

A SEMANTIC MODEL FOR ON-DEMAND EARTH OBSERVATION PRODUCTS

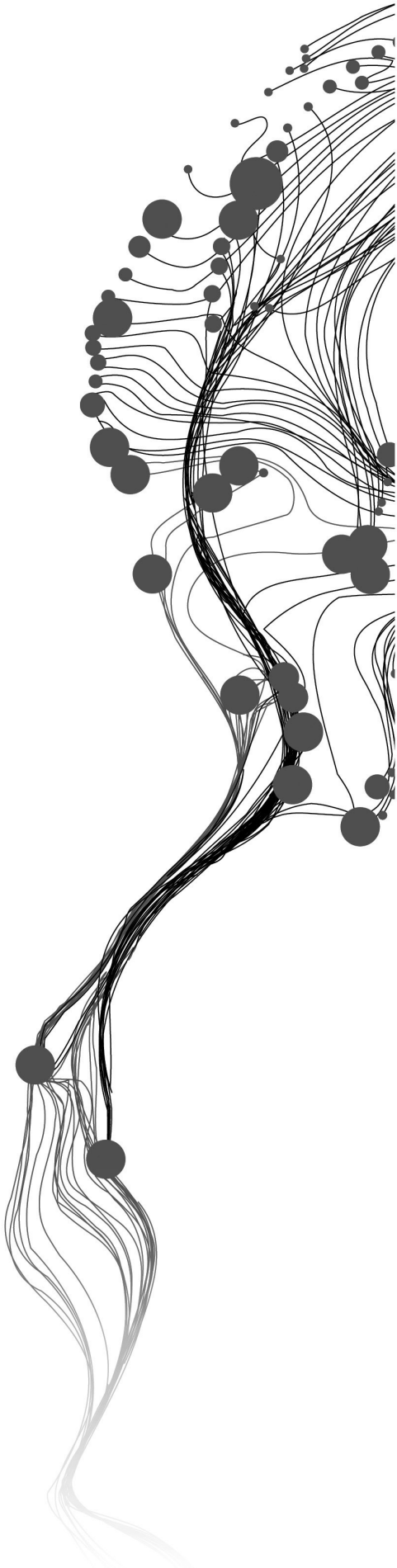
ANAM AKHTAR
July, 2021

SUPERVISORS:

Dr. Ir. Rob Lemmens, ITC GIP
Dr. M. Belgiu , ITC EOS

EXTERNAL GUIDE:

Sirko Schindler, German Aerospace Center (DLR), Germany



A SEMANTIC MODEL FOR ON-DEMAND EARTH OBSERVATION PRODUCTS

ANAM AKHTAR
Enschede, The Netherlands, July,
2021

Thesis submitted to the Faculty of Geo-information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Geoinformatics

SUPERVISORS:

Dr. Ir. Rob Lemmens, ITC GIP
Dr. M. Belgiu , ITC EOS

EXTERNAL GUIDE:

Sirko Schindler, German Aerospace Center (DLR), Germany

THESIS ASSESSMENT BOARD:

Dr. Ir. R.A. de By (chair)
Danny Vandenbroucke (External Examiner, KU Leuven)

Disclaimer

This document describes work undertaken as part of a programme of study at the Faculty of Geo-information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

ABSTRACT

Geoportals provide raw Earth Observation (EO) data which is subsequently processed to be used as EO products. One example of these products are spectral indices. Currently, geoportals provide only few spectral indices in their database although much more have been developed to serve various domains like water resources, urban planning, geology, etc. Decisions about which spectral indices to precompute is based on available storage and computing capacity as well as on previous demand. On-demand-computation can enable access to all the spectral indices without the need to precompute and store large amounts of EO products. For this purpose, comprehensive descriptions are needed that include all requirements and instructions to compute a spectral index.

The research thesis identifies the different components required to build a semantic model for on-demand geoprocessing of spectral indices. For this, the mathematical formula for each spectral index is mentioned with their dependencies identified. This model has a vision of calculating the spectral indices as per user's demand. To identify the user's demand, competency questions are formulated. This further describes the various ways, the user can interact with the semantic model. As a result of competency questions, it was concluded that the model not only requires the formula but also the raw EO datasets which was populated as an ontology with the help of Open Geospatial Consortium (OGC):17-003r2 specification. One added functionality was that a user can also search for a spectral index via applications. This part is displayed in the EO4GEO BoK which was verified through a survey. This survey resulted in a positive response from the participants. Further, the ability of the model to extract the required product was validated by posing SPARQL queries to the formulated competency questions. The model was able to give the requested spectral index and their parameters as a result of user's requirement. It is concluded that this model is successful in achieving its goal of describing the calculation of spectral indices explicitly.

Keywords

Ontology, Knowledge Graph, Spectral Indices, Semantics, Earth Observation Products

ACKNOWLEDGEMENTS

This research thesis is dedicated to my parents. My whole journey of masters is a success because of their unparalleled support.

I would like to sincerely thanks my supervisors - Dr. Ir. Rob Lemmens and Dr. Mariana Belgiu, for their guidance throughout my research phase. I am glad to be their student as they provided me assistance with patience and constructive advises which helped to sharpen my research.

I would like to extend my deepest gratitude to Sirko Schindler. His practical suggestions and insights have added value to this research. Thank you for believing in my abilities and for being such a generous mentor.

I gratefully acknowledge the participants who gave their time for the survey needed for this research. Special mention to Belinda and Jeroen, who have always helped me in almost all the administrative stuff at ITC. I am grateful for the professors who taught me here and my classmates who are a bunch of lovely people.

My sister and my close friends who were a call away or a just train away, have helped me alot to stay strong during my whole stay here. My life at ITC was an eye opener to lot of things and therefore, I am grateful to God for giving me this opportunity to be a class of 2021 of ITC. And most of all I am happy because God took me to places even before my dreams could reach.

TABLE OF CONTENTS

1	Introduction	1
1.1	Motivation and Problem Statement	1
1.2	Research Identification	3
1.3	Innovation	4
1.4	Thesis Structure	4
2	Literature review	5
2.1	Earth Observation data portals	5
2.2	Semantic Description	7
2.3	Fundamentals of Semantic Structure	9
2.4	Semantic Modelling	11
2.5	Related Work	13
2.5.1	Ontology Design	13
2.5.2	Application based semantic model	14
2.5.3	OGC Standard for EO Products	14
3	Implementation of Semantic Model	16
3.1	Requirement Analysis	16
3.1.1	Spectral Indices analysis	17
3.1.2	Use Case	17
3.1.3	Competency Question Analysis	20
3.1.4	Concept Identification	20
3.2	Building the T-box of Semantic model	20
3.3	Implementing the Model as a Linked Data Vocabulary	26
3.4	Building the A-box of Semantic model	27
3.5	Validation and Verification of the Semantic Model	33
4	Analysis & Discussion	35
4.1	Validation of the dependencies of the semantic model	35
4.2	Verification of the thematic domain of spectral indices	44
4.3	Answers to the research questions	47
4.4	Limitations	49
5	Conclusion and Outlook	50
	Appendices	55
A	Semantic model Survey	56
B	T-Box for semantic model	61
C	A-Box for semantic model	65
D	CQs as Displayed in the GraphDB Interface	67

LIST OF FIGURES

2.1	Timeline of technical solutions with respect to their degree of Interactivity (from Sudmanns et al., 2020).	6
2.2	Semantic Web architecture. (from Berners-Lee, 1999)	7
2.3	RDF Triple.	8
2.4	Example scenario for the Semantic Web. (from Maedche et al., 2002)	8
2.5	An example for RDF Schema.	10
2.6	A generalized example of steps involved in building a semantic layer.	11
2.7	Steps of the ontology design methodology. (from Hüsemann et al., 2005)	12
2.8	Ontology design for hydrologic monitoring. (from Wang et al., 2017)	13
2.9	Conformance and requirements class dependencies. (from Coene et al., 2020)	15
3.1	Workflow	16
3.2	Use case example: analyzing crop harvesting periods.	19
3.3	T-BOX to describe the components of spectral Indices.	21
3.4	T-Box to describe the thematic domains of the spectral indices.	22
3.5	T-Box for raw EO data. Derived from OGC: 17-003r2 (Coene et al., 2020).	24
3.6	Mathematical representation of NDVI in RDF form.	26
3.7	The dependencies for calculating NDVI.	28
3.8	The dependencies for calculating TSAVI. All the literals of Required Wavelength are omitted here for visual representation.	29
3.9	The dependencies for calculating MSAVI.	30
3.10	The dependencies for calculating CARI.	31
3.11	The interface for EO4GO Bok.	34
3.12	The connection of concepts of this semantic model and EO4GEO BoK.	34
4.1	Number of participants and their expertise.	45
4.2	Number of participants and their feedback to Q2 of Section 3.	46
A.1	As displayed in the survey	57
D.1	CQ1 as displayed on GraphDB.	67
D.2	CQ2 as displayed on GraphDB.	68
D.3	CQ3 as displayed on GraphDB.	68
D.4	CQ4 as displayed on GraphDB.	69
D.5	CQ5 as displayed on GraphDB.	70
D.6	CQ6 part 1 as displayed on GraphDB.	71
D.7	CQ6 part 2 as displayed on GraphDB.	72
D.8	CQ7 part 1 as displayed on GraphDB.	73
D.9	CQ7 part 2 as displayed on GraphDB.	74
D.10	CQ8 as displayed on GraphDB.	75
D.11	CQ9 as displayed on GraphDB.	76

LIST OF TABLES

3.1	Examples of spectral indices descriptions obtained through various sources. The displayed platform and sensor for each index is just an example and does not mean that it is the only option.	18
3.2	Concepts identified for the model depending on the competency questions. . . .	21
3.3	Object Properties of semantic model	25
3.4	Data Properties of semantic model.	25
3.5	Annotation Properties of semantic model	26
3.6	Spectral Indices associated with tasks.	32
3.7	Tasks associated with themes.	32
4.1	Input raw EO data.	36
4.2	Results for CQ1	37
4.3	Results for CQ2	37
4.4	Results for CQ3	38
4.5	Results for CQ4	39
4.6	Results for CQ5. Empty bindings omitted.	40
4.7	Results for CQ6 part 1	40
4.8	Results for CQ6 part 2	41
4.9	Results for CQ7 part 1	42
4.10	Results for CQ7 part 2	42
4.11	Results for CQ8	43
4.12	Results for CQ9	44

Chapter 1

Introduction

Our planet Earth has always been under watch since the evolution of a new era. This new era began in 1957 when Sputnik became the first satellite to be sent to space. Since then, thousands of satellites have followed (Union of Concerned Scientists, 2021). Later International Space Station (ISS) was launched in 1998 which became the largest single unit revolving around the earth. Today, more than 60 years later the advancement in science and technology by humans has given new hopes for the upcoming generation. It has brought the wonders of the universe to the palm of our hands. The satellites have improved the way of our living. The data from these platforms have helped humans in various fields of application like crop monitoring, water management, pollution monitoring, forest management, etc. The major concern is to enable data usage in such a way that it is useful for both the end-users and the data providers.

There are a lot of processes required from capturing the data through various platforms to providing it to end-users. It has always been a discussion between the scientists, developers, and analysts, etc. about processing, storing, and providing the EO data (Doldrina, 2015, Sudmanns et al., 2020). This data is usually provided on portals, which are websites providing multiple datasets on a single platform (Rose, 2004). For human interaction with this data, geospatial portals or geoportals are used. A geoportal "is a human interface to a collection of online geospatial information resources, including data sets and services" (Rose, 2004). Through a geoportal, end-user can interact with the EO data and products. This thesis focuses on providing EO data on geoportals that can be beneficial for user-centric purposes.

1.1 MOTIVATION AND PROBLEM STATEMENT

Every second data is collected by sensors onboard through various platforms like satellites, Unmanned Aerial Vehicles (UAVs), or ISS. The huge amount of geospatial data produced by these platforms needs proper and secure storage in the archives especially the ones acquired long ago as it is valuable for the future use. Moreover, the data can be further processed and is later known as derived or secondary products. These products are processed at different levels before being provided to the user. Examples of secondary products are National Land Cover Dataset¹, LANDFIRE², Global Land Survey³, or Land Surface Reflectance⁴ derived from land imagery of United States Geological Survey (USGS) (National Research Council, 2013). These products are a combination of different datasets from various sensors. Among these products are spectral indices which form the core to many EO services like water quality monitoring (Shafique et al., 2003) or detecting fire zones (Fornacca et al., 2018). For example, Normalized Differential Vegetation Index (NDVI) is used in agriculture for crop phenology monitoring (Boori et al., 2019) or for irrigation purposes (Sadgir et al., 2017).

¹<https://www.usgs.gov/centers/eros/science/national-land-cover-database>

²<https://nimbus.cr.usgs.gov/landing/>

³<https://www.usgs.gov/core-science-systems/nli/landsat/global-land-survey-gls>

⁴<https://www.usgs.gov/core-science-systems/nli/landsat/landsat-surface-reflectance>

This research focuses on spectral indices as the derived products of EO data. The information about the spectral indices is present on various online platforms like Land Surface Reflectance⁵ by USGS, Index Database⁶ (IDB), or L3Harris Geospatial⁷. Currently⁸, Land Surface Reflectance by USGS provide only seven spectral indices namely - NDVI, Enhanced Vegetation Index (EVI), Soil Adjusted Vegetation Index (SAVI), Modified Soil Adjusted Vegetation Index (MSAVI), Normalized Difference Moisture Index (NDMI), Normalized Burn Ratio (NBR), and Normalized Burn Ratio 2 (NBR2). L3Harris Geospatial provides a list of approximately 70 spectral indices whereas IDB, besides listing spectral indices, also connects them to sensors and possible applications. There are more than 500 spectral indices (Henrich et al., 2012). But it does not mean it is a complete list of spectral indices present today. However, most of them are not even included in the database of any organization. Even the existing spectral indices are generally hardcoded, this means that the indices are computed using pre-defined datasets and parameters. This can hamper the goal of geoportals as it reduces the user interaction. The geoportals should aim to maximize the user interaction by allowing them to set various constraints like spatial, temporal resolution, or even particular platform or sensor. Most of the geoportals allow such environment but not necessarily it meets all the requirements of the user. For example, a user requires spectral indices related to forestry for a particular area and sensor like Sentinel but the product available is hardcoded with another platform like Landsat for that area. Therefore, such a setting hampers the quality (in terms of demand) of the product required by the user. Also, mostly the products on geoportals are computed based on experience and previous usage of the EO products. This can sometime lead to data wastage when the user does not require some constraints imposed on the pre-computed data.

While making products available to users, data providers always strive to make their use efficient. Data providers of the geoportals have a huge amount of satellite-derived data ingested into their database. These include comprehensive details about their sensors, spectral resolution, or spatial-temporal resolutions (Dlamini et al., 2019). Oftentimes, non-expert users find it difficult to search through all these exhaustive data. Even an expert user has a hard time due to heterogeneity and sheer volume of the products. Thus, a user is oftentimes not able to find the required product on these geoportals even if they are present. This leads to wastage of resources which ultimately affects the storage capacity as a lot of products are precomputed by the system and stored in the database. A system here refers to an environment which is set by data providers using fixed algorithms. These algorithms can be, for example, to select a particular parameter like spectral range and a particular sensor to create a certain product. This is what makes the system static, and derived data created/stored in such a manner, limits the user as data providers can not simply create all possible products, but have to prioritize due to finite storage/ computing capacity. Therefore, on-demand geoprocessing of derived products is envisioned where the user will be able to choose the product to be executed by the system. This would save computing time, storage capacities, and at the same time broaden the range of available EO products.

The above can be achieved if the system is interactive to the user. For this, a semantic model of the derived products is proposed in this research. "A semantic model acts as a sort of glue between disparate, federated data sources, so they can be described as how they fit together" (Allemang et al., 2011). In other words, a semantic model is used to maintain a database containing information from different sources which is both machine readable and interactive with users. Data stored in this manner not only benefits on-demand geoprocessing of the derived products but also helps in linking to other resources. Semantic model can link different sources through 'Linked Data' (LD). "The term Linked Data refers to a set of best practices for publishing and connecting structured

⁵<https://www.usgs.gov/core-science-systems/nli/landsat/landsat-surface-reflectance-derived-spectral-indices>

⁶<https://www.indexdatabase.de/db/i.php>

⁷<https://www.l3harrisgeospatial.com/docs/alphabeticallistspectralindices.html>

⁸As of July 2021

data on the Web" (Bizer et al., 2011). An example of linked data can be seen on Linked Open Data (LOD) cloud⁹ which is a web of data linked together by several communities together. The linked data can be helpful here as different applications can be linked to different platforms like Landsat, or Sentinel, etc. Through these links a user can easily access the metadata of a platform which can provide him with more information about it like status (online/offline), its date of launch, or sensor onboard. etc. Overall, semantic models is an added value as it benefits both the users as well as the providers of EO products.

1.2 RESEARCH IDENTIFICATION

The semantic model developed here is a representation for populating several spectral indices which will be able to describe the inputs needed (like sensor type or application area) and their dependencies to compute the desired index. The user will choose an index from a catalogue of indices and the system will be able to describe an algorithm to compute the index. For explanation, assume that a user poses a query that leads to NDVI, the system will return an instruction set that conveys the user to load, for example, Sentinel 2 data, which can be found at a particular location, and finally use the band 8A (near-infrared) and band 4 (red) of that dataset to calculate the NDVI. The system will be able to extract knowledge from the semantic model. This knowledge will contain information about the indices and the raw EO data. The raw EO data is the information about the sensors, platforms, spatial and temporal resolution, or spectral characteristics. When all the requirements of a request for a particular index are present in the knowledge base, it will give information about the dependencies for calculation of the requested index. These dependencies can be spectral bands, or constants used to describe the formula of spectral indices. Such as an index can be best computed for a particular spectral range and hence, the list of sensors available to compute it. The semantic model developed can enable the on-demand geocomputing of the spectral indices.

The main objective of this research is to develop a semantic model for on-demand computation of indices using raw EO data.

Based on the above goal, the following research objectives and questions were formulated:

1. To develop a semantic model which can represent explicitly, the given information of indices and raw EO data in the knowledge base.
 - 1.1. What characteristics are available for the raw data?
 - 1.2. What requirements do the individual indices have?
 - 1.3. Are there other parameters involved that might not be determined automatically and require further user interaction?
 - 1.4. Can the model be provided with meaningful default values for those parameters?
2. To interpret the above knowledge into a semantic model which describes all the components required for it .
 - 2.1. What are the different components identified for the semantic model?
 - 2.2. What are the logical descriptions behind the above mentioned parts?
 - 2.3. What are the challenges in representing the semantic model?

⁹<https://www.lod-cloud.net/#>

3. To evaluate whether the semantic model can provide information about the required index selected by the user.
 - 3.1. In what ways, the user can interact with the semantic model that gives a meaningful output?
 - 3.2. Which parts can be fully automated and require no user input?
 - 3.3. What all information can the semantic model provide to its user?

1.3 INNOVATION

A semantic model for spectral indices will be developed in this research. Previously, there have been semantic models, for example by Wang et al.(2017) or Pareja-Lora(2012), in the field of EO but none that focuses mainly on spectral indices. Even if it does like in Wang et al.(2016), it does not deal with such an exhaustive list of spectral indices. It is suitable to add information for any index present in IDB database. This model describes the indices and their dependencies to develop a meaningful output. Here, the output would be the requested index as per user needs. This means the user can search for a particular index and define its spatial and temporal constraints. Once an index is selected, the necessary instructions can be triggered to retrieve the information about the spectral index. Here, the focus is mainly on non-expert users and therefore, they do not have to know the details involved in developing the product. The model requires the information about raw EO data which has been populated using the already developed model by OGC (Coene et al., 2020).

Further, the system will focus on the workflow required to generate the desired product. However, there is also need for the system to represent the loose coupling of raw data inputs. This is done to make the system flexible and provides better coverage or resolution, etc. as there are more source datasets to choose from.

1.4 THESIS STRUCTURE

The research thesis is divided into five chapters. The first chapter gives a background to the research problem and the research motivation. It also introduces the research objectives and research questions. The second chapter is Literature Review. It introduces to the concepts related to semantic model which are used later in the implementation chapter. It also displays the work already done before and their research gap. In general it gives the fundamental knowledge for building any semantic model. In the third chapter, the implementation of the semantic model is done. The model identifies the different components which is used for describing the calculation of spectral indices. The fourth is the analysis and discussion chapter which analyzes the semantic model through SPARQL queries and survey. It also includes the answers to the research questions. The fifth chapter is the conclusion and recommendation for future work.

Chapter 2

Literature review

The following concepts were studied to build the foundation of the proposed semantic model. The overview of the model has already been discussed and here has some key elements or theories and related work for the model.

2.1 EARTH OBSERVATION DATA PORTALS

The first geoportal was considered to be National Spatial Data Infrastructure (NSDI) initiated by Federal Geospatial Data Committee (FGDC), United States (US) in 1994. The datasets here was stored in a distributed database (maintained at different locations instead at a single place) which can be accessed by an interface. Later in 2003, Geospatial One-Stop (GOS) came as a web geoportal with metadata whose focus was more on the users and, therefore, it added a community where users can discuss issues with datasets. This made the user's requirement for a dataset comprehensible to the data providers. Recently, the geoportals have started distributing data based on themes such as in INSPIRE geoportal¹. INSPIRE IS provided by European Union (EU) Member States with themes like building, landuse, geology and many more in their knowledge base². Similarly, new features keep getting added with the emergence of new geoportals.

The EO data on these geoportals have been increasingly used for monitoring, detecting, and predicting various ecosystems. These ecosystems can be forests, deserts, mountains, or oceans. This shows the capability of Earth Observation (EO) technology and their far-reaching limits. These technologies are exhausted in capturing, storing, processing, and providing the data to the users. The data obtained from satellites are usually raw datasets which need processing like atmospheric corrections, geometric corrections, cloud or snow coverage description. After the processing, the data is then further processed which include combining different datasets into a single product known as a derived product of the EO data. The challenge for every geo-data provider is to enable best data access and data utilization for the users.

At present, there are various geoportals owned by international/ national organizations like Earth Data³ by the National Aeronautics and Space Administration (NASA), Earth Online⁴ by the European Space Agency (ESA), or EOWEB⁵ by Deutsche Zentrum für Luft- und Raumfahrt (DLR). Some of the data providers of the geoportals also provides exploitation platforms available. "An EO Exploitation Platform (EP) is a virtual, open and collaborative environment, which brings together EO and non-EO data, computing resources, tools to support data exploitation (processing, data mining, data analytics), algorithm development, collaboration and communication (e.g. social networks, fora), and market place functionalities" (ESA, 2019). Examples of

¹<https://inspire-geoportal.ec.europa.eu/>

²<https://inspire.ec.europa.eu/Themes/Data%20Specifications/2892>

³<https://earthdata.nasa.gov/>

⁴<https://earth.esa.int/eogateway/>

⁵<https://eoweb.dlr.de/egp/>

exploitation platforms are ESA thematic exploitation⁶, and CODE-DE⁷. ESA thematic exploitation platform provides various stakeholders for sharing their code/ Application Programming Interface (API). Similarly, CODE-DE provides Copernicus data about Germany. It is a German based geoinformation strategy where students and private users from any country can use as well as share data on this platform.

