieves enriched metadata related to the indices. Specifically, Table 2 samples metadata for NDVI, including essential links to well-known LOD databases such as Wikidata (Vrandečić and Krötzsch, 2014) and DBpedia (Auer et al., 2007), along with a link to the GeoNetwork catalog provided by the TRACES team [31]. This enrichment process has been applied to all the indices investigated in this research, which represents a significant contribution of our work.

```
SELECT ?property ?metada
WHERE {
  ?TI skos:hasTopConcept ?TIenv.
    ?TIenv skos:narrower ?domain.
    ?domain  skos:narrower ?indices.
    ?indices ?property ?metada .
    FILTER (?indices = traces-codelist:NDVI)
}
```

Code Listing 1. SPARQL query that retrieves indices metadata.

property	metadata
owl:sameAs	wikidata.org/wiki/Q718775
dct:description	"NDVI is the most common vegetation index in remote ..."
dct:description	"Formula: (N - R)/(N + R) ; Bands: ["R","N"] ; Bands ..."
skos:exactmatch	traces-gn.unepgrid.ch/geonetwork/...
dct:bibliographicCitation	ntrs.nasa.gov/citations/19740022614

Table 2. Output of the SPARQL query presented in Listing 1.

Regarding the implemented spatio-temporal queries, in the following listings, we present two queries that demonstrate the strengths of our final EO-RDF data cube design by retrieving geo-spatial information not present in the original dataset and by manipulating all the dimensions of the data cubes. Specifically, Query 2 retrieves the geographic coordinates of the municipalities belonging to the Grand-Geneve case study (Table 3). Since we have enriched each

[26]LIS measurements are unitless because they are calculated as a ratio of differences between two spectral bands, and the units are canceled during the calculation.
[27]bfs.admin.ch/bfs/en/home.html
[28]insee.fr/fr/accueil
[29]geo.ld.admin.ch/query
[30]steamerlod.imag.fr/repositories/TRACES
[31]traces-gn.unepgrid.ch/geonetwork/srv/fre/catalog.search#

municipality with official LOD resources, we can connect to the SPARQL endpoints of FSO (Swiss municipalities) and GeoChange (French municipalities) and run a federated GeoSPARQL query to extract their spatial coordinates. The reader should note that this information was not present in the original dataset and that, due to the LOD connections, the data is no longer isolated. This is one of the fundamental principles of the LOD paradigm.

```
SELECT DISTINCT ?name ?code ?coordinates
WHERE {
?area sett:studyArea traces-geo:GrandGeneve;
    tsn:hasName ?name;
    tsn:hasIdentifier ?code;
    owl:sameAs ?bounderies.
SERVICE <steamerlod.imag.fr/repositories/
                              geochange?>
 {
    ?bounderies geo:hasGeometry ?geometryFR.
    ?geometryFR geo:asWKT ?coordinates.
 }
SERVICE <geo.ld.admin.ch/query>
 {
    ?bounderies geo:hasGeometry ?geometryCH.
    ?geometryCH geo:asWKT ?coordinates.
 }
}
```

Code Listing 2. GeoSPARQL query that retrieves the geographic coordinates of municipalities in Grand-Geneve.

name	code	coordinates
"Allinges"	"FR74005"	"POLYGON ((6.450841005360..))
"Armoy"	"FR74020"	"POLYGON ((6.450841005360..))
"Cervens"	"FR74053"	"POLYGON ((6.450841005360..))
...	...	...

Table 3. Output of the SPARQL query presented in Listing 2.