However, the growing numbers of portals for handling EO data has been given by Sudmanns et al.(2020). The Figure 2.1 below illustrates how on-demand computing which is same as online processing has been increasing over time. The sources mentioned are not only geoportals but also the platforms or solutions that deals with the big geo-datasets. The degree of interactivity of the portals/ solutions displays three possibilities by which user can interact with it, namely - online processing, uploading, and downloading the data. It can be concluded from the figure that recent portals and solutions have increased the interactivity of the users and hence enhance their experience of using it.

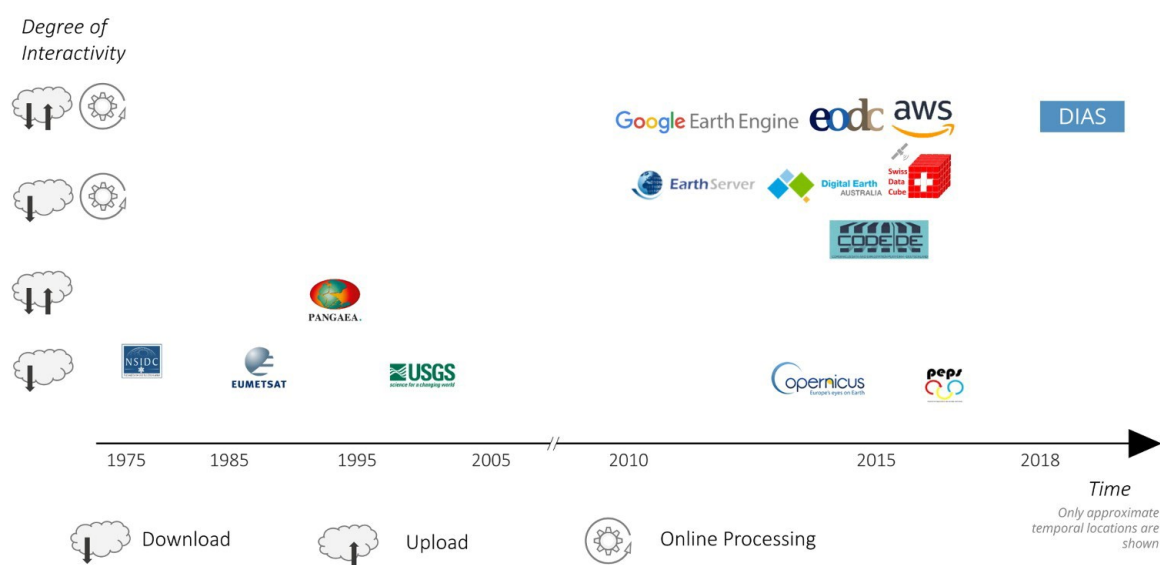


Figure 2.1: Timeline of technical solutions with respect to their degree of Interactivity (from Sudmanns et al., 2020).

Apart from geoportals there are also some online resources that provide remote sensing solution which can used to compute EO products. Google Earth Engine (GEE) is such an example. "GEE is a cloud-based platform for planetary-scale geospatial analysis which allows to process a variety of geographical data at scale and handle large geographical datasets" (World Bank, 2020). It has ready to use datasets which can be accessed via its web interface. It provides open and free access for educational purpose. Another example is IDB. IDB is a tool for computing the remote sensing indices (Henrich et al., 2012). It was developed by German researchers – Verene Henrich, Gunther Krauss, Christian Götze and Christofer Sandow and was launched in 2012. "It contains information on 519 indexes, 167 satellite sensors, 43 ways of index application, and 230 references" (Ukrainski, 2019). The purpose of this platform is to provide the exhaustive list of spectral indices present and it also allows to search indices according to sensors and applications.

⁶<https://eo4society.esa.int/thematic-exploitation-platforms-overview/>

⁷<https://code-de.org/en/>

The above-mentioned portals, solutions, or tools are just few examples of EO data exploitation. There are many such examples present today with techniques which improves their data availability. For this, their database structure requires proper storage and maintenance of data.

2.2 SEMANTIC DESCRIPTION

Semantic web provide rules to structure a database. The semantic web is network of interlinked information provided by a multitude of stakeholders. This web of data/ metadata is used by computers to perform meaningful tasks. “The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation.” (Berners-Lee et al., 2001). It enables people to store data on web, build vocabularies and add rules for handling data (W3C, 2015a). This can be done by following the semantic web stack containing languages and technologies Figure 2.2.

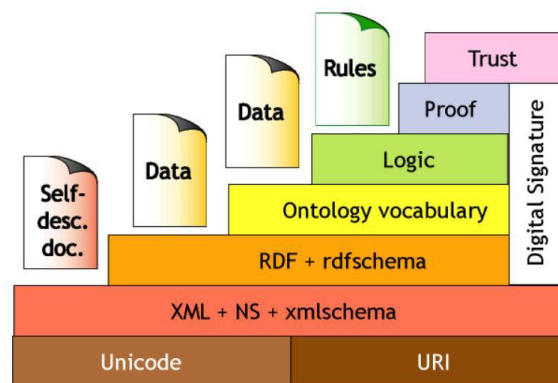


Figure 2.2: Semantic Web architecture. (from Berners-Lee, 1999)

The various parts of the semantic web architecture/stack are discussed briefly as following:

Syntax layer The syntax layer provides an interexchange of data throughout the semantic web architecture. The syntax can use different formats Extensible Markup Language (XML), or JavaScript Object Notation (JSON), etc. Namespace (NS) is the identify of a domain (knowledge base) on internet. Uniform Resource Identifier (URI) and Unicode are also essential in this layer. Unicode layer provides unique identity for every character to be used on any platform.

Resource Description Framework (RDF) / RDFschema (RDFs) The World Wide Web Consortium (W3C) defined RDF as – “A foundation for processing metadata; it provides interoperability between applications that exchange machine-understandable information on the Web” (Lassila et al., 1999).

RDF is a data model built on edge-node "graphs." Each link in a graph consists of three things (W3C, 2015a) as shown in (Figure 2.3):

- Subject (start node, instance, entity, feature);
- Predicate (verb, property, attribute, relationship, link); and
- Object (value, end node).

‘RDF Schema provides a data-modelling vocabulary for RDF data.’ (Brickley et al., 2014). In simple words, it gives a template for describing the abstract syntax of RDF, for example, Turtle (ttl), and JSON-LD (Brickley et al., 2014).

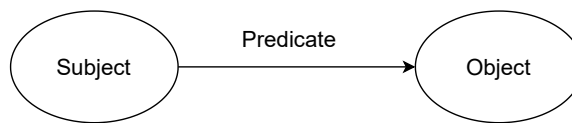


Figure 2.3: RDF Triple.

Ontology Layer Ontology layer is on top of RDF in the semantic web architecture. It consists of Ontology Web Language⁸ (OWL) based languages such as SPARQL which can be used to perform queries on any interface. OWL gives the first-order description logic. This allow to do basic classification like "all motorized vehicles with 4 wheels are cars" or "the mother of my mother is my grand-mother".

Logic Layer The logic layer provides a set of rules and actions to be performed on the data. OWL basically describes properties of entities and then assigns classes. The logic layer works more like if-then-rules.

Proof & Trust Layer The digital signature is used for verifying the source of standards and to protect the data. The proof & trust layer is still not really in use right now.

User Interface layer The last layer, User Interface layer is user centric and deals with presenting the information to the users.

Overall, the goal is to improve the usability of systems, their seamless & automatic integration, and the exchange of data across organizations, etc. The above forms the basis of any Semantic Web related process. The semantic web uses all these layers to generate a result which is understandable by both machines and people. An example of semantic web can be seen in Figure 2.4.

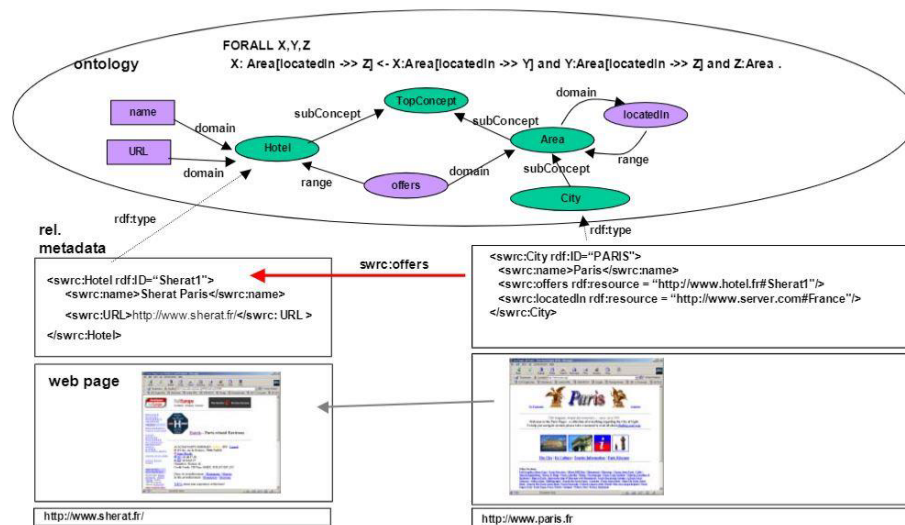


Figure 2.4: Example scenario for the Semantic Web. (from Maedche et al., 2002)

Maedche et al.(2002) gives an insight of a tourist portal where the user can search any specific hotel in Paris. Here, all the major layers like syntax, RDF/RDFs, ontology, and logic layers are displayed. It focuses mainly on the advantage of semantic for tourism. The authors introduced the concepts related to semantic web to implement their vision of understanding information

⁸<https://www.w3.org/TR/owl-guide/>

more on the system's side rather than the user's side. Out of many concepts, few were semantic search, or knowledge portals. Ontobroker (Angele et al., 1999) was one of the suggestions for semantic search which allows user to query the World Wide Web (WWW) (Maedche et al., 2002). Knowledge portals provide domain-specific information. Here, the knowledge portals enables the automated discovery of information. This paper paves the way for future in the sense where the loosely held information can be rich for querying using semantic web technology.

Although , Semantic Web is not recent technology but there are areas that still needs development or improvement like the trust and proof layer. This is one of the reason why the original structure is maintained because it still needs complete development.

The semantic rules include some fundamental components to structure it. These components have been discussed in the next section.

2.3 FUNDAMENTALS OF SEMANTIC STRUCTURE

The semantic structure makes use of knowledge graph. "The knowledge graph represents a collection of interlinked descriptions of entities – objects, events, or concepts. Knowledge graphs put data in context via linking and semantic metadata and this way provide a framework for data integration, unification, analytics and sharing" (Ontotext, 2016). They can be explored by imposing queries and also have semantic descriptions. The semantic description means to define terms and their connections which can be done by formal semantics. Formal semantic is used to represent abstract knowledge of the graph and has various components like:

1. Node/ Entity/ Concept/ Class
2. Blank Node
3. Relation/ properties
4. Literals
5. Instance/ Individual

Node is also known as entity, concept, or a class. These represent real world entity like vehicle, person, animal, etc. They contain unique Identity Document (ID) known as Uniform Resource Locator (URL) which is used to uniquely identify the various classes in the knowledge graph. Some of the nodes are a collection of different objects named class. The others are called blank nodes. Blank nodes are entities without global identifier and used for a better clarity of the classes. Relations are the connections between the classes. The type of relation defines how a class is related to another class. It can be data property, object property, or annotation property. The data property are properties that define relation between a class and its data (literal value). Object properties are the connection between two classes. Annotation properties are used as assertions for a specific concept.

Literals contain the value of a class and the type of value used which can be a string, numerical, or a Boolean value. These are predefined literals but it can also be defined manually. Instances or individuals are the entities which represents a particular example of a class. For clarity, (see Figure 2.5) lets take an example of a statement – "Amsterdam is the capital of Netherlands". It can also be inferred as "A city is the capital of a country". So, the 'city' and 'country' are the collection of real-world entities and hence are classes. Amsterdam and Netherlands become the instance for city and country, respectively. The data property of Netherlands is 'hasPopulation' and the literal value (xsd:integer) for it is 17,134,872. The object property between Amsterdam and Netherlands is 'isCapital'. Now getting everything together in an RDF schema (as already discussed before) which uses the triples – subject, predicate, and object. This example would look like Figure 2.5.

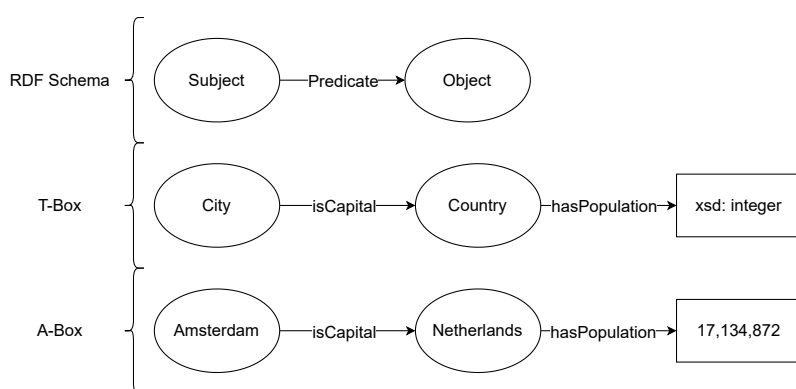


Figure 2.5: An example for RDF Schema.

Knowledge graphs can be separated into T-Box and A-Box. The T-Box known as Terminology Box is used to define classes, their properties and the literal types. The A-Box known as Assertion Box is used for adding instances. In this example, 'City' and 'Country' are the classes for the T-Box. 'xsd:integer' is literal while 'isCapital' and 'hasPopulation' are the properties for the T-box. 'Amsterdam', 'Netherlands' and '17,134,872' are the instances in the A-box. The T-box is a generalized concept of a semantic model and it is not necessary that every component gets reflected in the A box. The T-box and A-box are known as descriptive logic and are used to describe concepts using complex architecture. This gives meaning to the model. In RDF schema, there are different terminologies to describe nodes (such as `rdfs:Class`, or `rdfs:Literal`) and their relations (such as `rdfs:domain`, or `rdfs:range`). More detailed list has been displayed by W3C⁹.

These terms are generally used for materializing the model into an RDF format. RDF forms a base vocabulary which is extensible like RDF/XML. RDF uses semantic languages such as RDF/XML, Turtle, or JSON-LD. Turtle has gained popularity among these. There are many examples of RDF language serialization on Linked Data Vocabularies¹⁰. Apart from these, the RDF can be easily developed for any application as well as individual interests. One such example can be from NASA (Earth Data)¹¹ that provides vocabulary description of raw EO data (such as sensors, resolution data, etc.) in RDF format along with Comma Separated Value (CSV), JSON, and XML. Similarly like RDF, OWL also has pre-defined classes and relations which are not present in RDF, yet it uses some of the classes and properties from RDF.

A query language is used used to extract knowledge from the RDF such as SPARQL. "The query language, SPARQL 1.1¹² is a set of specifications that provide languages and protocols to query and manipulate RDF graph content on the Web or in an RDF store" (W3C, 2013). SPARQL make use of URI to identify classes or concepts and impose queries for data graph as RDF.

To develop a knowledge graph for the semantic model a contextual background is needed from building the components of semantic model to making it work for that purpose. This has been discussed in the next section.

⁹<https://www.w3.org/TR/rdf-schema/>

¹⁰<https://lov.linkeddata.es/dataset/lov>

¹¹<https://gcmd.earthdata.nasa.gov/static/kms/>

¹²<https://www.w3.org/TR/sparql11-query/>

2.4 SEMANTIC MODELLING

As already discussed, the fundamentals of semantic structure can be used to build a semantic model of any kind. Semantic modelling is the representation of various classes and their relation together towards a particular goal. Semantic modelling approaches make use of ontology models which explicitly represent any concept. The term ontology is borrowed from philosophy meaning "systematic account of existence" (Gruber, 1995). Thomas R. Gruber wrote two famous papers defining the term ontology - *Toward Principles for the Design of Ontologies Used for Knowledge Sharing* and *A Translation Approach to Portable Ontology Specifications*. He defined ontology as "an explicit specification of a conceptualization" (Gruber, 1995). Gruber's shared vocabulary during the web development phase in 1990's was adopted later by Tim Berners-Lee as Semantic Web or Web 3.0 with a vision of accumulating data on a single platform. Today, ontologies have become famous in the context of Semantic web as it provides an appropriate tool for sharing information, and naming convention for a particular application domain (Hüsemann et al., 2005).

With increasing data and their metadata, there is a need to handle these data properly which can be done with the help of ontology. It can handle data from different sources as already discussed. Building an ontology requires creativity and generally, it is done as shown in Figure 2.6.

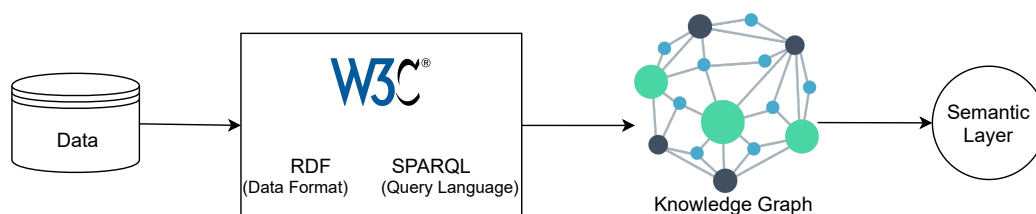


Figure 2.6: A generalized example of steps involved in building a semantic layer.

The data is structured with regard to data format, mostly RDF format (not always) by W3C. The other format might include Ontology Web Language(OWL)¹³. It is then converted into a knowledge graph which completes the ontology and hence, forms a semantic layer. The knowledge graph represents a collection of interlinked descriptions of entities – real-world objects and events, or abstract concepts which is understood both by humans and computers (that is, it is machine readable).

The above was just a generalized example. There are various methods an ontology can be built. Hüsemann et al.(2005) introduces a method for building an ontology as shown in Figure 2.7. Here, after every step there is evaluation to make sure that at any step the ontology model should not fail. This paper make use of multimedia domain in which a person wants to manage his/her multimedia data such as music files, image files, movie files, and ordinary files with metadata from their collection. The steps involved are:

Requirement Analysis identifies relevant requirements through different specifications. This can be organization analysis, process analysis, or data analysis. The organization analysis targets the potential users of the ontology. In this case, this can be a multimedia creator or a user. The process analysis identifies where the ontology will be used and for what purpose. Data analysis identifies various data sources that will be used to built the model.

¹³<https://www.w3.org/TR/owl-features/>

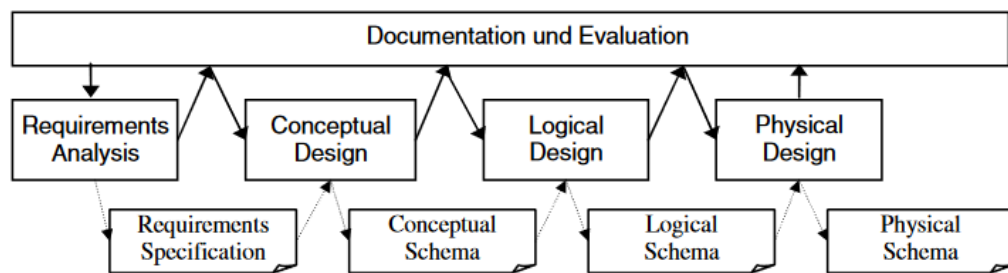


Figure 2.7: Steps of the ontology design methodology. (from Hüseman et al., 2005)

Requirement Specification makes use of the requirement analysis. It collects all the relevant concepts within the identified context of the ontology. This includes specifying a use case, formulating competency questions (CQs), specifying an initial glossary, refining the initial glossary, and evaluation. Use cases are used to specify a particular scenario. CQ are used to evaluate the completeness of the model at the end. CQs formulate various questions and answers them at the end of the design phase. An initial glossary is identifying different concepts used and their definition. Glossary refinement specifies whether a concept of the initial glossary is a literal or an instance of another concept. Last is the validation and verification of the model.

Conceptual Design The concepts identified in the Requirement Specification are given a naming convention. The properties connecting to different concepts are also given proper names. The concepts and their properties are differentiated from each other.

Conceptual Schema The conceptual design is given additional information like the range and the domain of the concepts and properties are identified. This completes the overall structure of the model.

Logical Design and Logical Schema This step converts the conceptual schema into an ontology language that can be RDF/XML, Turtle or JSON-LD.

Physical Design and Physical Schema This phase uses logical design for the particular application mentioned before.

The method of building an ontology here is systematic which captures every little details for building a semantic model. It not a domain specific development of ontology but using this method, any kind of ontology can be engineered. Even the evaluation method which uses CQs is feasible for verification of the semantic model.

Another ontology for a particular application is proposed by Wang et al.(2017). The author aimed to build an ontology for hydrologic monitoring. Here, the author defined the hydrologic monitoring as monitoring the movement of water in a hydrological system. This means the mobility of water from rainfall to surface runoff. This paper focused on the providing a proper standard for the semantics used for hydrologic monitoring. The ontology is proposed whose three concepts are identified, namely - Sensors, Observations, and Events. Sensors are basically the observers in hydrological observing which are further divided into several components. These components are not very briefly explained but it tells that the Sensor concept is used for defining these components. The observation concept is used to define the observed property by the sensor in a hydrological observations. These can be water temperature, rainfall, water quality, etc. Events

concepts is used to define the hydrologic events like floods, tsunami, debris flow, etc. for the hydrological monitoring. The concepts are related to each other as shown in Figure 2.8. The author gives a short description on the implementation of the ontology. The properties connecting the concepts are not well defined. The instances are not included and only the T-box is explained. The paper is focused on interoperability of the model but it lacks the proper reason to choose the concepts and properties. Moreover, there is no verification of this ontology.

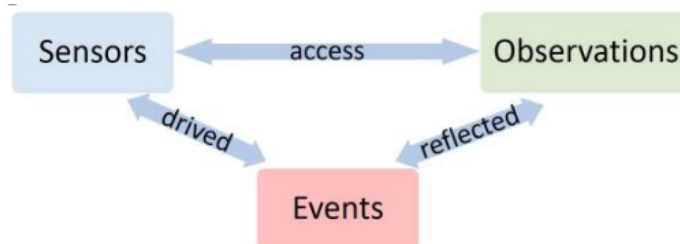


Figure 2.8: Ontology design for hydrologic monitoring. (from Wang et al., 2017)

2.5 RELATED WORK

All the above concepts can be used for semantic modelling where indices to be computed can be retrieved as per the user's requirement. Many research studies are dedicated to semantic web enabled geoprocessing of EO products.