Later, we perform a slice-type query (Listing 3) in which the spatial dimension is fixed on the community of Grand-Geneve, the temporal dimension is fixed on the year 1985, and the indice dimension is fixed at NDVI to retrieve the average mean values of the selected slice (Table 4). Then, by using the Python Folium library[32], we can plot the retrieved coordinates and average NDVI values of the municipalities on a heat map. Thus, to generate the final chart depicted in Figure 4, we have run the query multiple times, each time varying the year value, specifically 1985, 2010, and 2022, to enable temporal comparison. The initial findings reveal a notable expansion of vegetation across various landscapes, notably in natural, agricultural, and mountainous areas, as indicated by the darker green coloring. Moreover, there is a significant increase in the average vegetation values, rising from a maximum of 50% to 65% in recent years. To better explain the underlying factors driving this vegetation surge, works such as (Obuchowicz et al., 2023) suggest exploring the correlation between NDVI and other indices, such as surface temperature.

```
SELECT ?name ?code (AVG(?mean) AS ?avgmean)
WHERE {
  ?obs rdf:type qb:Observation;
       :dataSet :Seasonaly-LIS-dataset;
       :dimensionArea ?area;
       :dimensionTime ?time;
       :dimensionIndice codelist:NDVI;
       :measureMeanUnitless ?mean.
?area sett:studyArea traces-geo:GrandGeneve;
  tsn:hasName ?name;
  tsn:hasIdentifier ?code.
?time time:year "1985"^^xsd:gYear.
}
GROUP BY ?name ?code
ORDER BY ?name
```

Code Listing 3. SPARQL query that calculates the average NDVI values in the municipalities of Grand-Geneve in 2022.

name	code	avgmean
"Aire-la-Ville"	"CH6601"	"0.40374833"^^xsd:float
"Allinges"	"FR74005"	"0.5516"^^xsd:float
"Amancy"	"FR74007"	"0.5443302"^^xsd:float
"Ambilly"	"FR74008"	"0.37145054"^^xsd:float
...	...	...

Table 4. Output of the SPARQL query presented in Listing 3.

Therefore, we have to perform a variation of the previous slice-type query by exploiting the LST dataset and calculating the mean surface temperature in the Grand-Geneve community during the same periods (*i.e.,* 1985, 2010, and 2022). Figure 5 illustrates that urban municipalities, depicted in dark red, have experienced a pronounced increase in temperature, while in mountainous regions the phenomenon, known as "Elevation Dependent Warming" (mou, 2015), is observed; *i.e.,* temperature increases are more significant at higher elevations as the climate warms. As for average temperature values, considerable changes have been observed over the years. The minimum temperature has risen from -2°C in 1985 to 5°C in 2022, while the maximum temperature has increased from 12°C to 17°C. In addition, a clear visual correlation is observed between NDVI and temperature variables. This correlation suggests that increased vegetation activity or productivity is observed as a consequence of increased temperatures; in other words, higher temperatures create more favorable conditions for vegetation growth.

To conclude the exploration phase, in addition to SPARQL queries, users can also exploit GraphDB's visual interface to analyze the stored data. Thus, we strongly encourage interested users to visually explore the EO-RDF data cubes to deepen their understanding. For illustration, Figure 6 provides an overall view of the cube structure with real EO data. In particular, we can observe how the observations are interconnected in the cube through its shared dimen-