2.5.1 Ontology Design

Wang et al.(2016), for example, proposed an approach for monitoring soil moisture using the satellite sensors. The authors developed an ontology which described the EO metadata stored in a knowledge base with a spatiotemporal-spectral-enhanced structure. This ontology was developed as a vision to provide proper discovery of satellite data. Here, the EO metadata ontology results were displayed in a web application prototype. The ontology development was divided into two parts – one where a model was proposed based on observation metadata and the second one, where OWL was used to describe the model. Various soil related indices were computed and later checked for its ability to be discovered. The result from this model was evaluated based on precision, and using statistics F-measure. The precision was evaluated by comparing their model with already built semantic models like the NASA's Semantic Webfor Earth and Environmental Terminology (SWEET) ontology (DiGiuseppe et al., 2014), the W3C-developed Semantic Sensor Network (SSN) ontology (Compton et al., 2012), and the Sharing Environmental Education Knowledge Extensible Observation Ontology (Madin et al., 2007). This gives the precision of sensors and observations used. The F-measure gave the results on the the time required to get the desired results. The method of evaluation of the semantic model using F-measure can be incorporated but the evaluation is done for a spectral index which were actually computed. For precision evaluation, the existing ontologies mentioned in the paper do not describe the products from the EO satellites but the sensors and their observation. Therefore, this cannot be useful method of evaluation for this research's semantic model.

This work can be extended to different ways. First, this research is only limited to a specific use case which is agricultural soil monitoring, but can also be extended to various application fields like forest management, water resource management, etc. Second, the products produced here are displayed in the application prototype with no mention of downloading it but they cannot be downloaded, therefore, an interface that can enable direct downloading or on-demand downloading can be their future work.

2.5.2 Application based semantic model

Mandl et al.(2007) focuses on service-oriented architecture for on-demand geoprocessing of science products. This work is handled by NASA, Goddard Space Flight Center (GSFC). It gives a prototype for retrieving EO products related to disaster management. It is built on OGC Sensor Web Enablement (SWE) standards. In one of the scenarios for disaster management, an image of fire map is created. This workflow creates the JPEG fire map from the EO-1 Hyperion data (Mandl et al., 2007). As this paper is not only limited to EO products, but it also gives a prototype of retrieving Unmanned Aircraft System (UAS) products with an example of same fire map. The authors also proposed a general workflow of any disaster related themes. The user centric theme included-Fire, Flood, Spills, Volcanoes, Contaminants, Tsunamis, and Earthquakes. The user can select the required theme, its location and priority (like before event or ongoing event or after event). All these requirements are sent to BPMN which is used for the user interface. This then sends the signal to select from various workflows like Business Process Modeling Notation (BPEL), SensorML (Model Language) or openWFE (not actively in use) which executes the processes involved for the generation of products over the web. Here, the user can get products online or on their desktops directly. Although, the prototype is a good representation, it lacks to indicate the risks or limitations that might occur while developing the real workflow. Also, the research is just limited to a particular use case. The satellites involved are few which are only done for Global Earth Observing System of Systems (GEOSS). For the data implementation, it used integration of Web 2.0 and OGC SWE. A better approach would have been to use semantic web for the interoperability of the data for both user and system perspective.

2.5.3 OGC Standard for EO Products

The OGC standards are focused on technical audience which outlines the structure for specific software components. There is a list of OGC Standard implementation with OGC document code. The one related to EO products is serialized as OGC: 17-003r2. This document provides information in RDF format which includes classes and properties of the EO vocabulary. It is encoded as JSON-LD and is used as with a namespace as 'eop'. The structure of the document is given in Figure 2.9. The classes are colour coded as green with dependencies as orange. This document gives a specific standard for explaining raw EO products and their dependencies.

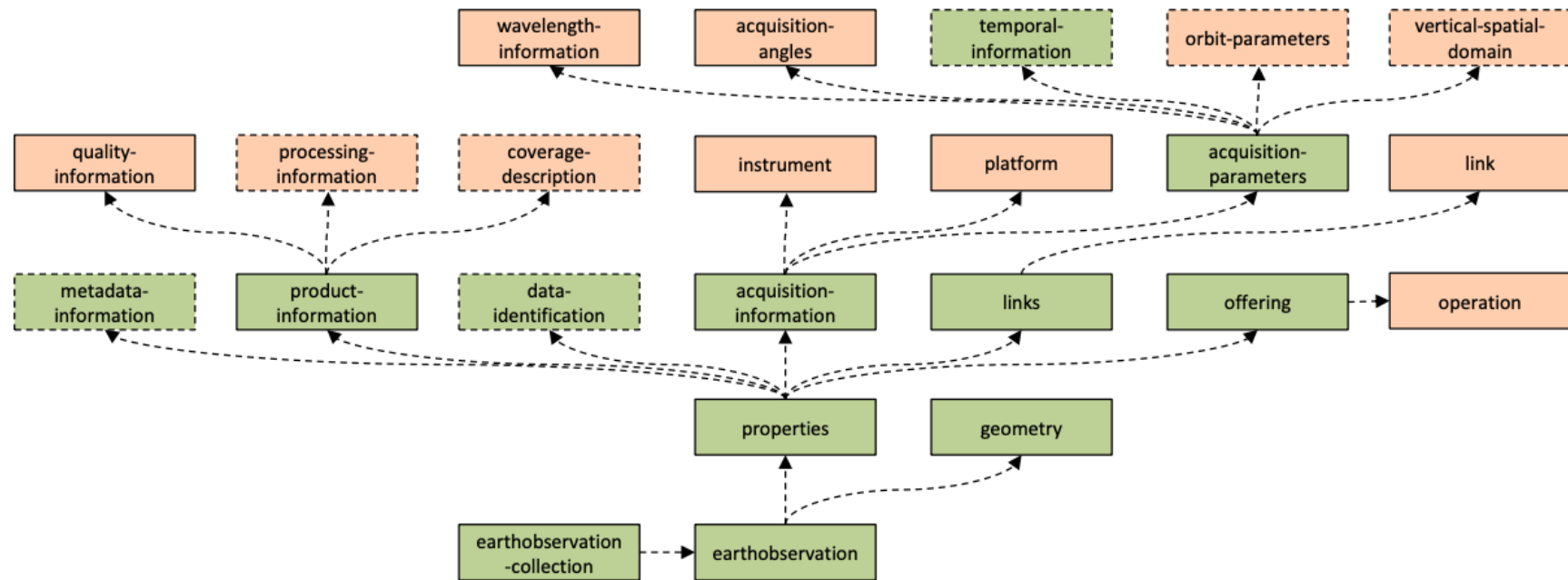


Figure 2.9: Conformance and requirements class dependencies. (from Coene et al., 2020)

Chapter 3

Implementation of Semantic Model

Using the above literature, a semantic model for on-demand EO products will be developed here. As the focus of this model is towards the spectral indices amongst the EO products so the criteria for its development is also specific. This criteria has been explained in detail in this chapter. The following steps as in Figure 3.1 are followed for developing the semantic model with its validation and verification.



Figure 3.1: Workflow

First, the analysis of the various components required for the spectral indices is done. This includes exploring different spectral indices on several platforms, and studying their requirements through the formula specified and/or their required sensors. The thematic domain for spectral indices are also discussed here. The use case analysis gives the background for possible applications of spectral indices and also identifies the different components required in the model. Second, the structure for T-Box (as mentioned in Section 2.3) of the semantic model which connects the components together to give an explicit representation of the domain. In a third step, this model is converted into an RDF-format of the model which is used to add instances. In fourth step, for illustration purpose selected examples describing spectral indices and their calculation are created as the A-box (as mentioned in Section 2.3) of the model. Alongside examples for raw EO products, those will serve as the basis for an evaluation later on (see Chapter 4). Finally, the inputs needed for validation and verification were briefly introduced here while it is described in detail in Chapter 4.

3.1 REQUIREMENT ANALYSIS

Semantic models are built for a particular purpose and intention. The overall goal here is the on-demand processing for EO data. Therefore, the requirement analysis is done to bring all the pieces of information together for the model. The objective here is to build a semantic model that can enable on-demand processing of spectral indices. The spectral indices are dependent on both the components required to calculate it and the raw EO data. These components can be input signals of a particular spectral wavelength depending on a particular sensor or anything the provider of the index specified. It should be noted that the semantic model for raw EO data has been already accomplished by OGC: 17-003r2 (Coene et al., 2020) (See Section 2.5.3). Therefore, the focus is more towards developing the components required for spectral indices. However, the model uses OGC: 17-003r2 later to populate it with the raw EO data as the model is dependent on this as well.

To make the conclusions concrete, several sources of information on spectral indices were studied. There are various existing platforms that provide information about spectral indices like Land Surface Reflectance¹. It provides Landsat derived spectral indices and their metadata but currently contains only few spectral indices². It also provides the metadata related to a spectral index. There is another online platform providing information about the indices by Harris Geospatial Solutions. They provide a range of approximately 70 indices with a background description of processing it on Environment for Visualizing Images (ENVI) software. Aforementioned IDB (Henrich et al., 2012) claims to provide the most exhaustive list of spectral indices on a single platform. It provides a list of spectral indices and maps those to specific sensors and applications. But these are not only options for exploring spectral indices. There are other sources too like scientific papers, like Bannari et al.(1995), which provide much detailed information about an index such as their applications, and required wavelengths or constants.

3.1.1 Spectral Indices analysis

Spectral indices are derived by a mathematical formula that is applied to a collection of spectral bands. The formula of an index can be used to infer information about the dependency of an index for a particular spectral range as well as other parameters such as constants in a formula. The spectral range is a specific part of the electromagnetic spectrum with a starting and ending wavelength. Some ranges are assigned a particular label such as Near Infra-Red (NIR), is a spectral range of starting and ending wavelength of approximately 700nm to 1300nm, respectively as stated on L3Harris Geospatial³ website. Some indices require additional constants. These constants can be just a numeric value or be defined as a nested formula themselves. This formula can have another set of dependencies. However, these values can also be dependent on a platform or a sensor. Platforms and sensors differ from each other. A platform forms a kind of base on which a sensor is mounted. For example, ISS can be a platform from where astronauts take images using their cameras. Landsat is a platform which carries different sensors like Landsat 1-3 Multi-spectral Scanner (MSS), Landsat 4-5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper Plus (ETM+), and Landsat 8 Operational Land Imager (OLI)/ Thermal Infrared Sensor (TIRS). For overall analysis of the spectral indices different aspects were considered. Table 3.1 provides the details for four example indices: This table shows a list of indices which are : NDVI, Transformed Soil Adjusted Vegetation Index (TSAVI), Modified Soil Adjusted Vegetation Index (MSAVI), and Chlorophyll Absorption Ratio Index (CARI). These indices were chosen because they demonstrate all kinds of dependencies for describing the components of the semantic model that were discovered throughout the literature research.

3.1.2 Use Case

The semantic model forms the basis of a knowledge graph. The knowledge graphs are a collection of structured vocabulary which is readable by machines as well as humans. These kind of knowledge graphs are basically referred to as ontologies. Ontologies are than just a vocabulary as they are used to define more complex structure for a particular area (W3C, 2015b). For example, the

¹<https://www.usgs.gov/core-science-systems/nli/landsat/landsat-surface-reflectance-derived-spectral-indices?>

²At the time of writing the thesis (June 2021), only seven indices are provided.

³<https://www.l3harrisgeospatial.com/Learn/Whitepapers/Whitepaper-Detail/ArtMID/17811/ArticleID/16162/Vegetation-Analysis-Using-Vegetation-Indices-in-ENVI>

Table 3.1: Examples of spectral indices descriptions obtained through various sources.
The displayed platform and sensor for each index is just an example and does not mean that it is the only option.

Index Name	Proposed By	Formula	Platform: Sensor	Spectral Bands & their Range (µm)	Application Area
NDVI	Rouse et al., 1974	$\frac{NIR-RED}{NIR+RED}$	Landsat: MSS (Multispectral Scanner)	Band 5 (RED) → 0.6 – 0.7 Band 6 (NIR) → 0.7 – 0.8	Vegetation
TSAVI	Baret et al., 1989	$\frac{b(NIR-b.RED-a)}{RED+b(NIR-a)+X(1+b^2)}$ $a = Slope(a : 0.8863 \text{ for AVIRIS})$ $b = Intercept(b : 55.0 \text{ for AVIRIS})$ $X = 0.08$	Airborn: AVIRIS (Airborne visible/ infrared imaging spectrometer)	Band 37 (RED) → 0.693 – 0.702 Band 78 (NIR) → 1.082 – 1.087	Vegetation & Soil
MSAVI	Qi et al., 1994	$\frac{(NIR-RED)(1+L)}{NIR+RED+L}$ $L = 1 - (s * NDVI * WDVI)$ $s = Slope$ <i>WDVI: Weighted Difference Vegetation Index</i> <i>NDVI: Normalized Difference Vegetation Index</i>	Landsat: MSS (Multispectral Scanner)	Band 5 (RED) → 0.6 – 0.7 Band 6 (NIR) → 0.7 – 0.8	Vegetation & Soil
CARI	Kim et al., 1994	$\frac{REDEDGE}{RED} \cdot \frac{\sqrt{(a.RED+RED+b)^2}}{(a^2+1)(0.5)}$ $a = \frac{REDEDGE5-GREEN}{150}$ $b = GREEN - \left(\frac{REDEDGE5-GREEN}{150,550} \right)$	Sentinel 2: MSI (Multispectral Instrument)	Band 5 (REDEDGE) → 0.705 Band 4 (RED) → 0.665 Band 3 (GREEN) → 0.560	Vegetation

definition of ontology can also be modified based on its purpose. Like it can be used for data integration, or data storage. Data integration helps users to combine or share knowledge from different sources together as displayed in LOD⁴ cloud. While data storage is used to organize and maintain data in a structured way. Here, both the definitions are well projected.

The ontologies are built with a specific purpose in mind that define a certain set of use cases they have to support. Hence, the use case analysis is done to understand the context and extract the knowledge necessary to build the fundamentals. For the use case let's take an example from agricultural background, a domain with a large user base and a substantial demand for EO products (Migdall et al., 2018). Companies and organizations keep building GIS tools that can help increase the efficiency of agriculture production, account for environmental factors, and reduce manual labour.

In our example, a GIS company got a project which deals with several aspects of crop monitoring. Crop monitoring in context of remote sensing includes the growth and the health of the crop. The company makes use of satellite data which involves processed EO data as spectral indices. The spectral indices can map changes in crop from seed plantation, seed sowing, crop growth to crop harvesting. Crop harvesting is collecting crops when they have been matured. Every crop has its own harvesting time period. Assume that an analyst in the company is given the task to monitor wheat crop for the harvesting month of May 2021 in Assen, Netherlands. For this, he uses a monthly temporal resolution that can extract the data for wheat. These indices should indicate whether a crop is ready for harvesting or not as the spectral reflectance curve for crop can indicate it (Arafat et al., 2013). The GIS analyst as a user communicates with the system to identify available indices related to crop monitoring. The system will find the list of indices according to the sensors present for a given wavelength. Depending on these sensors, the raw data will be available within a spatial and temporal resolution which gives the list of required spectral indices. The use case is illustrated in Figure 3.2.

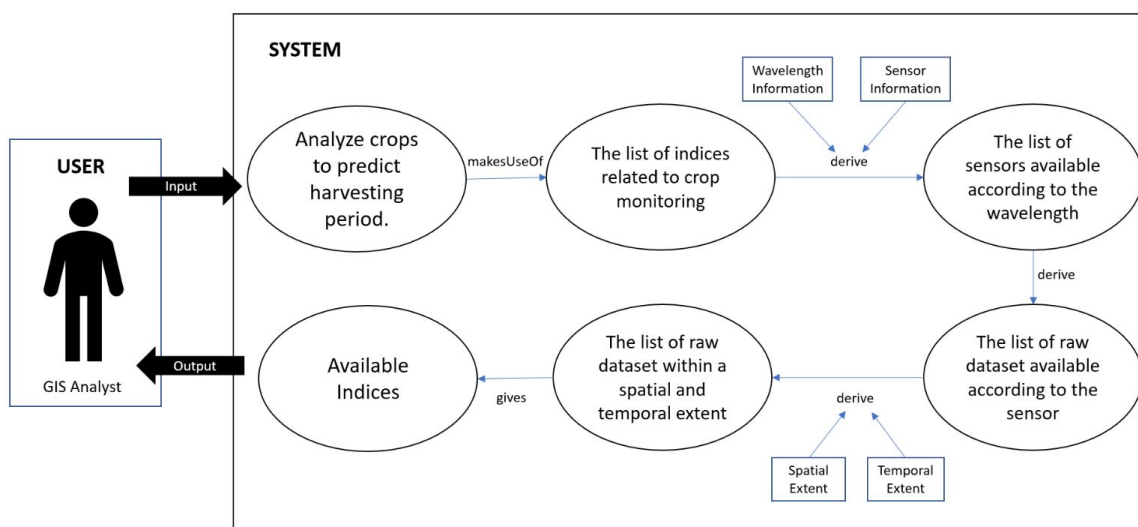


Figure 3.2: Use case example: analyzing crop harvesting periods.

⁴<https://lod-cloud.net/>

3.1.3 Competency Question Analysis

The use case outlines the process for generating a list of suitable spectral indices by the GIS analyst. It forms the basis for deriving requirements to the ontology. However, to define the relevant concepts for the ontology, competency questions are identified. The competency questions represent the different information needs throughout the use case. They can also be used to validate the completeness of the ontology at the end (Hüsemann et al., 2005). The competency questions for the above use case are as follows:

- CQ1 What tasks are present for agriculture?
- CQ2 Which spectral indices are available for crop monitoring?
- CQ3 Given that the user wants to know about the origin of the spectral index, what is the reference of that spectral index?
- CQ4 Which of the datasets are present for May 2021 assuming it the harvesting month of the crop?
- CQ5 Given that NDVI is chosen for crop monitoring, what constraints are used on the input data to compute it?
- CQ6 Given the constraints require another index like in MSAVI, is such an index available?
- CQ7 Given the constraints describe wavelengths best suited for an index like for TSAVI, which sensors and platforms provide compatible data?
- CQ8 Given the compatible platform like for Sentinel 2B, which raw datasets were created by that platform?
- CQ9 Given input datasets satisfying the constraints, are those available for the given spatial extent?

3.1.4 Concept Identification

As mentioned, the competency questions define the concepts related to an ontology. Therefore, the concepts and their connections were brainstormed based on the CQs. The result of this identification is shown in Table 3.2. For example, CQ9 states that the spatial extent of the dataset is required which was thought processed into a Geometry concept in Table 3.2. Similarly, in Table 3.2, the competency questions are connected to several concepts. This gives an purpose for using a concept. These concepts are explained in detail in the next section.

3.2 BUILDING THE T-BOX OF SEMANTIC MODEL

The main concepts identified above are used as classes/concepts for the semantic model, forming the T-box. The concepts needs to be connected to each other to give a meaningful representation of spectral indices calculation which is done through adding relations. The concepts with their relation form a conceptual schema (see Section 2.4) for the semantic model. This schema gives the descriptive structure to the semantic model. The different concepts connected to relations are described in detailed here. Furthermore, it gives the final overview of the semantic model (Figure 3.3). The T-BOX for this semantic model has been displayed in Appendix B. In the following, the formal definitions for each concept has been provided, before describing their relationships.

Table 3.2: Concepts identified for the model depending on the competency questions.

CQs	Concepts
CQ1	Theme; Task
CQ2	Index
CQ3	Reference
CQ4	Temporal Information
CQ5	Requirements
CQ6	Required Index
CQ7	Platform; Instrument; Wavelength Information
CQ8	Instrument; Wavelength Information; EO Product
CQ9	Geometry

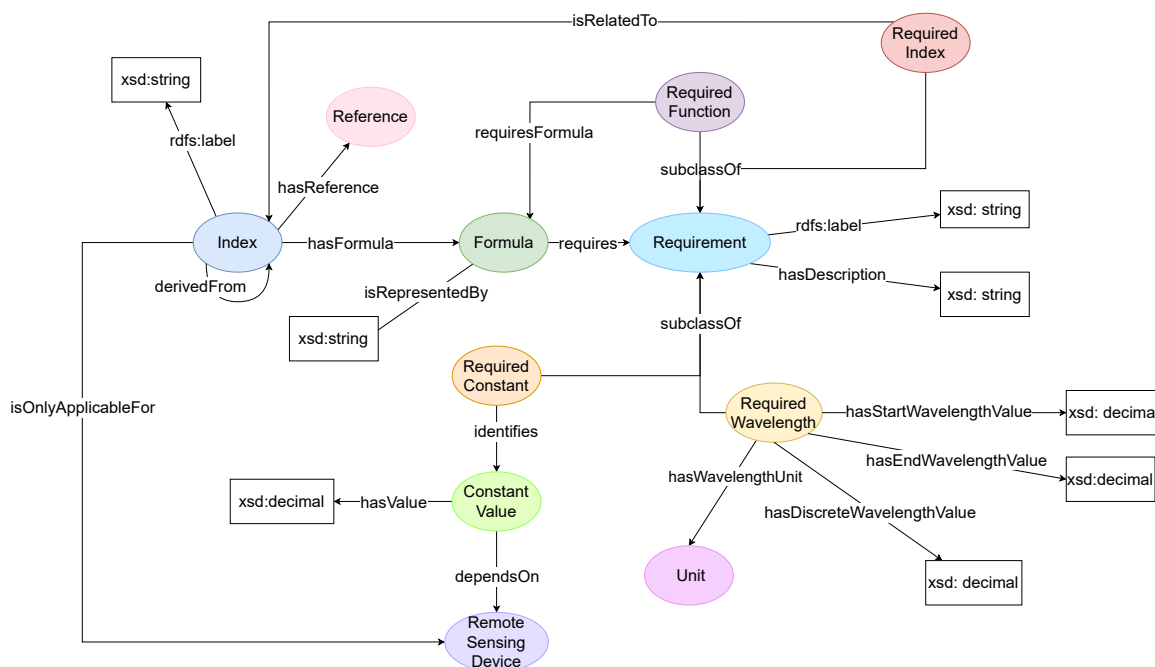


Figure 3.3: T-BOX to describe the components of spectral Indices.

Definition 3.2.1 (Index). The index refers to spectral indices. The spectral index is a mathematical formula which is a combination of one or more spectral bands. These can also include the constants as well.

The indices are the names of the spectral indices which includes the instruction to generate datasets for the given index. It can also be referred as indices, or spectral indices. The index should contain the description, formula, and task connected to it to answer questions of the CQs. The index also has a reference.

Definition 3.2.2 (Reference). The reference refers to the paper which displays the formula for a particular spectral index.

The reference of paper is provided for every index which gives the information about the origin of an index with its formula. It further explains the usability of an index.

The index consists of tasks.

Definition 3.2.3 (Task). Task refers to the function that a particular spectral index can perform.

Tasks are the more detailed information of a particular theme. The tasks includes the themes related to it. The T-Box for tasks and themes are displayed in Figure 3.4.

Definition 3.2.4 (Theme). The theme describe the possible environments in which the spectral indices could be used.

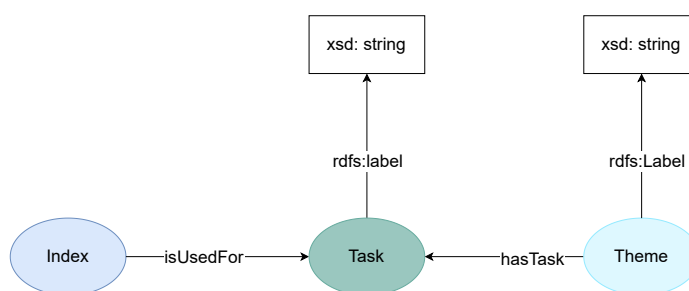


Figure 3.4: T-Box to describe the thematic domains of the spectral indices.

The following concepts describe the dependency of an index:

Definition 3.2.5 (Formula). Formula of an spectral index is a mathematical equation which gives the dependencies for it. It consists of spectral bands, or constants, or another index etc.

The formula of an index states the dependencies required for it such as spectral band, constants, and various other parameters as discussed below. These parameters are connected to a superclass concept Requirement.

Definition 3.2.6 (Requirement). The requirement concept is a superclass for all the concepts required by the formula.

These concepts are given below.

Definition 3.2.7 (Required Index). Required index is a not same as Index concept as this describes the dependency of an index on another index.

As discussed before, sometimes the formula of an index requires another formula. This is also displayed in Table 3.1.

Definition 3.2.8 (Required Function). The required function concept is used to define one of the dependencies of the Formula concept. A formula might require a constant that is further dependent on another formula. which is described by Require Function of the semantic model.

Required function is colour coded as purple as shown in . An example of this is shown for MSAVI in Table 3.1.

Definition 3.2.9 (Required Constant). A formula might require a constant. A constant can be a fixed value or a value dependent on a sensor or a platform.

The required constant concept is colour coded as orange in the semantic model as shown in Figure 3.3. Sometimes, they are the part of a formula but not a necessity as shown in Table 3.1.

Definition 3.2.10 (Required Wavelength). The required wavelength gives the information about the required spectral bands used to describe the computation of an index.

The required wavelength specify start and end wavelength of the spectral bands. The model has the capability of also defining discrete wavelength of the spectral bands.

The ontology describing the raw EO data is done by OGC:17-003r2 (Coene et al., 2020). The previously identified concepts in Table 3.2 shows that the model depends on the raw EO data as well. Therefore, the T-box derived from OGC:17-003r2 ontology was used for the model as shown in Figure 3.5. These include concepts like geometry, temporal information, platform, instrument, Earth observation.

The following definitions were taken from OGC:17-003r2 Coene et al.(2020) descriptions:

Definition 3.2.11 (Earth Observation). "This vocabulary defines the classes and properties required to model Earth Observation (EO) metadata for datasets (i.e. products or granules). It provides the models for the exchange of information describing EO datasets, both within and between different organisations" (Coene et al., 2020).

Definition 3.2.12 (Geometry). Geometry represents the spatial extent.

The concept is displayed in OGC and is connected to EarthObservation concept with a literal specifying the geographical extent.

Definition 3.2.13 (Temporal Information). "Contains the properties related to the start and end time of the acquisition of the data" (Coene et al., 2020).

The temporal information concept of OGC is connected to Acquisition Parameter concept with start and end temporal resolution.

Definition 3.2.14 (Platform). "Platform represents the various satellites/ missions. The Platform used for the acquisition. If more than one platform is used for creating the product, then the AcquisitionInformation object occurs more than once" (Coene et al., 2020).

The platform of the OGC is connected to Acquisition Information concept of OGC and to Remote Sensing Device concept of the semantic model.

Definition 3.2.15 (Instrument). "Instrument are sensors that are mounted on the platforms. The Instrument/Sensor used for the acquisition" (Coene et al., 2020). If more than one instrument is used for creating the product, then the AcquisitionInformation object occurs more than once.

The instrument of the OGC is connected to Acquisition Information concept of OGC and to Remote Sensing Device concept of the semantic model.

Definition 3.2.16 (Remote Sensing Device). The remote sensing device is a concept of semantic model with platform and instrument as subclass from OGC:17-003r2.

The platform and sensors are different from each other as already discussed. Some requirements of the spectral indices are dependent on a sensor and sometimes on a platform. Therefore, this concept was introduced as a superclass that can represent both of them.

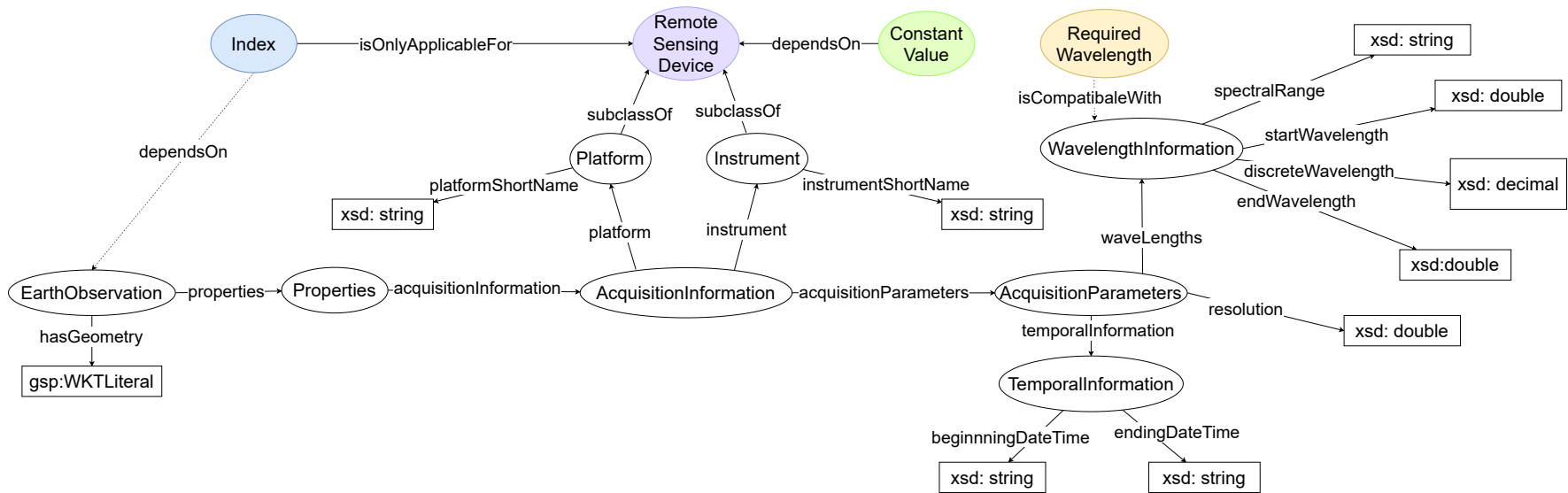


Figure 3.5: T-Box for raw EO data. Derived from OGC: 17-003r2 (Coene et al., 2020).

All the above concepts are connected to each other by properties. As mentioned before, there are different types of properties like Object properties and Data properties with their range and domain. The object properties are connection between concepts as shown in Table 3.3.

Table 3.3: Object Properties of semantic model

Object Property	Domain	Range
dependsOn	Constant Value	Remote Sensing Device
derivedFrom	Index	Index
hasTask	Theme	Task
isUsedFor	Index	Task
requiresFormula	Formula	Requirement
isRelatedTo	Required Index	Index
OnlyApplicableFor	Index	Remote Sensing Device
hasFormula	Required Function	Formula
hasWavelengthUnit	Required Wavelength	Unit

The semantic model also contains the data properties. The data properties are connection between concepts and literals. They are listed in Table 3.4:

Table 3.4: Data Properties of semantic model.

Data Property	Domain	Range
hasDescription	Requirement	xsd:string
hasDiscreteWavelengthValue	Required Wavelength	xsd:decimal
hasStartWavelengthValue	Required Wavelength	xsd:decimal
hasEndWavelengthValue	Required Wavelength	xsd:decimal
hasValue	Constant Value	xsd:decimal
isRepresentedBy	Formula	xsd:string

Apart from these properties, the annotation properties are used as well. "Annotation properties are used to place annotations on individuals, class names, property names, and ontology names" (Patel-Schneider et al., 2004). The annotation properties used are given below.

All the above concepts and properties are put together in a structure as shown in Figure 3.3. All the concepts defined by this thesis are coloured while the concepts reused from existing ontologies are uncoloured. The semantic model consists of three layers. The first layer can be seen in Figure 3.3 which describes the spectral indices and while doing so includes the dependencies. The next layer is of existing OGC: 17003r2 ontology (Figure 3.5) which involves raw EO data required by the ontology of spectral indices. The last layer displays the thematic domain of the semantic model (Figure 3.4).

Table 3.5: Annotation Properties of semantic model

Annotation Property	Description
eoi:hasReference	It includes the reference to papers describing an index. It is either a DOI (Digital Object Identifier) or full detail of a paper.
rdfs:label	This is a predefined vocabulary of RDFs as already discussed. Most of the concepts have label as shown in Figure 3.3.
dc:description	This is an annotation property from DC (Dublin Core) metadata specification. This contains all the elements describing the metadata.

3.3 IMPLEMENTING THE MODEL AS A LINKED DATA VOCABULARY

The concepts/ classes and properties mentioned above are used to build a linked data vocabulary for the calculation of spectral indices. It can be materialized into an RDF format. The software used here was Protege⁵ as it has an open license. It provides a clear interface for adding concepts and different properties. The namespace for the semantic model is the URL as ‘<http://w3id.org/eoi/>’ with the prefix ‘eoi’. The chosen namespace does not resolve yet, but can do so in the future.

The schematic diagram for spectral indices in Figure 3.3 is given more detailing while generating its RDF file. Also, possible ways for structuring the semantic model are discussed. In Figure 3.3 there is a simple representation of the formula but it can also be represented as Figure 3.6. For example, Equation 3.1 describes the calculation of NDVI. It can be represented in RDF as displayed in Figure 3.6.

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (3.1)$$

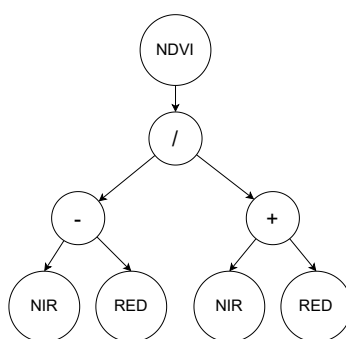


Figure 3.6: Mathematical representation of NDVI in RDF form.

⁵<https://protege.stanford.edu/>

The mathematical expressions for a linked data can be written with OpenMath (Wenzel et al., 2012). OpenMath is based on MathML and can be encoded as XML expressions for mathematical equations. It has its own RDF vocabulary⁶ which provides reasoning to the ontology. Every operator is described in OpenMath vocabulary. However, this complex method of representing formula is not adopted as this model is just describing the method to calculate an index but it is not actually calculating it. Hence, the representation of the formula is done via a RDF literal as shown in Figure 3.3.

Similarly, the 'Required wavelength' concept in Figure 3.3 has literals attached to it which gives information about the start, end, and discrete wavelength with their respective unit. The unit literal is adopted from Ontology of Units of Measure (OM) 2.0 ontology⁷. OM2 ontology provides classes, properties, and instances for using measures and units. These application area can be geometry, mechanics, or economics, etc. OM2 is not the only ontology describing measurement units. There are other ontologies as well such as Quantities, Units, Dimensions and dataTypes (QUDT)⁸. This also contains the same elements for units of measurement as OM2. The units description for the semantic model can be suited for both of these ontologies but OM2 is chosen over QUDT. This is because the concept description of OM2 is much better than QUDT. Sometimes, the description for QUDT gets confusing and therefore, proper definitions or descriptions of the concepts is important (Keil et al., 2019).

The required wavelength has a connection with raw EO data which is provided by OGC:17-003r2. This concept is connected to wavelength information concept of the OGC:17-003r2 as shown in Figure 3.3. The OGC:17-003r2 did not adopt any unit ontology into its structure and predetermined the unit for measuring wavelength as meter. But as the ontology by its definition says that a model should explicitly represent the information so the unit ontologies should be adopted wherever the units of measurement is required and this also gives flexibility to the ontologies.

The overall RDF format of the T-Box of semantic model can be seen in Appendix B.

3.4 BUILDING THE A-BOX OF SEMANTIC MODEL

The A-Box of the ontology provides the T-Box definition with instances representing actual spectral indices. The representation of the A-Box is done according to the T-Box. The instances used here are from the selected spectral indices as displayed in Table 3.1 - NDVI, TSAVI, MSAVI, and CARI. But these are not the only spectral indices that this semantic model can describe. These indices have been chosen to illustrate all important aspects of the developed model. The T-box concepts are reflected in A-Box with the coloured square boxes. These concepts have literals as shown in white boxes. The A-Box for this semantic model can be seen in Appendix C.

The following spectral indices are instances for Index concept. All indices have formula which define their dependencies. For example, NDVI is represented by a formula, which has two dependencies (Required Wavelength), namely NIR and RED as shown in Figure 3.7.

TSAVI depends on spectral ranges as well as constants which depends on platform and sensor. TSAVI requires a formula that is dependent on wavelength and constants (which are dependent on an instrument as well as have fixed value) shown in Figure 3.8.

MSAVI has a formula which depends on the spectral ranges and a constant that depends on other formula with its dependencies as displayed in the Table 3.1. The A-Box for MSAVI is shown in Figure 3.9.

⁶<http://openmath.org/cd/rdf#resourceset>

⁷<https://github.com/HajoRijgersberg/OM>

⁸<https://github.com/qudt/qudt-public-repo>

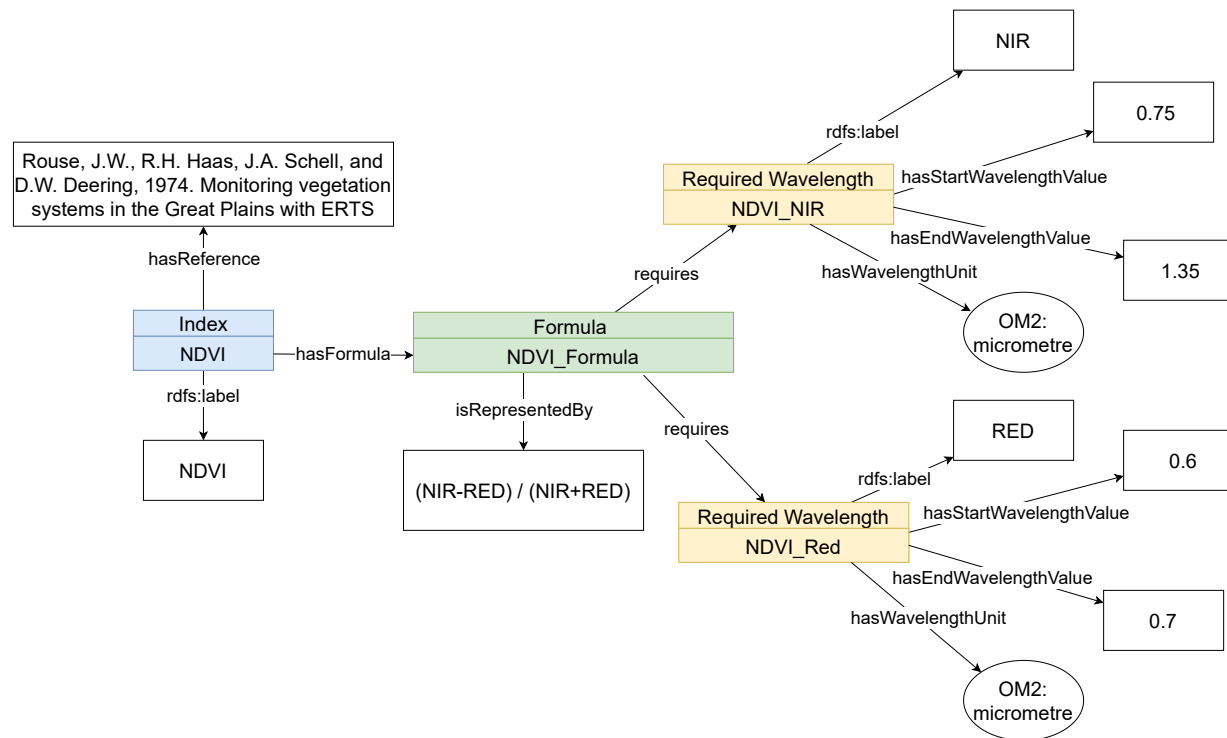


Figure 3.7: The dependencies for calculating NDVI.

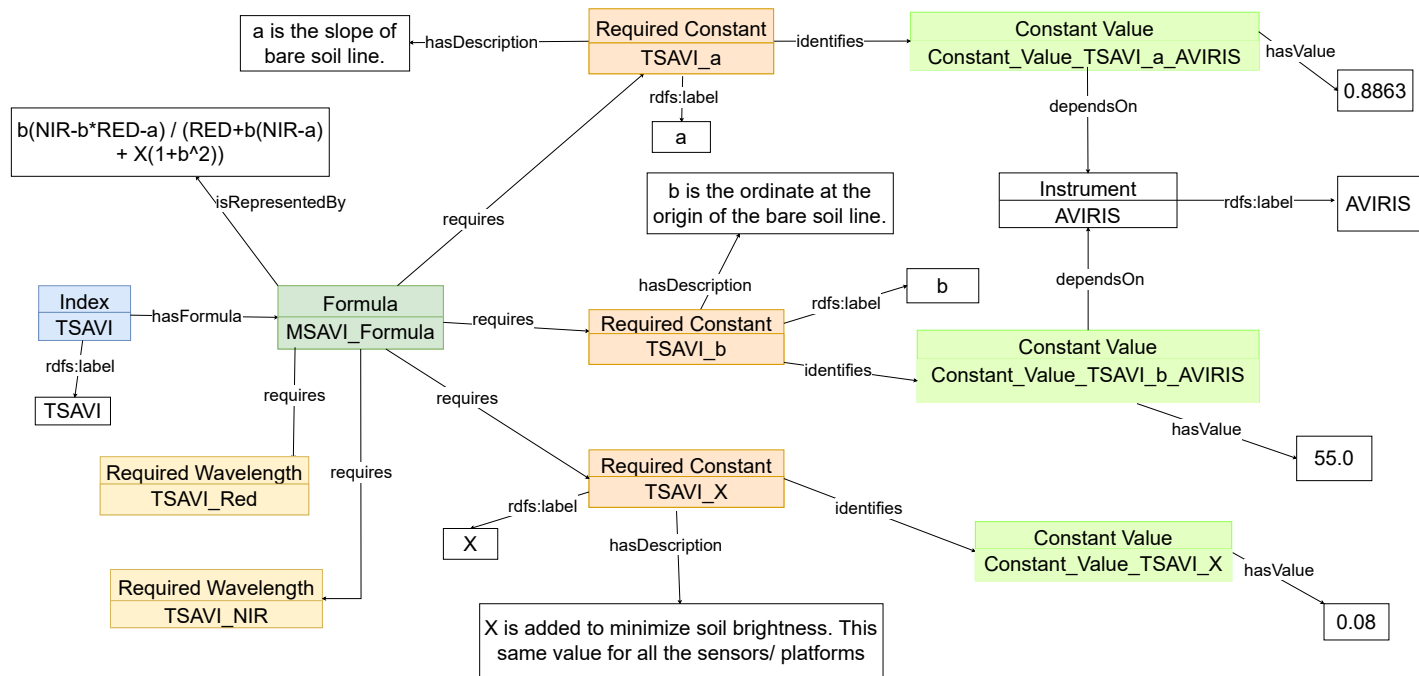


Figure 3.8: The dependencies for calculating TSAVI. All the literals of Required Wavelength are omitted here for visual representation.

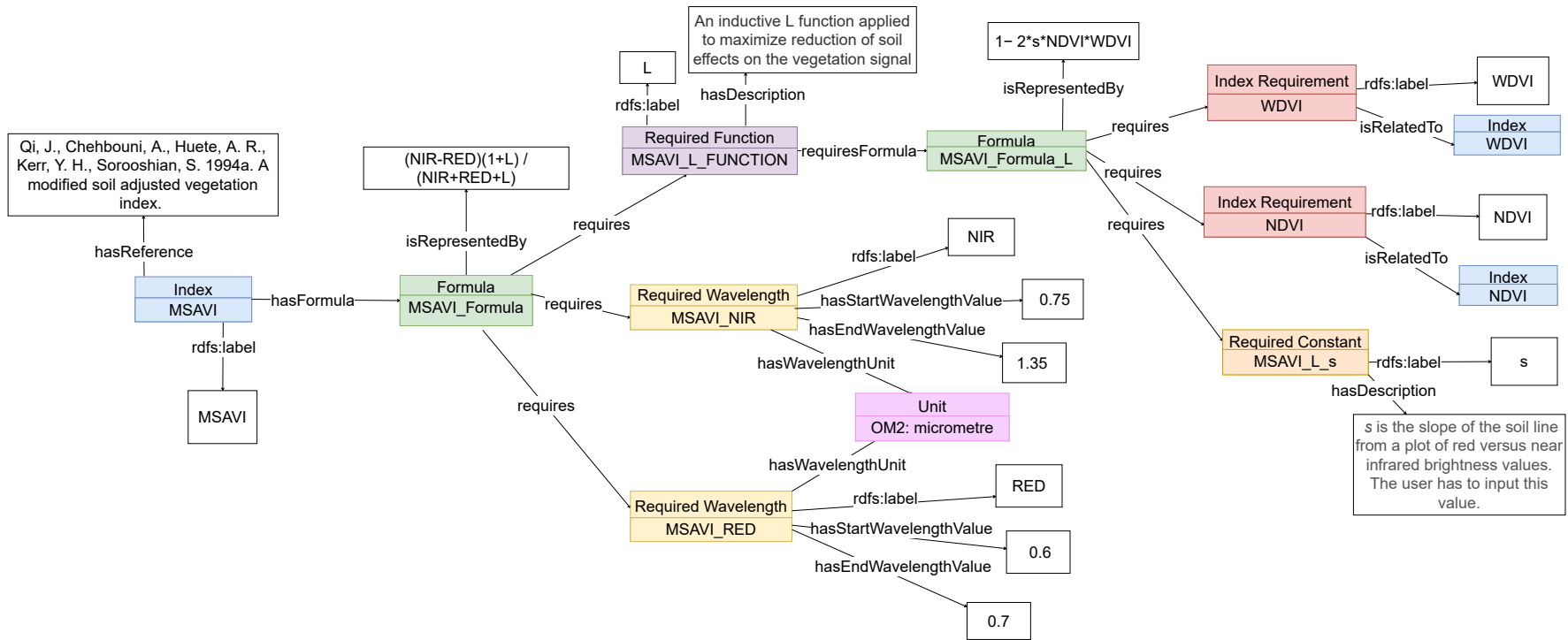


Figure 3.9: The dependencies for calculating MSAVI.