[32] python-visualization.github.io/folium/latest/#

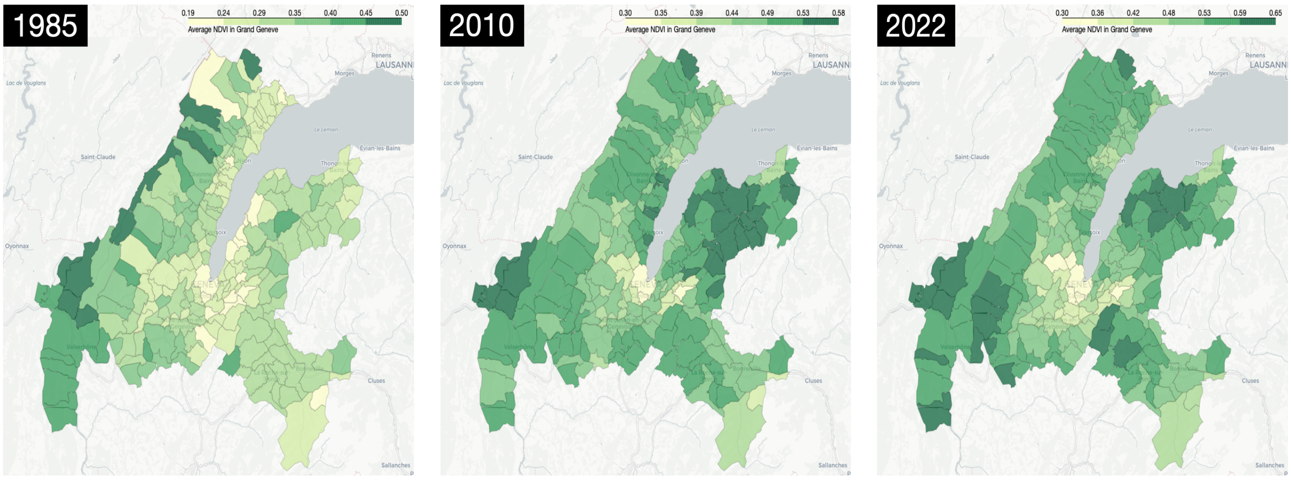


Figure 4. Average NDVI values in the municipalities of Grand-Geneve over the years 1985, 2010, and 2022 respectively.

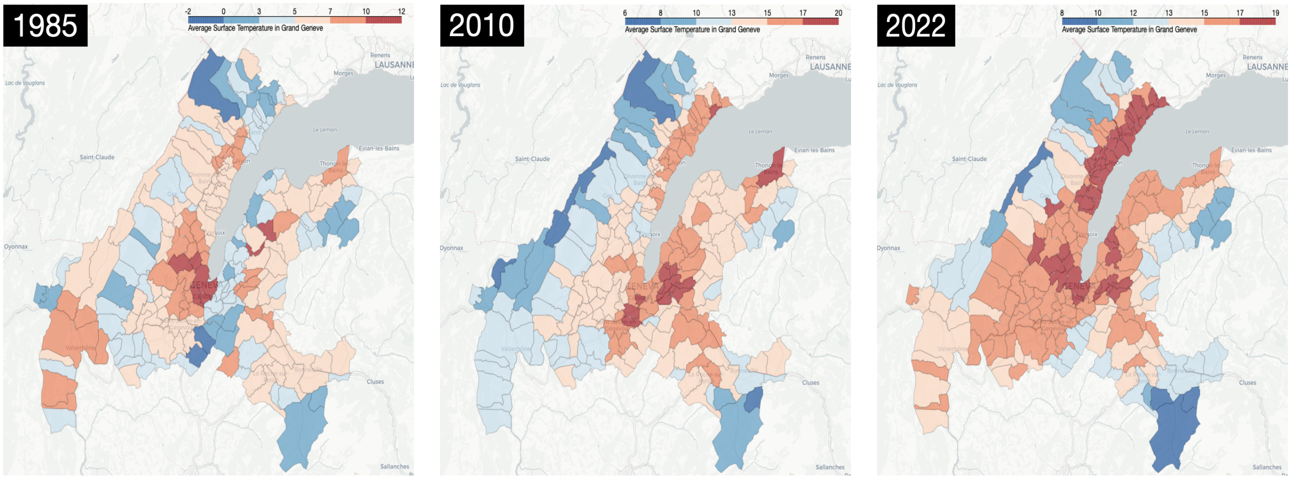


Figure 5. Average surface temperature in the municipalities of Grand-Geneve over the years 1985, 2010, and 2022 respectively.

sions, which are also connected to various LOD resources. Hence, through the effective utilization of queries and visualization, one can obtain insights that support the implementation of sustainable environmental practices and policies.

5 Discussion

5.1 Advantages and limitations of aggregating pixel EO data

The aggregation of EO data at the municipality level offers numerous advantages. First, the information obtained allows us to be as close as possible to the stakeholders, in particular citizens and policymakers interested in the environmental aspects of their municipality. In addition, by aggregating EO data, one can ensure its enrichment with local resources (such as legal text and urban planning), thus providing a more detailed characterization of the municipalities. Importantly, this aggregation facilitates comparative analyses of different municipalities over time, without incurring the high storage costs associated with pixel-level data storage. Furthermore, recent studies such as (Obuchowicz et al., 2023; Poussin et al., 2021) have successfully demonstrated the wealth of information that can be obtained at a national scale. These studies highlight the significance of capturing local environmental changes for informing policy decisions. However, one consequence of data aggregation is the loss of information. Common zonal statistics, such as standard deviation, may not be suitable for accurately analyzing the data. Therefore, specific strategies need to be employed to address this drawback. For instance, the incorporation of more sophisticated zonal statistics, including metrics such as z-score, mode, minimum, and maximum, can provide a more detailed picture of the environmental indices. Additionally, generating new EO-RDF data cubes at the Urban administration level can offer another avenue to compensate for information loss. Finally, in cases where users need more detailed data, they can access the pixels stored in the SDC through the enriched LOD connections of the indices (*i.e.,* by accessing the GeoNetwork catalogue). Whit these strategies, we aim to mitigate the limitation of our approach while