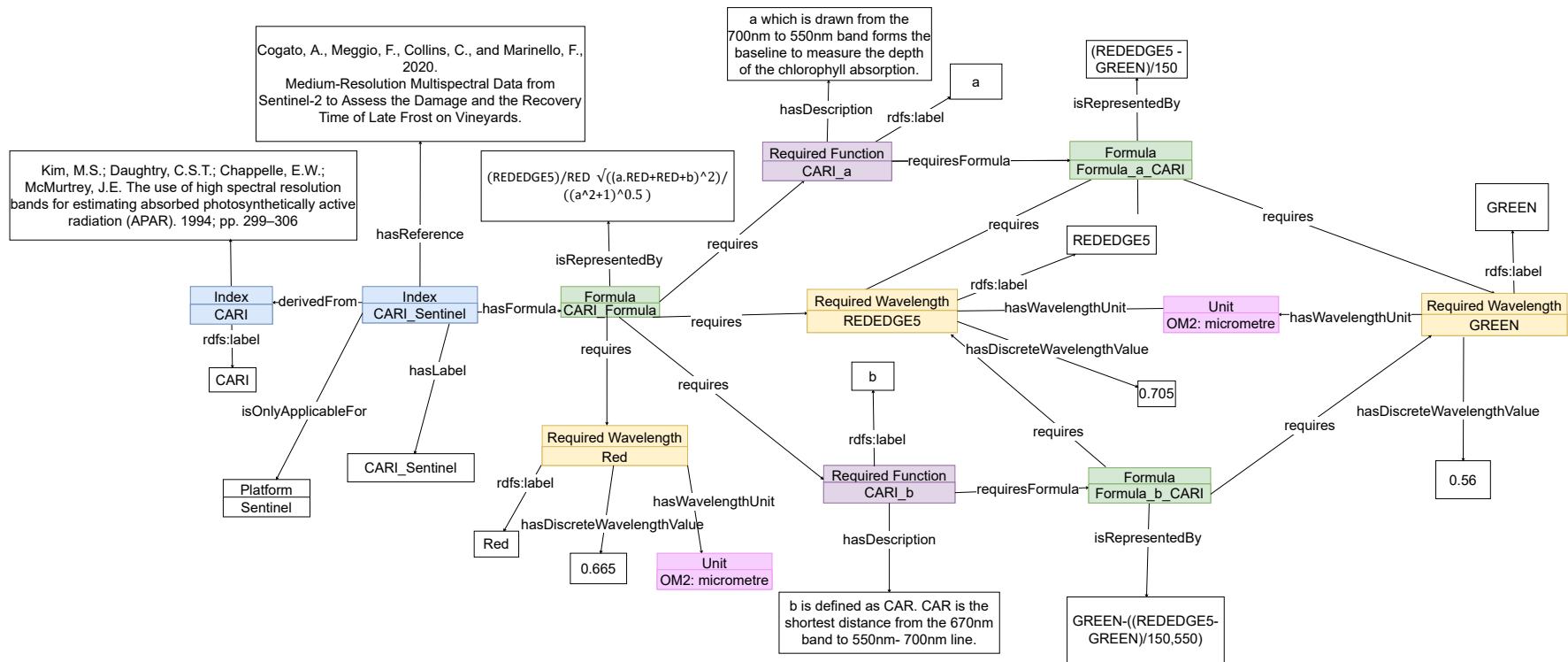


Figure 3.10: The dependencies for calculating CARI.

CARI formula displayed in the Table 3.1 is derived from the original formula for CARI proposed by Kim et al. (1994). The Figure 3.10 represents the T-box for CARI. This formula is specific for a particular platform, that is, Sentinel 2. Hence, the semantic model is capable to describe all the dependencies of the above spectral indices.

The tasks related to a spectral index are given in Table 3.6 in with their references.

Table 3.6: Spectral Indices associated with tasks.

Task	Index	Reference
Measure Biomass	NDVI TSAVI	Cabrera-Bosquet et al., 2011 Richardson et al., 1992
Assess Land Degradation	NDVI	Yengoh et al., 2015
Monitor Crops	NDVI MSAVI	Filgueiras et al., 2019 Richardson et al., 1992
Monitor Drought	MSAVI	Xue et al., 2017
Estimate Soil Moisture	TSAVI MSAVI	Richardson et al., 1992 Xue et al., 2017
Soil Erosion Analysis	MSAVI	Xue et al., 2017
Estimate Frost Damage	CARI	Cogato et al., 2020
Detect Vegetation Growth	NDVI	Xue et al., 2017
Quantify Forest supply	NDVI	Hyndavi et al., 2019
Detect Fire Zone	NDVI	Gabban et al., 2006
Estimate Chlorophyll Content	CARI	Xue et al., 2017

The tasks of the semantic model comes under different themes as given in Table 3.7.

Table 3.7: Tasks associated with themes.

Task	Theme
Measure Biomass	Vegetation; Agriculture
Assess Land Degradation	Agriculture; Soil
Monitor Crops	Agriculture
Monitor Drought	Soil
Estimate Soil Moisture	Soil
Soil Erosion Analysis	Soil
Estimate Frost Damage	Agriculture
Detect Vegetation Growth	Vegetation
Quantify Forest supply	Forest
Detect fire zone	Fire
Estimate Chlorophyll Content	Vegetation

3.5 VALIDATION AND VERIFICATION OF THE SEMANTIC MODEL

The main goal of any validation is to check for proper execution of the code. The validation of the developed semantic model is needed to check for the proper execution of T-Box and A-Box of the semantic model. The validation of the semantic model is done through a software GraphDB⁹ (free edition) which allows to import RDF and generate the outputs specified.

For this semantic model, the semantic model developed should be able to give the requested spectral index or the information about it like required constant, wavelength, function, or an index. The semantic model is also dependent on the raw EO data available. For this OGC: 17-003r2 ontology was used to populate the instances (raw EO data) (Table 4.1 in Analysis and Discussion chapter) taken from Earth Explorer USGS¹⁰ and Earth Online ESA¹¹. The OGC: 17-003r2 ontology was downloaded from OGC:17-003r2 (Appendix B.3) and was populated¹² by Protege. After importing both the RDF files (A-Box of semantic model and A-Box of the raw EO data) to GraphDB, the SPARQL queries were made to run on those RDF files. This has been discussed in detail in the next section.

The verification of includes the reviews, surveys, or inspections. This involves people who give their feedback as a user of something like a product. It helps to verify that the developed product is useful for the people it was developed.

The verification of the model developed uses only only task, theme, and index. This is because this semantic model developed weighs the complexity of the model towards the system and people do not need to know about the specified requirements to generate the spectral index. The tasks connected to an index and themes for the semantic model listed in Table 3.7 and Table 3.6 were verified through a survey. For this survey, the concepts index, tasks, and themes were displayed in EO4GEO BoK¹³ available on Living Textbook. The Living textbook is an online tool for displaying concept maps.

Aforementioned concepts of this semantic model were displayed in EO4GEO Bok because it provides existing concepts for possible applications of an index, which is an aspect that may help users to identify an appropriate index. So, following the idea of linked data, the model would use EO4GEO to cover this part of the model (similar as the OGC standard covered the description of raw EO data). The concepts of the semantic model were connected in EO4GEO Bok as displayed in Figure 3.12. The coloured concepts are from semantic model while white ones were already present in EO4GEO BoK.

This survey has been discussed in detail in the next chapter.

⁹<https://www.ontotext.com/products/graphdb/graphdb-free/>

¹⁰<https://earthdata.nasa.gov/>

¹¹<https://earth.esa.int/eogateway/>

¹²The A-Box for the raw EO data is not displayed here because of the length of this ontology but it will be included in the files submitted to the University

¹³<https://ltb.itc.utwente.nl/page/508/dashboard>

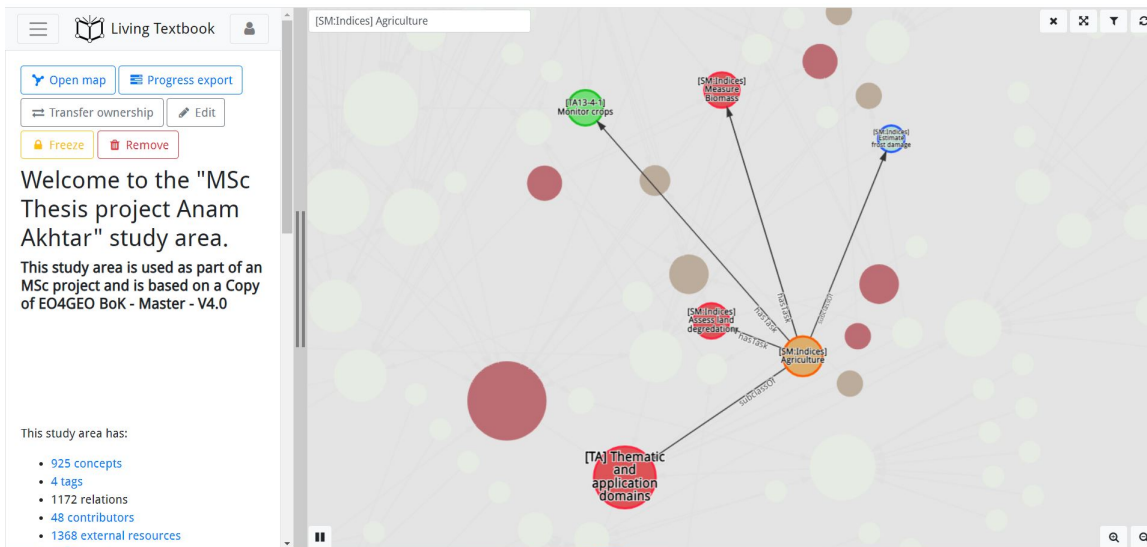


Figure 3.11: The interface for EO4GO Bok.

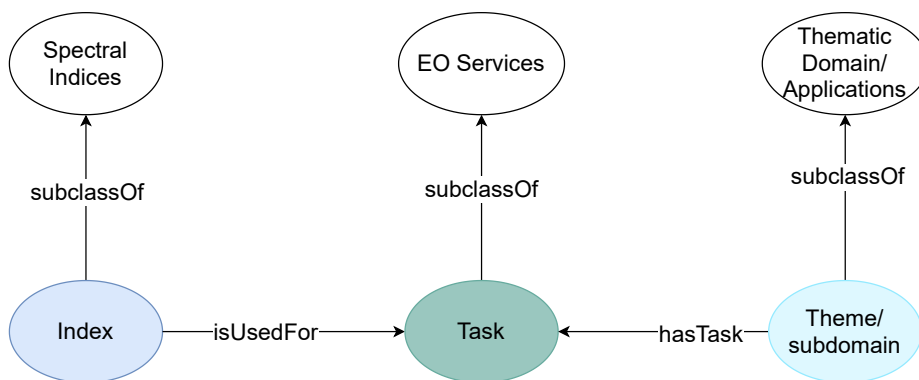


Figure 3.12: The connection of concepts of this semantic model and EO4GEO BoK.

Chapter 4

Analysis & Discussion

This chapter includes the validation and verification of the developed semantic model. It also reviews the research questions and objectives of the semantic model and discusses the limitations of the achieved results.

4.1 VALIDATION OF THE DEPENDENCIES OF THE SEMANTIC MODEL

Hüsemann et al.(2005) states that CQs are later used to validate the ontology, therefore, the CQs defined before for this semantic model are used to check the ability of the developed model to cover the relevant parts of the domain. Subsequently, the CQs will be translated to SPARQL queries. These queries are then run against the model including the example instances. This allows to verify that the CQs can indeed be phrased in and correctly answered by the developed model.

For validation only a few instances are used (see Section 3.4). The semantic model's description for calculation of the spectral indices depends on raw EO data as well. Therefore, the OGC:17003r2 ontology is populated with the instances representing real satellite products as displayed in Table 4.1. The products include information about platform and sensor with the spectral ranges used to acquire each one as well as other data such as acquisition time, and geographical boundaries. The products were chosen from all around the world, with different kinds of platforms, sensors, temporal, and spatial extents. Only few products are actually queried for, the other products are used to validate the model. These products serve as negative examples that queries should return and thus allow to check the accuracy of the queries.

Note that for clarity of the queries with their results have been displayed as listings and tables in this section. All the blank columns have been removed from the output tables for the clarity purpose. The queries and the results can also be seen on GraphDB interface in Appendix D.

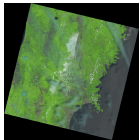
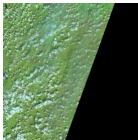

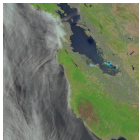
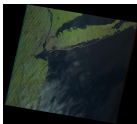
The geographical extent is represented by the geometry of the GeoSPARQL ontology¹ while the other terms related to platforms, sensors, and spectral bands are self explanatory.

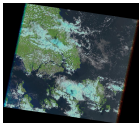


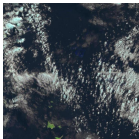

CQ1: What tasks are present for agriculture?

The tasks present for agriculture are Measure biomass, Monitor crop, Frost damage, and Assess Land Degradation. The SPARQL query imposed for CQ1 is given in Listing 4.1 which resulted in a list of tasks related to Agriculture as shown in Table 4.2. The query uses the `hasTask` property of the ontology to extract the tasks related to a theme. The query and results for CQ1 has also been displayed on GraphDB interface in Appendix D as Figure D.1.

¹<https://www.ogc.org/standards/geosparql>

Table 4.1: Input raw EO data.

Input image					
Product Name	Japan_20210603_Landsat	Netherlands_20210609_Sentinel2	Belgium_20210520_TerraSARX	Sanfrancisco_202210613_Sentinel2B	Newyork_20210507_Landsat7
Source	Earth Explorer USGS	Earth Explorer USGS	Earth Online ESA	Earth Explorer USGS	Earth Explorer USGS
Platform	Landsat 8	Sentinel 2B	TerraSAR-X	Sentinel2B	Landsat7
Sensor	OLI/TIRS	MSI	X-SAR	MSI	ETM+
Bounding Box	139.61038 39.95331, 142.32524 39.95407, 139.65118 37.82900, 142.28634 37.82970, 139.61038 39.95331	6.0047609 53.2119778, 7.6483415 53.2419635, 7.6785441 52.2551507, 6.0715951 52.2262109, 6.0047609 53.2119778	4.60642571349 50.8312542057, 4.60642571349 50.3109806637, 4.94746839754 50.3109806637, 4.94746839754 50.8312542057, 4.60642571349 50.8312542057	-123.0002276 37.9475896, -121.7506457 37.9409562, -121.7669948 36.9514795, -123.0002247 36.9578811, -123.0002276 37.9475896	-74.34298 41.30233, -72.10186 40.98409, -74.81792 39.69852, -72.62838 39.38779, -74.34298 41.30233
Begin Date	2021-06-03T01:14:31.6199750Z	2021-06-09T10:55:22.627Z	2021-05-20T17:36:12Z	2021-06-13T19:01:48.877Z	2021-05-07T14:42:05.7748750Z
End Date	2021-06-03T01:15:03.3899740Z	2021-06-09T11:08:11.083Z	2021-05-20T17:36:20Z	2021-06-13T19:04:56.481Z	2021-05-07T14:42:32.5288750Z

Input image					
Product Name	Singapore_20110820_Landsat4-5	Southafrika_20170114_TerraSARX	Lisbon_20000831_IRS1D	Greatbarrierreef_20210606_Sentinel2A	London_20210607_Sentinel2A
Source	Earth Explorer USGS	Earth Online ESA	Earth Online ESA	Earth Explorer USGS	Earth Explorer USGS
Platform	Landsat4-5	TerraSAR-X	IRS1D	Sentinel2A	Sentinel2A
Sensor	TM	X-SAR	PAN	MSI	MSI
Bounding Box	103.49383 2.35941, 105.13986 2.12003, 103.15585 0.76772, 104.80076 0.52850, 103.49383 2.35941	29.4154381094912 -26.6348936549586, 29.2747606276995 -26.0873281777721, 29.1071250793506 -26.122598375748, 29.2469007314022 -26.6703488028355, 29.4154381094912 -26.6348936549586	-9.1687456 38.6995454, -9.4036089 37.9505647, -9.1555823 37.9131648, -8.9183220 38.6617695, -9.1687456 38.6995454	150.11842 -20.7755038, 151.1724868 -20.7899668, 151.1601908 -21.7815368, 150.0990517 -21.7663169, 150.11842 -20.7755038	-0.0663842 52.3140378, 1.5399797 52.2631374, 1.4420952 51.2789252, -0.1297842 51.3280673, -0.0663842 52.3140378
Begin Date	2011-08-20T03:04:18.4588800Z	2017-01-14T16:32:12Z	2000-08-31T11:50:33Z	2021-06-06T00:11:10.455Z	2021-06-07T11:04:45.550Z
End Date	2011-08-20T03:04:45.0717500Z	2017-01-14T16:32:21Z	2000-08-31T11:50:33Z	2021-06-06T00:19:06.681Z	2021-06-07T11:17:34.005Z

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 select ?theme ?task ?taskLabel
5 where {
6   ?theme a eoi:Theme .
7   ?theme rdfs:label "Agriculture"@en.
8   ?theme eoi:hasTask ?task .
9   ?task rdfs:label ?taskLabel .
10 }

```

Listing 4.1: SPARQL query for CQ1

Table 4.2: Results for CQ1

theme	task	taskLabel
:Agriculture	:Assess_Land_Degradation	"Assess Land Degradation"@en
:Agriculture	:Estimate_Frost_Damage	"Estimate Frost Damage"@en
:Agriculture	:Measure_Biomass	"Measure Biomass"@en
:Agriculture	:Monitor_Crops	"Monitor Crops"@en

CQ2: Which spectral indices are available for crop monitoring?

The only spectral index available for crop monitoring is NDVI. The SPARQL query imposed for CQ2 is given in Listing 4.2 which resulted in a list of available spectral indices related to crop monitoring (only NDVI in this case) as shown in Table 4.3. The query uses `isUsedFor` property to extract the task related to an index. The query and results for CQ2 has also been displayed on GraphDB interface in Appendix D as Figure D.2.

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 select ?task ?index ?IndexName
4 where {
5   ?task a eoi:Task .
6   ?task rdfs:label "Monitor Crop"@en.
7   ?index eoi:isUsedFor ?task .
8   ?index rdfs:label ?IndexName .
9 }

```

Listing 4.2: SPARQL query for CQ2

Table 4.3: Results for CQ2

task	index	IndexName
:Monitor_Crop	:NDVI	"NDVI"@en

CQ3: Given that the user wants to know about the origin of the spectral index, what is the reference of that spectral index?

The SPARQL query imposed for CQ3 is given in Listing 4.3 which resulted in reference related to NDVI as shown in Table 4.4. The query uses the `hasReference` property to extract the reference related to an index. The query and results for CQ3 has also been displayed on GraphDB interface in Appendix D as Figure D.3.

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 select ?index ?reference
5 where {
6   ?index a eoi:Index .
7   ?index rdfs:label "NDVI"@en .
8   ?index eoi:hasReference ?reference .
9 }

```

Listing 4.3: SPARQL query for CQ3

Table 4.4: Results for CQ3

index	reference
:NDVI	"Rouse, J.W., R.H. Haas, J.A. Schell, and D.W. Deering, 1974. Monitoring vegetation systems in the Great Plains with ERTS"@en

CQ4: Which of the datasets are present for May 2021 assuming it the harvesting month of the crop?

The SPARQL query imposed for CQ4 is given in Listing 4.4 which resulted in a list of datasets for May 2021 as shown in Table 4.5. The query uses the concepts and properties of OGC:17-003r2 ontology whose namespace is prefixed with ‘eop’. The connection in the query is same as displayed in Figure 3.5. The query starts with the concept EarthObservation, followed by Properties, AcquisitionInformation, AcquisitionParameters, and TemporalInformation. The OGC:17-003r2 ontology uses the ical ontology² to extract the date and time of the dataset acquired. This date and time uses an intersection filter of SPARQL to extract data for May 2021. The query and results for CQ4 has also been displayed on GraphDB interface in Appendix D as Figure D.4.

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX eop: <http://www.opengis.net/ont/eo-geojson/1.0#>
3 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
4 PREFIX ical: <http://www.w3.org/2002/12/cal/ical#>
5 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
6 select ?dataset ?begin ?end ?DatasetName
7 where {
8   ?dataset a eop:EarthObservation ;
9   rdfs:label ?DatasetName .
10  ?dataset eop:properties ?property .
11  ?property eop:acquisitionInformation ?acinfo .
12  ?acinfo eop:acquisitionParameters ?acpara .
13  ?acpara eop:temporalInformation ?temp .
14  ?temp ical:beginningDateTime ?begin .
15  ?temp ical:endingDateTime ?end .
16  FILTER(
17    (?begin > "2021-05-01T00:00:00Z"^^xsd:dateTime)
18    && ("2021-05-31T23:59:00Z"^^xsd:dateTime > ?begin)
19    && ("2021-05-01T00:00:00Z"^^xsd:dateTime < ?end)
20    && (?end < "2021-05-31T23:59:00Z"^^xsd:dateTime)
21  )
22 }

```

Listing 4.4: SPARQL query for CQ4

²<http://www.w3.org/2002/12/cal/ical>

Table 4.5: Results for CQ4

dataset	begin	end	DatasetName
:PRODUCT/5	"2021-05-07T14:42:05.7748750Z"	"20210507T14:42:32.5288750Z"	"Newyork_20210507_Landsat7"
:PRODUCT/3	"2021-05-20T17:36:12Z"	"2021-05-20T17:36:20Z"	"Belgium_20210520_TerraSARX"

CQ5: Given that NDVI is chosen for crop monitoring, what constraints are used on the input data to compute it?

The SPARQL query imposed for CQ5 is given in Listing 4.5 which resulted in a list of constraints related to NDVI as shown in Table 4.6. The query uses the `requires` property to list the dependencies of the spectral indices. In this case, they are limited to wavelengths whose start and end values with their units are also displayed. The `BIND` function is used to specify what kind of requirement is used. In this case it is 'Required_Wavelength'. The `OPTIONAL` keyword is used because there are two ways the wavelengths can be populated in the model. One which have start and end wavelength. The other which have discrete wavelength. Note that the `OPTIONAL` keyword used can not only give the requirements for NDVI but also for other spectral indices as well. For NDVI, as there are only two required wavelengths, the query only gives the output for NIR and RED. The query and results for CQ5 has also been displayed on GraphDB interface in Appendix D as Figure D.5.

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
5 select ?index ?reqType ?requirement ?reqLabel
6        ?wvStart ?wvEnd ?wvUnit
7        ?nestedFormulaNode ?nestedFormula
8 where {
9   ?index rdfs:label "NDVI"@en ; eoi:hasFormula ?formula .
10  ?formula eoi:requires ?requirement .
11  ?requirement rdfs:label ?reqLabel .
12  OPTIONAL {
13    ?requirement a eoi:Required_Wavelength .
14    BIND( eoi:Required_Wavelength AS ?reqType ).
15    ?requirement eoi:hasWavelengthUnit ?wvUnit .
16    OPTIONAL {
17      ?requirement eoi:hasDiscreteWavelengthValue ?wvDiscValue ;
18    }
19    OPTIONAL {
20      ?requirement eoi:hasEndWavelengthValue ?wvEnd ;
21      eoi:hasStartWavelengthValue ?wvStart .
22    }
23  }
24  OPTIONAL {
25    ?requirement a eoi:Required_Function ;
26    eoi:requiresFormula ?nestedFormulaNode .
27    ?nestedFormulaNode eoi:isRepresentedBy ?nestedFormula .
28    BIND( eoi:Required_Function AS ?reqType ).
29  }
30  OPTIONAL {
31    ?requirement a eoi:Required_Constant .
32    BIND( eoi:Required_Constant AS ?reqType ).
33  }
34  OPTIONAL {
35    ?requirement a eoi:Required_Index .
36    BIND( eoi:Required_Index AS ?reqType ).
37  }
38 }

```

Listing 4.5: SPARQL query for CQ5

Table 4.6: Results for CQ5. Empty bindings omitted.

index	reqType	requirement	reqLabel	wvStart	wvEnd	wvUnit
:NDVI	:Required_Wavelength	:NDVI_NIR	"NIR"	"0.75"	"1.35"	om2:micrometre
:NDVI	:Required_Wavelength	:NDVI_RED	"RED"	"0.6"	"0.7"	om2:micrometre

CQ6: Given constraints require another index like in MSAVI, is such an index available?

The SPARQL queries imposed for CQ6 is divided into two parts. The first one gives the list of constraints, as given in Listing 4.6 which resulted in a list of constraints related to MSAVI formula as shown in Table 4.7. One constraint, :MSAVI_L_Function, is described by a formula itself, so it has to be checked as well. The second query shows that a function is required which shows the two other indices as requirements in Listing 4.7. This resulted in a list of required indices related to the function as shown in Table 4.8. Note that second query for the function only checks for the required index and not everything present in the formula. The query and results for CQ6 part 1 and part 2 has also been displayed on GraphDB interface in Appendix D as Figure D.6 and Figure D.6, respectively.

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 PREFIX eop: <http://www.opengis.net/ont/eo-geojson/1.0#>
5 select ?index ?MSAVIformula ?indexRequirement ?reqType
6 where {
7   ?index a eoi:Index;
8         rdfs:label "MSAVI"@en;
9         eoi:hasFormula ?formula.
10  ?formula eoi:isRepresentedBy ?MSAVIformula.
11  ?formula eoi:requires ?indexRequirement.
12  ?indexRequirement a ?reqType.
13  ?reqType rdfs:subClassOf eoi:Requirement.
14 }

```

Listing 4.6: SPARQL query for CQ6 part 1

Table 4.7: Results for CQ6 part 1

index	MSAVIformula	indexRequirement	reqType
:MSAVI	"(NIR-RED)(1+L)/(NIR+RED+L)"	:MSAVI_L_Function	:Required_Function
:MSAVI	"(NIR-RED)(1+L)/(NIR+RED+L)"	:MSAVI_NIR	:Required_Wavelength
:MSAVI	"(NIR-RED)(1+L)/(NIR+RED+L)"	:MSAVI_RED	:Required_Wavelength

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 PREFIX eop: <http://www.opengis.net/ont/eo-geojson/1.0#>
5 select ?index ?nestedFormula ?indexRequirement ?requiredIndex ?ReqIndexName
6 where {
7   ?index a eoi:Index;
8         rdfs:label "MSAVI"@en;
9         eoi:hasFormula ?formula.
10  ?formula      eoi:requires      ?funcRequirement .
11  ?funcRequirement a      eoi:Required_Function .
12  ?funcRequirement eoi:requiresFormula ?funcformula .
13  ?funcformula      eoi:isRepresentedBy ?nestedFormula .
14  ?funcformula      eoi:requires      ?indexRequirement .
15  ?indexRequirement a      eoi:Required_Index .
16  ?indexRequirement eoi:isRelatedTo  ?requiredIndex .
17  ?requiredIndex    rdfs:label      ?ReqIndexName .
18 }

```

Listing 4.7: SPARQL query for CQ6 part 2

Table 4.8: Results for CQ6 part 2

index	nestedFormula	indexRequirement	requiredIndex	?ReqIndexName
:MSAVI	"1-2*s*NDVI*WDVI"	:MSAVI_L_Index_Req_NDVI	:NDVI	"NDVI"@en
:MSAVI	"1-2*s*NDVI*WDVI"	:MSAVI_L_Index_Req_WDVI	:WDVI	"WDVI"@en

CQ7: Given the constraints describes wavelengths necessary for an index like TSAVI, which sensors and platforms provide compatible data?

The SPARQL query imposed for CQ7 is divided into two parts. One where the wavelength constraints required for TSAVI are extracted using the query shown in Listing 4.8 which resulted in Table 4.9. Another one is to extract the sensors and their platforms available for the wavelength range of TSAVI extracted in part one as shown in Listing 4.9 with the result displayed in Table 4.10. The unit used for the wavelength is micrometer for this semantic model. While OGC: 17-003 fixed their wavelength unit to meter. Therefore, if comparison is done for the raw EO data, the units in the semantic model needs to be converted to from micrometer to meter. The use of OPTIONAL keyword is already explained in CQ5. The query and results for CQ7 part 1 and part 2 has also been displayed on GraphDB interface in Appendix D as Figure D.8 and Figure D.9, respectively.

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX eop: <http://www.opengis.net/ont/eo-geojson/1.0#>
4 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
5 select ?index ?require ?start ?end ?Discrete ?WaveUnit where {
6   ?index a eoi:Index rdfs:label "TSAVI"@en; eoi:hasFormula ?formula.
7   ?formula eoi:requires ?require .
8   ?require a eoi:Required_Wavelength; eoi:hasWavelengthUnit ?WaveUnit;
9   OPTIONAL { ?require eoi:hasDiscreteWavelengthValue ?Discrete; }
10  OPTIONAL {
11    ?require eoi:hasEndWavelengthValue ?end ;
12    eoi:hasStartWavelengthValue ?start .
13  }
14 }

```

Listing 4.8: SPARQL query for CQ7 part 1

Table 4.9: Results for CQ7 part 1

index	require	start	end	WaveUnit
:TSAVI	:TSAVI_NIR	"0.75"^^xsd:decimal	"1.35"^^xsd:decimal	om2:micrometre
:TSAVI	:TSAVI_RED	"0.6"^^xsd:decimal	"0.7"^^xsd:decimal	om2:micrometre

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX eop: <http://www.opengis.net/ont/eo-geojson/1.0#>
4 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
5 select distinct ?sensor ?platformName ?sensorName
6 where {
7   ?aqi      a                eop:AcquisitionInformation;
8           eop:instrument ?sensor;
9           eop:platform   ?plat .
10  ?sensor  eop:instrumentShortName ?sensorName .
11  ?plat    eop:platformShortName ?platformName .
12  ?aqi     eop:acquisitionParameters ?aqipara .
13  ?aqipara eop:wavelengths ?Wavelength .
14  OPTIONAL {
15    ?Wavelength eop:discreteWavelength ?wvDiscValue ;
16  }
17  OPTIONAL {
18    ?Wavelength eop:endWavelength ?wvEnd ;
19    eop:startWavelength ?wvStart .
20  }
21  VALUES (?start ?end) {
22    ("0.00000060"^^xsd:double "0.00000070"^^xsd:double)
23    ("0.00000075"^^xsd:double "0.00000135"^^xsd:double)
24  }
25  FILTER(
26    (?start < ?wvEnd)
27    && (?end > ?wvStart)
28  )
29 }

```

Listing 4.9: SPARQL query for CQ7 part 2

Table 4.10: Results for CQ7 part 2

sensor	platformName	sensorName
:INST/OLI%2FTIRS	"Landsat8"	"OLI/TIRS"
:INST/TM	"Landsat4-5"	"TM"
:INST/ETM%2B	"Landsat7"	"ETM+"
:INST/PAN	"IRS"	"PAN"

CQ8: Given the compatible platform like for Sentinel 2B, which raw datasets were created by that platform?

The SPARQL query imposed for CQ8 is given in Listing 4.10 which resulted in a list of datasets available for Sentinel 2B as shown in Table 4.11. The query uses platformShortName to identify the platform Sentinel 2B. The query and results for CQ8 has also been displayed on GraphDB interface in Appendix D as Figure D.10.

```

1 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
2 PREFIX eop: <http://www.opengis.net/ont/eo-geojson/1.0#>
3 select distinct ?dataset ?datasetName
4 where {
5   ?dataset a eop:EarthObservation; eop:properties ?prop.
6   ?prop eop:acquisitionInformation ?aqc.
7   ?aqc eop:platform ?inst.
8   ?inst eop:platformShortName "Sentinel2B".
9   ?dataset rdfs:label ?datasetName.
10 }

```

Listing 4.10: SPARQL query for CQ8

Table 4.11: Results for CQ8

dataset	datasetName
:PRODUCT/2	"Netherlands_20210609_Sentinel2B"
:PRODUCT/4	"Sanfrancisco_202210613_Sentinel2B"

CQ9: Given input datasets satisfying the constraints, are those available for the given spatial extent?

The SPARQL query imposed for CQ8 is given in Listing 4.11 which resulted in a list of datasets available as shown in Table 4.12. So, as the CQ9 state that the resulted dataset in CQ8 should also satisfy the specified spatial extent. The use case stated Assen, Netherlands. So, the bounding box was fetched from Bounding Box site³ in OGC Well-known text (WKT) format which used for representing geometries. The query uses the gsp ontology to identify the geometry of the datasets with a given constrain that the product must be of Sentinel 2B as stated in CQ8. The geof ontology provides the function which checks for the intersecting geometries. As only two datasets should be checked so VALUES keyword in line 10 of the listing is used. The query and results for CQ9 has also been displayed on GraphDB interface in Appendix D as Figure D.11.