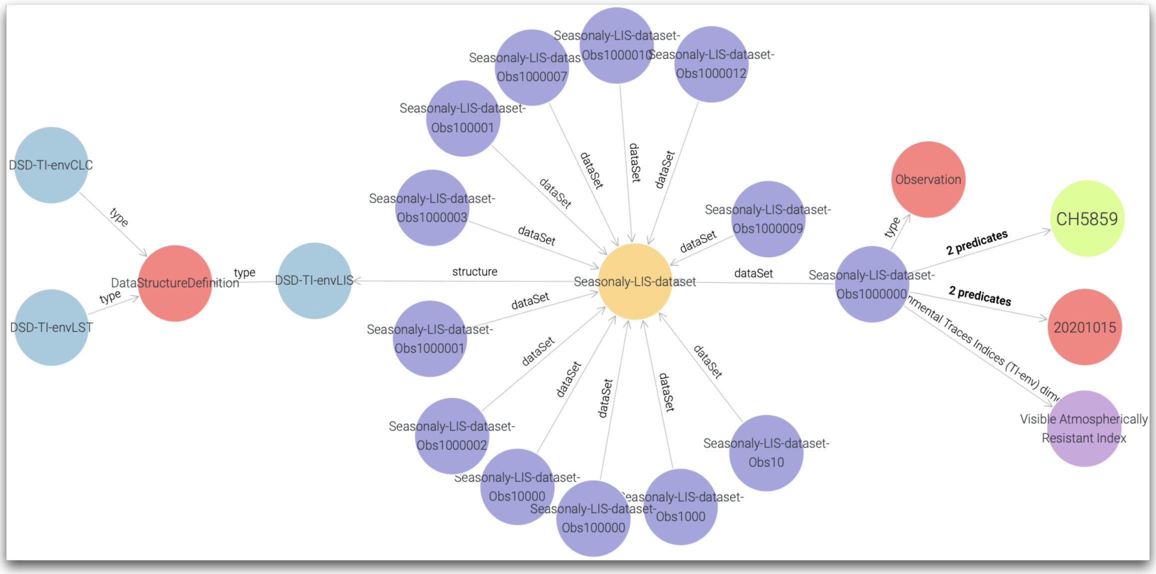


Figure 6. Overview of the EO-RDF data cubes structure. The three generated datasets are interconnected through their DSD. In turn, each dataset (*e.g.,* Seasonally-LIS-dataset) comprises a series of observations (*e.g.,* Obs1000000) with its corresponding dimensions, *i.e.,* space (CH5859), time (2020/10/15), and indices (Visible Atmospherically Resistant (VAR)).

maximizing the utility of aggregated EO data for various stakeholders.

5.2 Adaptability to dynamic municipal boundaries

The boundaries of geographic territories such as municipalities change over time. For this reason, we use our TSN and TSN-Change[33] ontologies for the description of territorial units and their changes over time (*e.g.,* name changes, code changes, mergers of two municipalities or divisions, etc.). This approach allows us to represent in the graph the evolutionary trajectories of these administrative units, along with their associated environmental changes. Furthermore, it is worth mentioning to interested users that, in this work, Swiss municipalities' boundaries correspond to the 2022 version, while French municipalities correspond to the 2019 version. Information related to the different versions of the municipalities such as population, boundaries coordinates, etc, can be accessed through SPARQL queries. Basic examples are currently available in our repository (Milon-Flores et al., 2024).