```

1 PREFIX geof: <http://www.opengis.net/def/function/geosparql/>
2 PREFIX eoi: <http://w3id.org/eoi/>
3 PREFIX eop: <http://www.opengis.net/ont/eo-geojson/1.0#>
4 PREFIX gsp: <http://www.opengis.net/ont/geosparql#>
5 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
6 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
7 PREFIX eoip: <http://w3id.org/eoi#PRODUCT/>
8 SELECT ?data ?dataName ?WKT1
9 WHERE {
10   VALUES ?data { eoip:2 eoip:4 }
11   ?data a eop:EarthObservation;
12   rdfs:label ?dataName;
13   gsp:hasGeometry ?geom1.
14   ?geom1 gsp:asWKT ?WKT1.
15   FILTER( geof:sfIntersects(?WKT1, ""POLYGON((
16     6.3967 53.0407, 6.7592 53.0407,
17     6.7592 52.923, 6.3967 52.923,
18     6.3967 53.0407))""^^gsp:wktLiteral)).
19 }

```

Listing 4.11: SPARQL query for CQ9

³<https://boundingbox.klokantech.com/>

Table 4.12: Results for CQ9

data	dataName	WKT1
:PRODUCT/2	"Netherlands_20210609_Sentinel2B"	"Polygon((6.0047609 53.2119778, 7.6483415 53.2419635, 7.6785441 52.2551507, 6.0715951 52.2262109, 6.0047609 53.2119778))"

All the CQs gave the required outputs. These outputs were based on different requirements related to an index and sometimes related to EO data. The ontology for describing spectral indices was successful in data integrating as along with raw EO data ontology (OGC-17003r2), it was able to give the desired outputs by the users. Overall, the structure of the semantic model to describe the components of spectral indices with its thematic descriptions have proven to be efficient in delivering the spectral indices as EO product.

4.2 VERIFICATION OF THE THEMATIC DOMAIN OF SPECTRAL INDICES

The semantic model also identifies spectral indices related to a thematic domain or application domain. The tasks and themes of a spectral index for the semantic model listed in Table 3.7 and Table 3.6 were verified through a survey. This survey was divided into four main sections. The first dealt with questions related to the structuring of the model for index, tasks, and themes. In the second to fourth section, participants were asked to evaluate the model from three different perspectives: as a consumer, as a producer, and as a non-expert consumer of spectral indices. The entire survey form can be seen in Appendix A. This survey was done online on Maptionnaire tool⁴.

In total, eight participants appeared for the survey. The survey recorded their expertise as well and amongst them, most of the participants were from a Geoinformatics background as shown in Figure 4.1.

The following questions and answers were a part of section one.

Q1. Do you think the above mentioned three layers (coloured) of the semantic model are enough to describe possible uses of spectral indices?

There was only one no as an answer with a suggestion that an index also has alternative applications (or tasks) it could be beneficial for. So, 'isUsefulFor' should be considered as a property instead. This suggestion can be incorporated in the semantic model by replacing the 'isUsedFor' property to 'IsUsefulFor'.

Q2. Currently, tasks are modelled as a kind of EO services and applications. An example of this is given in the link provided below. Do you agree with this view? Please use the link below.

There were two participants disagreeing with it but one of them did not understand the question properly. The other participant suggested - 'A task requires a verb because you do something. An application does not. Perhaps this needs reconsideration?' There were tasks already listed in EO4GEO BoK as EO services and applications. This was the reason to add the tasks of the semantic model developed here were added to it. If the contributors of EO4GEO BoK think that this is the connection best fit for them then nobody can negate it.

⁴<https://maptionnaire.com/>



Figure 4.1: Number of participants and their expertise.

Q3. As the survey participant would you be able to add new tasks and subdomains, based on your expertise.

All the participants agreed to this.

In Section two, the following questions were asked.

Q1. Crop monitoring is placed under the Agriculture theme. Do you think that this connection is right? Please use the link below.

There was one feedback with a no with a suggestion that ‘it could be categorized under remote sensing theme, and the reason is that the main science (or expertise) used here is remote sensing’. This could be incorporated but then again remote sensing also has different themes. Therefore, to satisfy the participant’s suggestion, the themes of the semantic model can be grouped as a subclass of Remote sensing domains. However, the majority agreed with the current approach.

Q2. Assuming that he choose one index, for example MSAVI for crop monitoring, does the description for MSAVI provide enough information to calculate it? Please use the link below.

There was one feedback with a no with a suggestion that -‘Perhaps the information can be separated for a specialist and a novice. The IDB also included links to the spectral bands to use per sensor.’ This is something that this research thesis do not want in the system because this would mean linking an index to the particular band of a sensor. It will contradict the motivation of this research thesis as this would mean the sensors/ platforms are hard-coded into the model. The aim is that users get information which is dynamically asserted by the system irrespective of the spectral bands of the sensors.

In section three, the following questions were asked.

Q1. Are the description for tasks and themes sufficient enough to describe the new spectral index? In this case the new index is related to soil texture. Please use the link below.

There was one feedback with a no with a suggestion that -‘to be able to add a new index with task, and domain and relationships there should be a ‘page’ with a how to and what to do and what not to do. Perhaps with an example as starting from Soil might be appropriate but typing texture in the search leaves me puzzled.’ According to the participant, the interface EO4GEO was not suitable for this. The participant wants an interface that can not only display the connections to the model but also should have a clickable link to a new page. This page should contain a guide to the user describing various the workflow of the model which is basically a tutorial.

Q2. As a developer of the spectral index, would it be useful to specify tasks or application areas where the index can not be used for?

This question provided interesting feedback. Half of the participants agreed and half of them disagreed as shown in Figure 4.2. Among the four participants who voted for no, three of them gave reasons as well. The first one being - 'Because it would be misleading'. It can be inferred that the participant is concerned that the user might misuse the spectral index because of his/her confusion. The other two reasons are that - 'it can be a rather long list' and 'to avoid redundancy'. But there were equal amount of participants agreeing to it. So, it can be concluded that people should not have to decide for every task, but can explicitly negate some tasks that have proven erroneous in the past. There might still be a third group of tasks for which no information was added so far, so they might belong either to the white or the blacklist.

Q1. Are the spectral indices related to Agriculture domain correct?

There was one feedback with a no with a feedback that - 'I cannot assess. With Agriculture I see three hastask. I would have to go into these and check whether these are correct and I would have think of other tasks that are missing (perhaps already in the model but not linked).' The participant gave answer for something that is not asked so maybe he / she did not understand the question well as the question does not ask for missing data but the correct one.

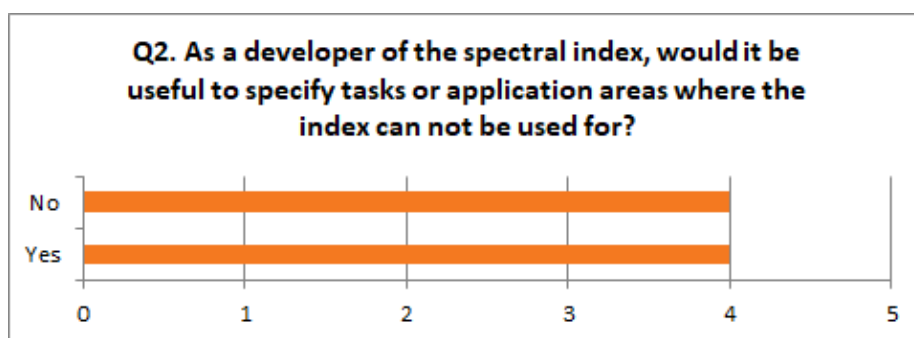


Figure 4.2: Number of participants and their feedback to Q2 of Section 3.

Q2. Are the tasks related to a particular index correct? These are the tasks under Agriculture like - NDVI: Monitor Crops, NDVI: Assess Land Degradation, NDVI: Measure Biomass, TSAVI: Measure Biomass, and MSAVI: Monitor Crops.

There was one feedback with a no with a feedback that - 'Could be. I am not the expert to define this. How to use NDVI in the various tasks is quite different to capture in a single text.' By just using 'isUsedFor' property, the participant can tag different tasks to the spectral indices without explaining anything in text as the semantic models are self-explanatory.

Do you have any feedback while exploring the semantic model displayed in EO4GEO BoK?

Only one response was given - 'I must admit that my role in this survey was not always clear to me. Was I supposed to be the non-expert, or the expert.' Different roles (expert/non-expert) were assigned to the same surveyor. This might be a point of confusion for the participant.

Through the survey it can be concluded that the majority participants agreed upon the structuring of the model for index, theme, and tasks. Some suggestions were good and can be incorporated in future like the one where isUsefulFor relation is suggested instead of isUsedFor. The most challenging one to conclude was the one with same for and against. Therefore, summarizing for both was a bit challenging. Moreover, the criticisms also added value as it projected the overall goal of semantic model where one of participants wanted the sensors to be added in the formula. Thus, feedbacks help to reflect upon the work done in this thesis. Overall, it can be

stated that the semantic model for the user interaction explicitly describes the thematic domain of spectral indices as it had a majority positive response by the participants. This verification was useful because by specifying different use cases in the survey, the participants replied as the user of the spectral indices connected to tasks and themes.

4.3 ANSWERS TO THE RESEARCH QUESTIONS

The research objectives and questions were identified in Section 1.2. These are answered here:

1. To develop a semantic model which can represent explicitly, the given information of indices and raw EO data in the knowledge base.

- 1.1. What characteristics are available for the raw data?

The geoportals provide abundance of metadata for EO products like orbit parameters, polarisation channel, sensor, platform, temporal resolution, cloud cover, and spatial extent, etc. But not all metadata is needed to describe the semantic model for spectral indices. Therefore, to analyze the characteristics of raw EO data needed, a use case was used where CQs were formulated. These CQs identified few concepts related to EO data that are required for this semantic model. The characteristics needed for raw EO data are the temporal and spatial information of the dataset along with the type of platform, and instrument required to capture it. These instrument also includes the wavelength information which gives the start and end wavelength of their spectral bands. Hence, it can be concluded that the characteristics for raw data are Temporal information, Wavelength information, Spatial boundary, Platform, Instrument, and Wavelength information.

- 1.2. What requirements do the individual indices have?

The information related to spectral indices can be found on many online resources but all of them have almost common grounds for defining their requirements. The requirements for spectral indices are specified in Section 3.2 where the requirement concept is further explained in details. The requirements depend on the formula of a spectral index. The different requirements for the spectral indices are constant values, other indices, function, and wavelength information of the spectral bands. The spectral bands contain more information like the start wavelength, end wavelength, and the respective unit.

- 1.3. Are there other parameters involved that might not be determined automatically and require further user interaction?

Yes, there can be parameters that require user interaction. One such example is explained for MSAVI in Figure 3.9. The required constant concept MSAVI_L_S requires the value for slope. The exact value depends on a sensor and so can not be given as a fixed value within the model.

- 1.4. Can the model be provided with meaningful default values for those parameters?

No, the model cannot provide the default values but it describes the kind of parameter required. Considering the MSAVI example again. The model describes required constant value and also specifies that the user needs to input the value for the slope.

2. To interpret the above knowledge into a semantic model which describes all the components required for it .
 - 2.1. What are the different components identified for the semantic model? As the intention of this model is to describe the spectral indices whose ultimate aim is on-demand computation of the spectral indices, therefore the classes, properties, and instances are chosen accordingly. The dependent classes for this purpose are Index, Formula, Required wavelength, Required index, Required function, and Required constants. This model also depends on the classes of OGC: 17-003r2 (Figure 3.5). Platform and Instrument classes are connected via Remote Sensing device superclass to the semantic model. The other classes from OGC: 17-003r2 are EarthObservation, WavelengthInformation, TemporalInformation with intermediate classes connecting them as Properties, AcquisitionInformation, and AcquisitionParameters. The properties connecting them is displayed in Figure 3.5. The classes related to thematic domain are tasks and themes which is displayed in Table 3.6 and Table 3.7, respectively. The properties connecting these classes are identified in Table 3.3, Table 3.4, and Table 3.5. In a production setting this model would need to be instantiated with a wide selection of indices. For this thesis only four examples were chosen as discussed in Section 4.1.
 - 2.2. What are the logical descriptions behind the above mentioned parts? The logical description is the reason behind the selection of the above stated components for the semantic model. The logical description for selecting only the above components is explained in Section 3.1. This section introduced the context of the semantic model, that is, spectral indices. The use case analysis identified the different CQs related to it. These CQs gave components for the semantic model and eventually also conveyed the reason for using only those concepts. The CQs with the concepts are given in Table 3.2. The glossary for the concepts are mentioned in Section 3.2.
 - 2.3. What are the challenges in representing the semantic model? The most challenging part was to add the tasks and themes related to the spectral indices because people with different backgrounds might use some other structure as well as different semantics for them. This is well reflected in the survey that participants had their own way of defining the semantic model for spectral indices. The survey gave an impression that not all would agree to the structure of the model and at the same time some of their constructive remarks were beneficial to reflect the goal of this research.
The other challenge was collecting information about spectral indices from different sources. As only the original papers for the spectral indices are mentioned here so getting access to them was a bit difficult especially CARI (Kim et al., 1994).
3. To evaluate whether the semantic model can provide information about the required index selected by the user.
 - 3.1. In what ways, the user can interact with the semantic model that gives a meaningful output? Currently, there are two ways a user user can interact with the model more might be added in the future. One is through SPARQL queries to the model itself as done for the CQs in Section 4.1. Another one is through the EO4GEO BoK. There are total eleven queries for CQs which returns the meaningful output. Some of these queries have been visualized in Figure 3.2. In this figure, the user can interact with the system from any way possible. There are four possible ways displayed in the figure. They are - list of indices related to a task, the list sensors according to the desired spec-

tral range, list of datasets available according to the sensors, and the list of raw datasets available within a spatial and temporal extent. However, the model is not limited to the predefined queries as CQs just are examples of different possibilities to interact with the model.

The EO4GEO BoK displays the thematic domains of the spectral indices. Only the index, tasks, and themes concepts are displayed here. The connections for EO4GEO BoK has been displayed in Section 3.5 in Figure 3.12. Only few instances are added to them for illustration purpose but an exhaustive list can be added to the semantic model in the same manner.

- 3.2. Which parts can be fully automated and require no user input? The user can decide the manner he/she wants to interact with the model. The various ways are discussed above. One such example can be that the user gives the name of the spectral index and the system automatically calculates it without the user giving any other inputs. But there can be exceptions to this, as seen in MSAVI. One of the constant value needs to be given by the user based on the platform/sensor he chooses.
- 3.3. What information can the semantic model provide to its user? The semantic model provides information on the dependencies of a spectral index. Like the different requirements needed to calculate the index. These requirements can have all or one of the listed dependencies - wavelength, function, constant value, or another index. Not only the dependencies but also the thematic domains which includes the tasks as well as themes for the spectral indices. Both the dependencies as well as the thematic domain are explained in detail in Section 3.2.

4.4 LIMITATIONS

The following are the few limitations for the semantic model developed.

- For illustration purpose, this model uses only few examples of spectral indices, tasks, themes, and raw EO products. In order to provide a sufficient overview, more indices, themes, and raw EO products have to be added.
- The model only exploits a limited number of characteristics of the raw EO data but other instructions can be add through the OGC: 17-003r3. This provides more instructions like the coverage description, or product information, etc as seen in Figure 2.9. The coverage description includes the cloud cover and snow cover. According to OGC: 17-003r3, it describes the percentage of cloud cover and snow cover on the satellite image. This would help users to choose a product which free from cloud/ snow cover or have low percentage of them. The product information gives a description of the EO product when it became available, its mission, and version, etc. This can help the users to know detailed information about the product.
- The model displayed shows that the processing instructions are dependent on formula but it can also be described beyond just a single formula. This model is limited to simple formula products In this case, it describes the formula for spectral indices but it cannot describe complex EO products. For example, elevation models cannot be described by the semantic model. This is because the calculation of elevation models depends on iterative process and this semantic model cannot describe iterative process.

Chapter 5

Conclusion and Outlook

The concept of ontology is not new to the computer science domain but it has recently gained popularity in other domains like petroleum, mechanics, or EO. In any domain, it sure has been a boon in imparting knowledge that is both readable by the machines and humans. Ontologies can be developed for any purpose which needs communication between machines and humans. One such purpose has been displayed here presently. The semantic model proposed here is used to describe the semantics for the spectral index required by the user. This model contributes to the vision of automatically calculating EO products as per the user's requirements. This requirement of users has been termed as on-demand geoprocessing of the EO product here. The user interaction with the semantic model has been outline in form of CQs. Users do not need to know about the back end calculation of the spectral indices but they need to give inputs wherever needed like the name of the spectral index, or the temporal/ spatial resolution of the required dataset. The user can also input the sensors or platforms required for the spectral indices. Currently, the semantic model does not automatically calculate the spectral indices but it does give a description on how to do so. It displays the formula for an index and its requirements which can be any combination of functions, another index, constant values, and wavelength. The corresponding raw EO datasets are described using the OGC:17-003r2 ontology. To validate the model, SPARQL queries were created based on the posed CQs and successfully executed on an example set of instances. All the CQs identified could be answered entirely which conveys that the model is practically useful for on-demand description of spectral indices.

A subset of the model was displayed in the EO4GEO Bok. This included only the index, task, and the theme of the semantic model. This representation was used to conduct a survey to verify the assignment of tasks and themes to indices. The survey resulted in a majority of the participants giving a positive feedback. Even most of suggestions provided can be incorporated in the semantic model which can be seen as a future work.

Overall, this semantic model is a piece of contribution to the science by saving energy and other resources used while producing EO to the users. The code/queries used can be re-used by different available platforms like ESA (European Space Agency) thematic exploitation, CODE-DE, etc. Therefore, as the codes developed for EO product in this thesis are working, it can be shared via these platforms.

This research is also significant in producing EO products that do not require processing data on another platforms/ softwares like ArcGIS, or ERDAS, etc. This would help the user to focus on his/ her project or research rather than spending time in cleaning the data to get a desired output.

For this research only few instances are taken but as this model can describe any number of spectral indices, therefore, more instances should be added in the future. The code developed here can be used for back end programming of an interface which works on-demand geoprocessing of the spectral indices. Moreover, the list of spectral indices from IDB website can be used for this interface. Thus, the semantic model developed here sets the foundation for iterative on-demand computation of the spectral indices.

Bibliography

- Allemang, D. & Hendler, J. (2011). Semantic web application architecture. *Semantic web for the working ontologist* (pp. 51–60).
<https://doi.org/10.1016/b978-0-12-385965-5.10004-4>
- Angele, J., Decker, S., Erdmann, M., Schnurr, H.-P., Staab, S., Studer, R., Witt, A. & Fensel, D. (1999). On2broker: Semantic-based access to information sources at the www. *WebNet World Conference on the WWW and Internet*, 366–371.
- Arafat, S., Aboelghar, M. & Farg, E. (2013). Crop discrimination using field hyper spectral remotely sensed data. *Advances in Remote Sensing*, 02, 63–70. <https://doi.org/10.4236/ars.2013.22009>
- Bannari, A., Morin, D., Bonn, F. & Huete, A. R. (1995). A review of vegetation indices. *Remote Sensing Reviews*, 13(1-2), 95–120. <https://doi.org/10.1080/02757259509532298>
- Baret, F., Guyot, G. & Major, D. (1989). Tsavi: A vegetation index which minimizes soil brightness effects on lai and apar estimation. *12th Canadian Symposium on Remote Sensing Geoscience and Remote Sensing Symposium*, 3, 1355–1358.
<https://doi.org/10.1109/IGARSS.1989.576128>
- Berners-Lee, T. (1999). *Weaving the web: The original design and ultimate destiny of the world wide web by its inventor*. Harper San Francisco.
- Berners-Lee, T., Hendler, J. & Lassila, O. (2001). A new form of web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific american*, 284(5), 34–43.
- Bizer, C., Heath, T. & Berners-Lee, T. (2011). Linked data: The story so far. *Semantic services, interoperability and web applications: Emerging concepts* (pp. 205–227). IGI global.
- Boori, M., Choudhary, K., Paringer, R., Sharma, A., Kupriyanov, A. & Corgne, S. (2019). Monitoring Crop Phenology Using NDVI Time Series from Sentinel 2 Satellite Data. *2019 5th International Conference on Frontiers of Signal Processing (ICFSP)*, 62–66.
<https://doi.org/10.1109/icfsp48124.2019.8938078>
- Brickley, D. & Guha, R. (2014). *RDF Schema 1.1. Retrieved from W3C Recommendation*.
<https://www.w3.org/TR/rdf-schema/>
- Cabrera-Bosquet, L., Molero, G., Stellacci, A., Bort, J., Nogüç½s, S. & Araus, J. (2011). Ndvi as a potential tool for predicting biomass, plant nitrogen content and growth in wheat genotypes subjected to different water and nitrogen conditions. *Cereal Research Communications*, 39, 147–159.
<https://doi.org/10.1556/CRC.39.2011.1.15>
- Coene, Y., Voges, U. & Barois, O. (2020). *Ogc eo dataset metadata geojson(-ld) encoding standard*.
<https://docs.openeospatial.org/is/17-003r2/17-003r2.html>
- Cogato, A., Meggio, F., Collins, C. & Marinello, F. (2020). Medium-resolution multispectral data from sentinel-2 to assess the damage and the recovery time of late frost on vineyards. *Remote Sensing*, 12(11), 1896.

- Compton, M., Barnaghi, P., Bermudez, L., Garcí½a-Castro, R., Corcho, O., Cox, S., Graybeal, J., Hauswirth, M., Henson, C., Herzog, A., Huang, V., Janowicz, K., Kelsey, W. D., Le Phuoc, D., Lefort, L., Leggieri, M., Neuhaus, H., Nikolov, A., Page, K., ... Taylor, K. (2012).
The SSN ontology of the W3C semantic sensor network incubator group.
Journal of Web Semantics, 17, 25–32. <https://doi.org/10.1016/j.websem.2012.05.003>
- DiGiuseppe, N., Pouchard, L. C. & Noy, N. F. (2014).
Sweet ontology coverage for earth system sciences. *Earth Science Informatics*, 7(4), 249–264.
- Dlamini, S. N., Beloconi, A., Mabaso, S., Vounatsou, P., Impouma, B. & Fall, I. S. (2019).
Review of remotely sensed data products for disease mapping and epidemiology.
Remote Sensing Applications: Society and Environment, 14, 108–118.
<https://doi.org/10.1016/j.rsase.2019.02.005>
- Doldrina, C. (2015). Open data and earth observations: The case of opening up access to and use of earth observation data through the global earth observation system of systems.
J. Intell. Prop. Info. Tech. & Elec. Com. L., 6, 73.
- ESA. (2019). *Thematic exploitation platforms overview*.
<https://eo4society.esa.int/thematic-exploitation-platforms-overview/>
- Filgueiras, R., Mantovani, E. C., Althoff, D., Fernandes Filho, E. I. & Cunha, F. F. d. (2019).
Crop ndvi monitoring based on sentinel 1. *Remote Sensing*, 11(12).
<https://doi.org/10.3390/rs11121441>
- Fornacca, D., Ren, G. & Xiao, W. (2018). Evaluating the best spectral indices for the detection of burn scars at several post-fire dates in a mountainous region of northwest yunnan, china.
Remote Sensing, 10(8), 1196.
- Gabban, A., San-Miguel-Ayanz, J. & Viegas, D. X. (2006). On the suitability of the use of normalized difference vegetation index for forest fire risk assessment.
International Journal of Remote Sensing, 27(22), 5095–5102.
<https://doi.org/10.1080/01431160500185656>
- Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing?
International journal of human-computer studies, 43(5-6), 907–928.
- Henrich, V., Krauss, G., Götze, C. & Sandow, C. (2012). *IDB Index Database*.
<https://www.indexdatabase.de>
- Hüsemann, B. & Vossen, G. (2005). Ontology engineering from a database perspective.
Annual Asian Computing Science Conference, 49–63.
- Hyndavi, A., James, L., Anjaneyulu, R. V. G., Suresh, S., Venkateswara Rao, C. & Bothale, V. M. (2019). Evolution of value addition process for generation of normalised difference vegetation index (ndvi) product - a case study. *2019 IEEE Recent Advances in Geoscience and Remote Sensing : Technologies, Standards and Applications (TENGARSS)*, 86–91.
<https://doi.org/10.1109/TENGARSS48957.2019.8976035>
- Keil, J. M. & Schindler, S. (2019).
Comparison and evaluation of ontologies for units of measurement.
Semantic Web, 10(1), 33–51. <https://doi.org/10.3233/SW-180310>
- Kim, M. S., Daughtry, C., Chappelle, E., McMurtry, J. & Walthall, C. (1994).
The use of high spectral resolution bands for estimating absorbed photosynthetically active radiation (a par).
- Lassila, O. & Swick, R. (1999). *Resource Description Framework (RDF) Model and Syntax Specification. W3C Recommendation*.
<https://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>
-

- Madin, J., Bowers, S., Schildhauer, M., Krivov, S., Pennington, D. & Villa, F. (2007). An ontology for describing and synthesizing ecological observation data. *Ecological Informatics*, 2(3), 279–296. <https://doi.org/10.1016/j.ecoinf.2007.05.004>
- Maedche, A. & Staab, S. (2002). Applying semantic web technologies for tourism information systems. *Information and Communication Technologies in Tourism*, 311–319.
- Mandl, D., Sohlberg, R., Justice, C., Ungar, S., Ames, T., Frye, S., Chien, S., Tran, D., Cappelaere, P. & Derezinski, L. (2007). Sensor webs with a service-oriented architecture for on-demand science products. *Atmospheric and Environmental Remote Sensing Data Processing and Utilization III: Readiness for GEOSS*, 6684, 668414.
- Migdall, S., Brüggemann, L. & Bach, H. (2018). Earth observation in agriculture. *Satellite-based earth observation* (pp. 85–93). Springer.
- National Research Council. (2013). *Landsat and beyond: Sustaining and enhancing the nation's land imaging program*. The National Academies Press. <https://doi.org/10.17226/18420>
- Ontotext. (2016). *What is a knowledge graph*. <https://www.ontotext.com/knowledgehub/fundamentals/what-is-a-knowledge-graph/>
- Pareja-Lora, A. (2012). *Ontotag - a linguistic and ontological annotation model suitable for the semantic web* (Doctoral dissertation). Universidad Politécnica de Madrid.
- Patel-Schneider, P. F. & Horrocks, I. (2004). *OWL Web Ontology Language Semantics and Abstract Syntax Section 2. Abstract Syntax*. Retrieved from *W3C Recommendation*. <https://www.w3.org/TR/owl-semantics/syntax.html>
- Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H. & Sorooshian, S. (1994). A modified soil adjusted vegetation index. 48, 119–126. [https://doi.org/10.1016/0034-4257\(94\)90134-1](https://doi.org/10.1016/0034-4257(94)90134-1)
- Richardson, A. J. & Everitt, J. H. (1992). Using spectral vegetation indices to estimate rangeland productivity. *Geocarto International*, 7(1), 63–69. <https://doi.org/10.1080/10106049209354353>
- Rose, L. (2004). Geospatial portal reference architecture: A community guide to implementing standards-based geospatial portals. *OpenGIS Discussion Paper, OGC*, 04–039.
- Rouse, J., Haas, R. H., Schell, J. A., Deering, D. W. et al. (1974). Monitoring vegetation systems in the great plains with erts. *NASA special publication*, 351(1974), 309.
- Sadgir, P. & Gatkal, S. (2017). NDVI: Irrigation Performance Evaluation Using RS and GIS. *Journal of Water Resource Engineering and Management*, 4, 17–21.
- Shafique, N. A., Fulk, F., Autrey, B. C., Flotemersch, J. et al. (2003). Hyperspectral remote sensing of water quality parameters for large rivers in the ohio river basin. *First interagency conference on research in the watershed, Benson, AZ*, 216–221.
- Sudmanns, M., Tiede, D., Lang, S., Bergstedt, H., Trost, G., Augustin, H., Baraldi, A. & Blaschke, T. (2020). Big earth data: Disruptive changes in earth observation data management and analysis? *International Journal of Digital Earth*, 13(7), 832–850. <https://doi.org/10.1080/17538947.2019.1585976>
- Ukrainski, P. (2019). *IDB: a remote sensing indices database*. <http://www.50northspatial.org/idb-remote-sensing-indices-database/>
- Union of Concerned Scientists. (2021). *UCS satellite database*. <https://www.ucsusa.org/resources/satellite-database>

- W3C. (2013). *SPARQL 1.1 Overview*. Retrieved from W3C Recommendation.
<https://www.w3.org/TR/2013/REC-sparql11-overview-20130321/#sparql11-protocol>
- W3C. (2015a). *Semantic web*. <https://www.w3.org/standards/semanticweb/>
- W3C. (2015b). *VOCABULARIES*. <https://www.w3.org/standards/semanticweb/ontology>
- Wang, Chen, Chen, Yang & Li. (2016). Earth observation metadata ontology model for spatiotemporal-spectral semantic-enhanced satellite observation discovery: A case study of soil moisture monitoring. *GIScience & Remote Sensing*, 53(1), 22–44.
<https://doi.org/10.1080/15481603.2015.1092490>
- Wang, Wang, W. & Chen, N. (2017). Building an ontology for hydrologic monitoring. *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 6232–6234. <https://doi.org/10.1109/IGARSS.2017.8128433>
- Wenzel, K. & Reinhardt, H. (2012).
Mathematical computations for linked data applications with openmath.
Proceedings of the 24th Workshop on OpenMath, 38–48.
- World Bank. (2020). *Introduction to Google Earth Engine (GEE)*. https://worldbank.github.io/OpenNightLights/tutorials/mod2_4_introduction_to_GEE.html
- Xue, J. & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of sensors*, 2017.
- Yengoh, G., Dent, D., Olsson, L., Tengberg, A. & Tucker, C. (2015).
Applications of ndvi for land degradation assessment.
https://doi.org/10.1007/978-3-319-24112-8_3

Appendices

Appendix A

Semantic model Survey

We have developed a semantic model to describe the various components of the spectral indices, their dependencies, and uses. The vision is a system where users select from various parameters for their work and the system can automatically calculate the required spectral index. The model itself is a description which uses various raw Earth Observation characteristics like sensors, the wavelength, or platforms. There is one component of the system where a user can make use of the application domain and related tasks to identify a suitable index.

Three different use cases have been built to verify how useful this semantic model will be while searching among the indices. First is from a consumer's perspective, second from a producer's perspective, and third a non-expert user's view.

Please choose your expertise area and give your feedback. Thank you for rendering your time!

YOUR DETAILS

1. Your expertise is in:

- Water Resources
- Forestry
- Urban Planning
- Geology
- Computer Science
- Geoinformatics

If other, please specify: _____

USE CASE: GENERAL*Question 1*

One goal of the semantic model is to identify an index related to a thematic domain or application domain. This part is displayed in EO4GEO BoK available on Living Textbook. All the concepts of semantic model starts with [SM: Indices]. The link to EO4GEO - <https://ltb.itc.utwente.nl/page/508/dashboard>. On the upper right corner of the page, there is an "Open Map" button. Please click this button while answering any question containing links. The following figure outlines the general structure. Note, that each connection can appear multiple times. So e.g., an index can have an arbitrary number of associated tasks:

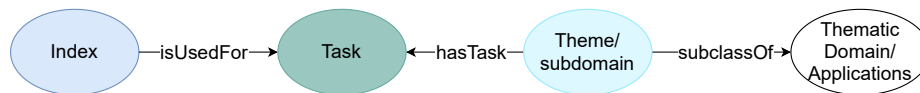


Figure A.1: As displayed in the survey

- Index: Refers to spectral indices.
- Task: Refers to the application an index is useful for.
- Theme/subdomain: Refers to the broader picture of task.
- Thematic Domain/Applications: Refers to the part where the semantic model is linked to the EO4GEO BoK.

Limitation: For the purpose of this evaluation, only four spectral indices were added.

1. Do you think the above mentioned three layers (coloured) of the semantic model are enough to describe possible uses of spectral indices?

- Yes
 No

If No, Please specify why: _____

2. Currently, tasks are modelled as a kind of EO services and applications. An example of this is given in the link provided below. Do you agree with this view? Please use the link below.

- Yes
 No

If No, Please specify why: _____

Please refer to this link to answer Q2 - <https://ltb.itc.utwente.nl/page/508/concept/98930>

3. As the survey participant would you be able to add new tasks and subdomains, based on your expertise.

- Yes
 No

If No, Please specify why: _____

USE CASE: CONSUMER

A consumer is the person who makes use of already existing data. In the context of the semantic model, a consumer is dealing with the a database of spectral indices. The consumer of spectral indices can be GIS analysts, Urban planners, Geologists, etc. using the indices for their particular tasks. They can also be sub categorized on their skills as expert or non-expert consumers of the spectral indices.

Question 2

Example Scenario: A GIS analyst is working for a company. This company got a project about crop monitoring and health. He is a part of the project and he needs to analyze the various factors related to crop health. He chooses spectral indices for his work. So, he is searching for various spectral indices that can solve his problem.

1. Crop monitoring is placed under the Agriculture theme. Do you think that this connection is right? Please use the link below.
 - Yes
 - No

If No, Please specify why: _____

Please refer to this link while answering Q1 (use open map button) -
<https://ltb.itc.utwente.nl/page/508/concept/98938>

2. Assuming that he choose one index, for example MSAVI for crop monitoring, does the description for MSAVI provide enough information to calculate it? Please use the link below.
 - Yes
 - No

If No, Please specify why: _____

Please refer to this link while answering Q2 (use open map button) -
<https://ltb.itc.utwente.nl/page/508/concept/98925>

USE CASE: PROVIDER

Question 3

A producer is the person who develops and provides the new information. In the context of the semantic model. The providers of the spectral indices can be scientists, or remote sensing developers who are expert in their domain area like forestry, agriculture, etc. They experiment with the combination of spectral bands and come up with new spectral indices.

Example Scenario: A remote sensing researcher developed a spectral index related to soil texture. He wants that his index should be easily available to the user. He is searching through the semantic model in EO4GEOBoK to best fit his index.

1. Are the description for tasks and themes sufficient enough to describe the new spectral index? In this case the new index is related to soil texture. Please use the link below.
 - Yes
 - No

If No, Please specify why: _____

Please refer to the this link to answer Q1 (use Open Map button) -
<https://ltb.itc.utwente.nl/page/508/concept/98939>

2. As a developer of the spectral index, would it be useful to specify tasks or application areas where the index can not be used for?
 - Yes
 - No

Please specify why: _____

USE CASE: NON EXPERT USER

Question 4

A bachelors student is searching for different spectral indices for Agriculture domain. He/she has the objective to list all possible uses for the given indices and summarize which index is best suited for a particular task. He/she is searching through the spectral indices of semantic model displayed in EO4GEO BoK.

1. Are the spectral indices related to Agriculture domain correct?

- Yes
- No

If No, Please specify why: _____

2. Are the tasks related to a particular index correct? These are the tasks under Agriculture like - NDVI: Monitor Crops, NDVI: Assess Land Degradation, NDVI: Measure Biomass, TSAVI: Measure Biomass, and MSAVI: Monitor Crops.

- Yes
- No

If No, Please specify alternatives/reasons: _____

FEEDBACK

Do you have any feedback while exploring the semantic model displayed in EO4GEO BoK?

Thank you again for rendering your time.

Appendix B

T-Box for semantic model

```

@prefix : <http://w3id.org/eoi/> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@base :> .

<http://w3id.org/eoi/>
  rdf:type owl:Ontology ;
  <http://purl.org/dc/terms/description>
    "Semantic model for on-demand Earth Observation Products."@en ;
  <http://purl.org/dc/terms/publisher> "Anam Akhtar"@en ;
  <http://purl.org/dc/terms/title> "Spectral Indices Model"@en ;
  owl:versionInfo "1.0"@en .

#####
# Annotation properties
#####

### http://w3id.org/eoi/hasReference
:hasReference rdf:type owl:AnnotationProperty .

### http://purl.org/dc/terms/description
<http://purl.org/dc/terms/description> rdf:type owl:AnnotationProperty .

### http://purl.org/dc/terms/issued
<http://purl.org/dc/terms/issued> rdf:type owl:AnnotationProperty .

### http://purl.org/dc/terms/modified
<http://purl.org/dc/terms/modified> rdf:type owl:AnnotationProperty .

### http://purl.org/dc/terms/publisher
<http://purl.org/dc/terms/publisher> rdf:type owl:AnnotationProperty .

### http://purl.org/dc/terms/title
<http://purl.org/dc/terms/title> rdf:type owl:AnnotationProperty .

### http://www.w3.org/2004/02/skos/core#note
<http://www.w3.org/2004/02/skos/core#note> rdf:type owl:AnnotationProperty .

#####
# Datatypes
#####

### http://www.w3.org/2001/XMLSchema#date
xsd:date rdf:type rdfs:Datatype .

#####
# Object Properties
#####

### http://w3id.org/eoi/dependsOn
:dependsOn rdf:type owl:ObjectProperty ;

```

```

    rdfs:domain :Constant_Value ;
    rdfs:range :Remote_Sensing_Device .