6 Conclusions and future work

In this paper, we presented the LEODS framework, focused on the integration and semantic enrichment of EO data in LOD Cloud. Through a processing chain, our framework describes how to embody EO data into open-linked data cubes. During the process, EO data, originally at the pixel level, was aggregated at the lowest administrative division level to be closer to stakeholders and ensure its contextualization with other local Web resources. Subsequently, a careful modeling of the cubes was conducted generically enough to ensure their future integration with heterogeneous data sharing spatiotemporal dimensions and with various indices beyond the EO domain. To illustrate the advantages of our approach, we have implemented the LEODS framework with real study areas. As a result, three EO-RDF data cubes, not longer isolated but enriched with metadata, connected to various LOD resources, following FAIR principles, and preserving the advantages of traditional EODCs are now available to stakeholders in a SPARQL endpoint. The structure of the cubes incorporates a model generic enough to guarantee its integration with other resources through its spatio-temporal dimensions, as well as to ensure the inclusion of indices beyond the scope of the EO. Furthermore, predefined queries were provided to stakeholders to facilitate their exploration and analysis of the data. We expect that both expert and non-expert users can benefit from our approach. As an example, the produced cubes, incorporating relevant information on the environmental aspects of the selected municipalities, can be consulted and explored by analysts, citizens, policymakers, or interested organizations to propose effective environmental policies. In future work, we aim to use the EO-RDF data cubes as a first step in creating a KG that describes the environmental trajectories of municipalities. Additionally, we intend to add more indices to our existing EO-RDF data cubes and to continue to enrich them with new LOD resources.

[33] lig-tdcge.imag.fr/tsnchange/index.html

7 Data and Software Availability

All three EO-RDF data cubes, in Turtle format, are available at github.com/DanielaFe7-personal/Traces-EO-RDF-data-cubes with a permissive license. Supplementary material, such as SPARQL queries and relevant information about the used dataset, is also available in the GitHub repository. In addition, the knowledge graph is available in an open SPARQL endpoint (steamerlod.imag.fr/repositories/TRACES) based on a publicly accessible GraphDB repository[34] (version 8.4.1).

8 Appendix

8.1 Glossary

AIM	Agricultural Information Model
CLC	Corin Land Cover
DSD	Data Structure Definition
EO	Earth Observation
EODC	Earth Observation Data Cube
FAIR	Findable, Accessible, Interoperable, and Reusable
FSO	Federal Statistical Office
HA	Hectares
INSEE	National Institute of Statistics and Economic Studies
IGN	National Institute of Geographic Information
IUCN	International Union for Conservation of Nature
IPCC	Intergovernmental Panel on Climate Change
KG	Knowledge Graph
LEODS	Linked Earth Observation Data Series
LIS	Landsat land cover indices
LOD	Linked Open Data
LST	Landsat land surface temperature indices
NDBI	Normalized Difference Built-Up Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NDSI	Normalized Difference Snow Index
OGC	Open Geospatial Consortium
OWL	Web Ontology Language
QB	RDF data cube
RDF	Resource Description Framework
SDC	Swiss Data Cube
SDMX	Statistical Data and Metadata eXchange
SDW	Spatial Data on the Web
SKOS	Simple Knowledge Organisation System
SOSA	Sensor, Observation, Sampler, and Actuator
SSN	Semantic Sensor Network
ST	Surface Temperature
SW	Semantic Web
TSN	Territorial Statistical Nomenclature Ontology
VAR	Visible Atmospherically Resistant Index
W3C	World Wide Web Consortium

[34] http://steamerlod.imag.fr/

References

Elevation-dependent warming in mountain regions of the world, Nature climate change, 5, 424–430, 2015.

Appel, M. and Pebesma, E.: On-demand processing of data cubes from satellite image collections with the gdalcubes library, Data, 4, 92, 2019.

Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., and Ives, Z.: DBpedia: A Nucleus for a Web of Open Data, in: The Semantic Web, edited by Aberer, K., Choi, K.-S., Noy, N., Allemang, D., Lee, K.-I., Nixon, L., Golbeck, J., Mika, P., Maynard, D., Mizoguchi, R., Schreiber, G., and Cudré-Mauroux, P., pp. 722–735, Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

Augustin, H., Sudmanns, M., Tiede, D., Lang, S., and Baraldi, A.: Semantic Earth observation data cubes, Data, 4, 102, 2019.

Ayadi, N. Y., Faron, C., Michel, F., Gandon, F., and Corby, O.: A Model for Meteorological Knowledge Graphs: Application to Météo-France Data, in: ICWE 2022-22nd International Conference on Web Engineering, 2022.

Baumann, P.: The Datacube Manifesto, available at earthserver.eu/tech/datacube-manifesto, 2017.

Bayerl, S. and Granitzer, M.: Data-transformation on historical data using the RDF data cube vocabulary, in: Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business, pp. 1–8, 2015.

Bernard, C., Villanova-Oliver, M., Gensel, J., and Dao, H.: Modeling Changes in Territorial Partitions Over Time: Ontologies TSN and TSN-change, in: Proceedings of the 33rd Annual ACM Symposium on Applied Computing, SAC '18, pp. 866–875, ACM, https://doi.org/10.1145/3167132.3167227, 2018.

Brizhinev, D., Toyer, S., Taylor, K., and Zhang, Z.: Publishing and using Earth observation data with the RDF data cube and the discrete global grid system, W3C Working Group Note and OGC Discussion Paper W3C, 20170928, 16–125, 2017.

Casey, S., Doody, P., and Shields, A.: An Ontology-Based System for Cancer Registry Data, in: 2022 33rd Irish Signals and Systems Conference (ISSC), pp. 1–6, https://doi.org/10.1109/ISSC55427.2022.9826197, 2022.

Cyganiak, R., Wood, D., and Lanthaler, M.: RDF 1.1 Concepts and Abstract Syntax, https://www.w3.org/TR/rdf11-concepts/, 2014.

Datcu, M., Daschiel, H., Pelizzari, A., Quartulli, M., Galoppo, A., Colapicchioni, A., Pastori, M., Seidel, K., Marchetti, P., and D'Elia, S.: Information mining in remote sensing image archives: system concepts, 41, 2923–2936, https://doi.org/10.1109/TGRS.2003.817197, 2003.

de Sousa, L. M.: Spatial Linked Data Infrastructures, Zenodo, https://doi.org/10.5281/zenodo.10113504, 2023.

Giuliani, G., Chatenoux, B., De Bono, A., Rodila, D., Richard, J.-P., Allenbach, K., Dao, H., and 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, 1, 100–117, 2017.

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

Hogan, A.: The Semantic Web: Two decades on, Semantic Web, 11, 169–185, https://doi.org/10.3233/SW-190387, 2020.

Koubarakis, M. and Kyzirakos, K.: Modeling and querying metadata in the semantic sensor web: The model stRDF and the query language stSPARQL, in: Extended Semantic Web Conference, pp. 425–439, Springer, 2010.

Koubarakis, M., Kyzirakos, K., Nikolaou, C., Garbis, G., Bereta, K., Smeros, P., Gianakopoulou, S., Dogani, K., Karpathiotaki, M., and Vlachopoulos, I.: Linked Earth Observation Data: The Projects TELEIOS and LEO, in: Proceedings of the Linking Geospatial Data Conference, 2014.

Koubarakis, M., Kyzirakos, K., Nikolaou, C., Garbis, G., Bereta, K., Dogani, R., Giannakopoulou, S., Smeros, P., Savva, D., Stamoulis, G., et al.: Managing big, linked, and open earth-observation data: Using the teleios\/leo software stack, IEEE Geoscience and Remote Sensing Magazine, 4, 23–37, 2016.

Kyzirakos, K., Savva, D., Vlachopoulos, I., Vasileiou, A., Karalis, N., Koubarakis, M., and Manegold, S.: GeoTriples: Transforming geospatial data into RDF graphs using R2RML and RML mappings, Journal of Web Semantics, 52, 16–32, 2018.

Lefort, L., Bobruk, J., Haller, A., Taylor, K., and Woolf, A.: A Linked Sensor Data Cube for a 100 Year Homogenised Daily Temperature Dataset., in: SSN, pp. 1–16, 2012.

Leroux, H. and Lefort, L.: Using CDISC ODM and the RDF Data Cube for the Semantic Enrichment of Longitudinal Clinical Trial Data., in: SWAT4LS, 2012.

Lewis, A., Oliver, 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 Sensing of Environment, 202, 276–292, 2017.

Milon-Flores, D. F., Bernard, C., Gensel, J., Giuliani, G., Chatenoux, B., and Dao, H.: LEODS framework repository, https://github.com/DanielaFe7-personal/Traces-EO-RDF-data-cubes, 2024.

Nativi, S., Mazzetti, P., and Craglia, M.: A view-based model of data-cube to support big earth data systems interoperability, Big Earth Data, 1, 75–99, 2017.

Nikolaou, C., Dogani, K., Bereta, K., Garbis, G., Karpathiotakis, M., Kyzirakos, K., and Koubarakis, M.: Sextant: Visualizing time-evolving linked geospatial data, Journal of Web Semantics, 35, 35–52, 2015.

Obuchowicz, C., Poussin, C., and Giuliani, G.: Change in observed long-term greening across Switzerland–evidence from a three decades NDVI time-series and its relationship with climate and land cover factors, Big Earth Data, pp. 1–32, 2023.

Palma, R., Roussaki, I., Döhmen, T., Atkinson, R., Brahma, S., Lange, C., Routis, G., Plociennik, M., and Mueller, S.: Agricultural Information Model, in: Information and Communication Technologies for Agriculture—Theme III: Decision, pp. 3–36, Springer, 2022.

Patel, A. and Jain, S.: Present and future of semantic web technologies: a research statement, International Journal of Computers and Applications, 43, 413–422, https://doi.org/10.1080/1206212X.2019.1570666, 2021.

Poussin, C., Massot, A., Ginzler, C., Weber, D., Chatenoux, B., Lacroix, P., Piller, T., Nguyen, L., and Giuliani, G.: Drying conditions in Switzerland–indication from a 35-year Landsat time-series analysis of vegetation water content estimates to support SDGs, Big Earth Data, 5, 445–475, 2021.

Richard Cyganiak, D. R. and Tennison, J.: The RDF Data Cube Vocabulary, W3c recommendation, W3C, https://www.w3.org/TR/vocab-data-cube/, 2014.

Rodriguez, T. N. and Hogan, A.: COVIDCube: An RDF Data Cube for Exploring Among-Country COVID-19 Correlations, vol. 2980 of *CEUR Workshop Proceedings*, CEUR-WS.org, http://ceur-ws.org/Vol-2980/paper395.pdf, 2021.

Salomón, R. L., Peters, R. L., Zweifel, R., Sass-Klaassen, U. G., Stegehuis, A. I., Smiljanic, M., Poyatos, R., Babst, F., Cienciala, E., Fonti, P., et al.: The 2018 European heatwave led to stem dehydration but not to consistent growth reductions in forests, Nature communications, 13, 28, 2022.

Schuldt, B., Buras, A., Arend, M., Vitasse, Y., Beierkuhnlein, C., Damm, A., Gharun, M., Grams, T. E., Hauck, M., Hajek, P., et al.: A first assessment of the impact of the extreme 2018 summer drought on Central European forests, Basic and Applied Ecology, 45, 86–103, 2020.

Simoes, R., Camara, G., Queiroz, G., Souza, F., Andrade, P. R., Santos, L., Carvalho, A., and Ferreira, K.: Satellite image time series analysis for big earth observation data, 13, 2428, publisher: MDPI, 2021.

Sudmanns, M., Augustin, H., van der Meer, L., Baraldi, A., and Tiede, D.: The Austrian semantic EO data cube infrastructure, Remote Sensing, 13, 4807, 2021.

Tran, B.-H., Aussenac-Gilles, N., Comparot, C., and Trojahn, C.: An Approach for Integrating Earth Observation, Change Detection and Contextual Data for Semantic Search, pp. 3115–3118, IEEE, 2020a.

Tran, B.-H., Aussenac-Gilles, N., Comparot, C., and Trojahn, C.: Semantic integration of raster data for earth observation: An RDF dataset of territorial unit versions with their land cover, ISPRS International Journal of Geo-Information, 9, 503, 2020b.

Ubaldi, B.: Open government data: Towards empirical analysis of open government data initiatives, 2013.

Van Der Meer, L., Sudmanns, M., Augustin, H., Baraldi, A., and Tiede, D.: SEMANTIC QUERYING IN EARTH OBSERVATION DATA CUBES, XLVIII-4/W1-2022, 503–510, https://doi.org/10.5194/isprs-archives-XLVIII-4-W1-2022-503-2022, 2022.

Vrandečić, D. and Krötzsch, M.: Wikidata: A Free Collaborative Knowledgebase, 57, 78–85, https://doi.org/10.1145/2629489, 2014.