### http://w3id.org/eoi/derivedFrom
:derivedFrom rdf:type owl:ObjectProperty ;
    rdfs:domain :Index ;
    rdfs:range :Index .

### http://w3id.org/eoi/hasFormula
:hasFormula rdf:type owl:ObjectProperty ;
    rdfs:domain :Required_Function ;
    rdfs:range :Formula .

### http://w3id.org/eoi/hasTask
:hasTask rdf:type owl:ObjectProperty ;
    rdfs:domain :Theme ;
    rdfs:range :Task .

### http://w3id.org/eoi/hasWavelengthUnit
:hasWavelengthUnit rdf:type owl:ObjectProperty ;
    rdfs:domain :Required_Wavelength ;
    rdfs:range :Unit .

### http://w3id.org/eoi/identifies
:identifies rdf:type owl:ObjectProperty ;
    rdfs:domain :Required_Constant ;
    rdfs:range :Constant_Value .

### http://w3id.org/eoi/isOnlyApplicableFor
:isOnlyApplicableFor rdf:type owl:ObjectProperty ;
    rdfs:domain :Index ;
    rdfs:range :Remote_Sensing_Device .

### http://w3id.org/eoi/isRelatedTo
:isRelatedTo rdf:type owl:ObjectProperty ;
    rdfs:domain :Required_Index ;
    rdfs:range :Index .

### http://w3id.org/eoi/isUsedFor
:isUsedFor rdf:type owl:ObjectProperty ;
    rdfs:domain :Index ;
    rdfs:range :Task .

### http://w3id.org/eoi/requires
:requires> rdf:type owl:ObjectProperty ;
    rdfs:domain :Formula ;
    rdfs:range :Requirement .

### http://w3id.org/eoi/requiresFormula
:requiresFormula rdf:type owl:ObjectProperty ;
    rdfs:domain :Index ;
    rdfs:range :Formula .

#####
# Data properties
#####

### http://w3id.org/eoi/hasDescription
:hasDescription rdf:type owl:DatatypeProperty ;
    rdfs:domain :Requirement ;
    rdfs:range xsd:string .

### http://w3id.org/eoi/hasDiscreteWavelengthValue
:hasDiscreteWavelengthValue rdf:type owl:DatatypeProperty ;
    rdfs:domain :Required_Wavelength ;

```

```
        rdfs:range xsd:decimal .

### http://w3id.org/eoi/hasEndWavelengthValue
:hasEndWavelengthValue rdf:type owl:DatatypeProperty ;
                        rdfs:domain :Required_Wavelength ;
                        rdfs:range xsd:decimal .

### http://w3id.org/eoi/hasStartWavelengthValue
:hasStartWavelengthValue rdf:type owl:DatatypeProperty ;
                        rdfs:domain :Required_Wavelength ;
                        rdfs:range xsd:decimal .

### http://w3id.org/eoi/hasValue
:hasValue rdf:type owl:DatatypeProperty ;
          rdfs:domain :Constant_Value ;
          rdfs:range xsd:decimal .

### http://w3id.org/eoi/isRepresentedBy
:isRepresentedBy rdf:type owl:DatatypeProperty ;
                rdfs:domain :Formula ;
                rdfs:range xsd:string .

#####
#   Classes
#####

### http://w3id.org/eoi/Constant_Value
:Constant_Value rdf:type owl:Class .

### http://w3id.org/eoi/Formula
:Formula rdf:type owl:Class .

### http://w3id.org/eoi/Index
:Index rdf:type owl:Class .

### http://w3id.org/eoi/Remote_Sensing_Device
:Remote_Sensing_Device rdf:type owl:Class .

### http://w3id.org/eoi/Required_Constant
:Required_Constant rdf:type owl:Class ;
                  rdfs:subClassOf :Requirement .

### http://w3id.org/eoi/Required_Function
:Required_Function rdf:type owl:Class ;
                  rdfs:subClassOf :Requirement .

### http://w3id.org/eoi/Required_Index
:Required_Index rdf:type owl:Class ;
                rdfs:subClassOf :Requirement .

### http://w3id.org/eoi/Required_Wavelength
:Required_Wavelength rdf:type owl:Class ;
                    rdfs:subClassOf :Requirement .

### http://w3id.org/eoi/Requirement
:Requirement rdf:type owl:Class .

### http://w3id.org/eoi/Task
:Task rdf:type owl:Class .

### http://w3id.org/eoi/Thematic_Application
:Thematic_Application rdf:type owl:Class .

### http://w3id.org/eoi/Theme
:Theme rdf:type owl:Class ;
       rdfs:subClassOf :Thematic_Application .

### http://w3id.org/eoi/Unit
```



```

:Unit rdf:type owl:Class .

### http://www.opengis.net/ont/eo-geojson/1.0/Instrument
<http://www.opengis.net/ont/eo-geojson/1.0/Instrument>
  rdf:type owl:Class ;
  rdfs:subClassOf :Remote_Sensing_Device ;
  owl:disjointWith <http://www.opengis.net/ont/eo-geojson/1.0/Platform> .

### http://www.opengis.net/ont/eo-geojson/1.0/Platform
<http://www.opengis.net/ont/eo-geojson/1.0/Platform>
  rdf:type owl:Class ;
  rdfs:subClassOf :Remote_Sensing_Device .

#####
#   General axioms
#####

[ rdf:type owl:AllDisjointClasses ;
  owl:members ( :Formula
                  :Index
                  :Requirement
                )
] .

[ rdf:type owl:AllDisjointClasses ;
  owl:members ( :Required_Constant
                  :Required_Function
                  :Required_Index
                  :Required_Wavelength
                )
] .

### Generated by the OWL API (version 4.5.9.2019-02-01T07:24:44Z)
### https://github.com/owlcs/owlapi

```

Appendix C

A-Box for semantic model

The below contains the examples for only one sample of A-Box for Index, Formula Requirement, Required Function, Required Index, Required Constant, and Required Wavelength with task and theme of the semantic model.

```

@prefix : <http://w3id.org/eoi/>.
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@base <http://w3id.org/eoi/> .

<http://w3id.org/eoi/>
  rdf:type owl:Ontology ;
  <http://purl.org/dc/terms/description>
    "Semantic model for on-demand Earth Observation Products."@en ;
  <http://purl.org/dc/terms/publisher> "Anam Akhtar"@en ;
  <http://purl.org/dc/terms/title> "Spectral Indices Model"@en ;
  owl:versionInfo "1.0"@en .

#A-Box for Index
### http://w3id.org/eoi/NDVI
<http://w3id.org/eoi/NDVI> rdf:type owl:NamedIndividual ,
  <http://w3id.org/eoi/Index> ;
<http://w3id.org/eoi/hasFormula> <http://w3id.org/eoi/NDVI_Formula> ;
<http://w3id.org/eoi/isUsedFor> <http://w3id.org/eoi/Assess_Land_Degradation> ,
  <http://w3id.org/eoi/Detect_Fire_Zone> ,
  <http://w3id.org/eoi/Detect_Vegetation_Growth> ,
  <http://w3id.org/eoi/Measure_Biomass> ,
  <http://w3id.org/eoi/Monitor_Crop> ,
  <http://w3id.org/eoi/Quantify_Forest_Supply> ;
<http://w3id.org/eoi/hasReference> "Rouse, J.W., R.H. Haas, J.A. Schell, and
D.W. Deering, 1974. Monitoring vegetation systems in the Great Plains with ERTS"@en ;}
  rdfs:label "NDVI"@en .

#A-Box for Formula Requirement
### http://w3id.org/eoi/NDVI_Formula
<http://w3id.org/eoi/NDVI_Formula> rdf:type owl:NamedIndividual ,
  <http://w3id.org/eoi/Formula> ;
<http://w3id.org/eoi/requires> <http://w3id.org/eoi/NDVI_NIR> ,
  <http://w3id.org/eoi/NDVI_RED> ;
<http://w3id.org/eoi/isRepresentedBy> "(NIR-RED)/(NIR+RED)" ;
  rdfs:label "NDVI formula"@en .

#A-Box for Required Index
### http://purl.org/eoi/MSAVI_L_Index_Req_NDVI
:MSAVI_L_Index_Req_NDVI rdf:type owl:NamedIndividual ,
  <http://w3id.org/eoi/Required_Index> ;
<http://w3id.org/eoi/isRelatedTo> <http://w3id.org/eoi/NDVI> .

#A-Box for Required Function
### http://w3id.org/eoi/MSAVI_L_Function
<http://w3id.org/eoi/MSAVI_L_Function> rdf:type owl:NamedIndividual ,
  <http://w3id.org/eoi/Required_Function> ;

```

```
<http://w3id.org/eoi/requiresFormula> <http://w3id.org/eoi/MSAVI_Formula_L> ;
<http://w3id.org/eoi/hasDescription> "An inductive MSAVI_L function applied
to maximize reduction of soil effects on the vegetation signal."^^xsd:string ;
    rdfs:label "L"@en .
```

```
#A-Box for Required Constant
### http://w3id.org/eoi/MSAVI_L_s
<http://w3id.org/eoi/MSAVI_L_s> rdf:type owl:NamedIndividual ,
    <http://w3id.org/eoi/Required_Constant> ;
<http://w3id.org/eoi/hasDescription> "s is the slope of the soil line from a
plot of red versus near infrared brightness values. The user has to input
this value."^^xsd:string ;
    rdfs:label "s"@en .
```

```
#A-Box for Required Wavelength
### http://w3id.org/eoi/CARI_GREEN
<http://w3id.org/eoi/CARI_GREEN> rdf:type owl:NamedIndividual ,
    <http://w3id.org/eoi/Required_Wavelength> ;
<http://w3id.org/eoi/hasWavelengthUnit>
    <http://www.ontology-of-units-of-measure.org/resource/om-2/micrometre> ;
    <http://w3id.org/eoi/hasDiscreteWavelengthValue> 0.56 ;
    <http://purl.org/dc/terms/description> "GREEN is the
spectral range of Sentinel 2 for CARI."@en ;
    rdfs:label "GREEN"@en .
```

```
#A-Box for Theme
### http://w3id.org/eoi/Forestry
<http://w3id.org/eoi/Forestry> rdf:type owl:NamedIndividual ,
    <http://w3id.org/eoi/Theme> ;
<http://w3id.org/eoi/hasTask> <http://w3id.org/eoi/Quantify_Forest_Supply> ;
<http://purl.org/dc/terms/description> "The Forestry is defined by the
semantic model as a theme as well as subdomain of Thematic and
application domains."@en ;
    rdfs:label "Forestry"@en .
```

```
#A-Box for Task
### http://w3id.org/eoi/Detect_Vegetation_Growth
<http://w3id.org/eoi/Detect_Vegetation_Growth> rdf:type owl:NamedIndividual ,
    <http://w3id.org/eoi/Task> ;
<http://purl.org/dc/terms/description> "Vegetaion growth can refer to the growth of any
kind of flora like plants, shrubs, or trees. Vegetation growth includes the detection of
various green vegetation growing. This task is related to many vegetation indices and some
of them are displayed in the semantic model as shown here. It has a theme or subdomain
- Vegetation."@en ;
    rdfs:label "Detect Vegetation Growth"@en .
```

Appendix D

CQs as Displayed in the GraphDB Interface

The screenshot displays the GraphDB interface. At the top, a SPARQL query is entered in a text area:

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 select ?theme ?task ?taskLabel
5 where {
6   ?theme a eoi:Theme .
7   ?theme rdfs:label "Agriculture"@en.
8   ?theme eoi:hasTask ?task.
9   ?task rdfs:label ?taskLabel .
10 }
11

```

Below the query editor, there are several controls: a 'Run' button, a 'keyboard shortcuts' link, and a row of tabs: 'Table', 'Raw Response', 'Pivot Table', and 'Google Chart'. A 'Download as' button is also present. A status bar indicates 'Showing results from 1 to 4 of 4. Query took 0.1s, minutes ago.'

The results are displayed in a table with the following columns: 'theme', 'task', and 'taskLabel'. The table contains four rows of data:

	theme	task	taskLabel
1	:Agriculture	:Assess_Land_Degradation	"Assess Land Degradation"@en
2	:Agriculture	:Estimate_Frost_Damage	"Estimate Frost Damage"@en
3	:Agriculture	:Measure_Biomass	"Measure Biomass"@en
4	:Agriculture	:Monitor_Crop	"Monitor Crop"@en

Figure D.1: CQ1 as displayed on GraphDB.

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 select ?task ?index ?IndexName
4 where {
5     ?task a eoi:Task.
6     ?task rdfs:label "Monitor Crop"@en.
7     ?index eoi:isUsedFor ?task.
8     ?index rdfs:label ?IndexName.
9 }
    
```

Run

Press Alt+Enter to autocomplete

Table Raw Response Pivot Table Google Chart Download as

Filter query results Showing results from 1 to 1 of 1. Query took 0.1s, moments ago.

	task	index	IndexName
1	:Monitor_Crop	:NDVI	"NDVI"@en

Figure D.2: CQ2 as displayed on GraphDB.

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 select ?index ?reference
5 where {
6     ?index a eoi:Index.
7     ?index rdfs:label "NDVI"@en.
8     ?index eoi:hasReference ?reference.
9 }
    
```

Run

Press Alt+Enter to autocomplete

Table Raw Response Pivot Table Google Chart Download as

Filter query results Showing results from 1 to 1 of 1. Query took 0.1s, moments ago.

	index	reference
1	:NDVI	"Rouse, J.W., R.H. Haas, J.A. Schell, and D.W. Deering, 1974. Monitoring vegetation systems in the Great Plains with ERTS"@en

Figure D.3: CQ3 as displayed on GraphDB.

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX eop: <http://www.opengis.net/ont/eo-geojson/1.0#>
3 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
4 PREFIX ical: <http://www.w3.org/2002/12/cal/ical#>
5 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
6 select ?dataset ?begin ?end ?DatasetName
7 where {
8     ?dataset a eop:EarthObservation;
9             rdfs:label ?DatasetName.
10    ?dataset eop:properties ?property.
11    ?property eop:acquisitionInformation ?aqcinfo.
12    ?aqcinfo eop:acquisitionParameters ?aqcpara.
13    ?aqcpara eop:temporalInformation ?temp.
14    ?temp ical:beginningDateTime ?begin.
15    ?temp ical:endingDateTime ?end.
16    FILTER(
17        (?begin > "2021-05-01T00:00:00Z"^^xsd:dateTime)
18        && ("2021-05-31T23:59:00Z"^^xsd:dateTime > ?begin)
19        && ("2021-05-01T00:00:00Z"^^xsd:dateTime < ?end)
20        && (?end < "2021-05-31T23:59:00Z"^^xsd:dateTime)
21    )
22 }
    
```

Run

Press Alt+Enter to autocomplete

Table Raw Response Pivot Table Google Chart
Download as ▾

Filter query results
Showing results from 1 to 2 of 2. Query took 0.1s, minutes ago.

	dataset	begin	end	DatasetName
1	http://w3id.org/eoi#PRODUCT/5	"2021-05-07T14:42:05.7748750Z"^^xsd:dateTime	"2021-05-07T14:42:32.5288750Z"^^xsd:dateTime	"Newyork_20210507_Landsat7"
2	http://w3id.org/eoi#PRODUCT/3	"2021-05-20T17:36:12Z"^^xsd:dateTime	"2021-05-20T17:36:20Z"^^xsd:dateTime	"Belgium_20210520_TerraSARX"

Figure D.4: CQ4 as displayed on GraphDB.

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
5 select ?index ?reqType ?requirement ?reqLabel
6         ?wvStart ?wvEnd ?wvUnit
7         ?nestedFormulaNode ?nestedFormula
8 where {
9   ?index rdfs:label "NDVI"@en.
10  ?index eoi:hasFormula ?formula.
11  ?formula eoi:requires ?requirement.
12  ?requirement rdfs:label ?reqLabel .
13  OPTIONAL {
14    ?requirement a eoi:Required_Wavelength .
15    BIND( eoi:Required_Wavelength AS ?reqType ).
16    ?requirement eoi:hasWavelengthUnit ?wvUnit .
17    OPTIONAL {
18      ?requirement eoi:hasDiscreteWavelengthValue ?wvDiscValue;
19    }
20    OPTIONAL {
21      ?requirement eoi:hasEndWavelengthValue ?wvEnd ;
22      eoi:hasStartWavelengthValue ?wvStart .
23    }
24  }
25  OPTIONAL {
26    ?requirement a eoi:Required_Function ;
27    eoi:requiresFormula ?nestedFormulaNode .
28    ?nestedFormulaNode eoi:isRepresentedBy ?nestedFormula .
29    BIND( eoi:Required_Function AS ?reqType ).
30  }
31  OPTIONAL {
32    ?requirement a eoi:Required_Constant .
33    BIND( eoi:Required_Constant AS ?reqType ).
34  }
35  OPTIONAL {
36    ?requirement a eoi:Required_Index .
37    BIND( eoi:Required_Index AS ?reqType ).
38  }
39 }

```



Run

keyboard shortcuts

- Table
- Raw Response
- Pivot Table
- Google Chart

Download as

Filter query results Showing results from 1 to 2 of 2. Query took 0.1s, minutes ago.

	index	reqType	requirement	reqLabel	wvStart	wvEnd	wvUnit	nestedFormu laNode	nestedFormu la
1	:NDVI	:Required_Wav	:NDVI_NIR	"NIR"@en	"0.75"^^xsd:decimal	"1.35"^^xsd:decimal	http://www.ontology-of-units-of-measure.org/resource/2/micrometre		
2	:NDVI	:Required_Wav	:NDVI_RED	"RED"@en	"0.6"^^xsd:decimal	"0.7"^^xsd:decimal	http://www.ontology-of-units-of-measure.org/resource/2/micrometre		

Figure D.5: CQ5 as displayed on GraphDB.

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 PREFIX eop: <http://www.opengis.net/ont/eo-geojson/1.0#>
5 select ?index ?MSAVIformula ?indexRequirement ?reqType
6 where {
7     ?index a eoi:Index;
8         rdfs:label "MSAVI"@en;
9         eoi:hasFormula ?formula.
10    ?formula eoi:isRepresentedBy ?MSAVIformula.
11    ?formula eoi:requires ?indexRequirement.
12    ?indexRequirement a ?reqType.
13    ?reqType rdfs:subClassOf eoi:Requirement.
14 }

```

Run

Press Alt+Enter to autocomplete

Table Raw Response Pivot Table Google Chart Download as

Filter query results Showing results from 1 to 3 of 3. Query took 0.1s, moments ago.

	index	MSAVIformula	indexRequirement	reqType
1	:MSAVI	"(NIR-RED)/(1+L)/(NIR+RED+L)"	:MSAVI_L_Function	:Required_Function
2	:MSAVI	"(NIR-RED)/(1+L)/(NIR+RED+L)"	:MSAVI_NIR	:Required_Wavelength
3	:MSAVI	"(NIR-RED)/(1+L)/(NIR+RED+L)"	:MSAVI_RED	:Required_Wavelength

Figure D.6: CQ6 part 1 as displayed on GraphDB.


```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 PREFIX eop: <http://www.opengis.net/ont/eo-geojson/1.0#>
5 select ?index ?nestedFormula ?indexRequirement ?requiredIndex ?ReqIndexName
6 where {
7   ?index a eoi:Index;
8     rdfs:label "MSAVI"@en;
9     eoi:hasFormula ?formula.
10  ?formula eoi:requires ?funcRequirement.
11  ?funcRequirement a eoi:Required_Function.
12  ?funcRequirement eoi:requiresFormula ?funcformula.
13  ?funcformula eoi:isRepresentedBy ?nestedFormula.
14  ?funcformula eoi:requires ?indexRequirement.
15  ?indexRequirement a eoi:Required_Index.
16  ?indexRequirement eoi:isRelatedTo ?requiredIndex.
17  ?requiredIndex rdfs:label ?ReqIndexName.
18 }

```

Press Alt+Enter to autocomplete

Table Raw Response Pivot Table Google Chart Download as

Filter query results Showing results from 1 to 2 of 2. Query took 0.1s, minutes ago.

	index	nestedFormula	indexRequirement	requiredIndex	ReqIndexName
1	:MSAVI	"1- 2*s*NDVI*WDVI"	:MSAVI_L_Index_Req_ND :NDVI		"NDVI"@en
2	:MSAVI	"1- 2*s*NDVI*WDVI"	:MSAVI_L_Index_Req_WD :WDVI		"WDVI"@en

Figure D.7: CQ6 part 2 as displayed on GraphDB.

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 PREFIX eop: <http://www.opengis.net/ont/eo-geojson/1.0#>
5 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
6 select ?index ?require ?start ?end ?Discrete ?WaveUnit where {
7   ?index a eoi:Index;
8         rdfs:label "TSAVI"@en;
9         eoi:hasFormula ?formula.
10  ?formula eoi:requires ?require.
11  ?require a eoi:Required_Wavelength;
12         eoi:hasWavelengthUnit ?WaveUnit;
13  OPTIONAL {
14    ?require eoi:hasDiscreteWavelengthValue ?Discrete;
15  }
16  OPTIONAL {
17    ?require eoi:hasEndWavelengthValue ?end ;
18            eoi:hasStartWavelengthValue ?start .
19  }
20 }
    
```

Run
keyboard shortcuts

Table Raw Response Pivot Table Google Chart Download as

Filter query results Showing results from 1 to 2 of 2. Query took 0.1s, moments ago.

	index	require	start	end	Discrete	WaveUnit
1	:TSAVI	:TSAVI_NIR	"0.75"^^xsd:decimal	"1.35"^^xsd:decimal		http://www.ontology-of-units-of-measure.org/resource/2/micrometre
2	:TSAVI	:TSAVI_RED	"0.6"^^xsd:decimal	"0.7"^^xsd:decimal		http://www.ontology-of-units-of-measure.org/resource/2/micrometre

Figure D.8: CQ7 part 1 as displayed on GraphDB.

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 PREFIX eop: <http://www.opengis.net/ont/eo-geojson/1.0#>
5 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
6 select distinct ?sensor ?platformName ?sensorName
7 where {
8
9   ?aqi a          eop:AcquisitionInformation;
10      eop:instrument ?sensor;
11      eop:platform ?plat.
12
13   ?sensor eop:instrumentShortName ?sensorName.
14   ?plat eop:platformShortName ?platformName.
15   ?aqi eop:acquisitionParameters ?aqipara.
16   ?aqipara eop:wavelengths ?Wavelength.
17
18   OPTIONAL {
19     ?Wavelength eop:discreteWavelength ?wvDiscValue ;
20   }
21
22   OPTIONAL {
23     ?Wavelength eop:endWavelength ?wvEnd ;
24     eop:startWavelength ?wvStart .
25   }
26
27   VALUES (?start ?end) {
28     ("0.00000060"^^xsd:double "0.00000070"^^xsd:double)
29     ("0.00000075"^^xsd:double "0.00000135"^^xsd:double)
30   }
31   FILTER(
32     (?start < ?wvEnd)
33     && (?end > ?wvStart)
34   )
35 }

```



Run

Press Alt+Enter keyboard shortcuts

Table Raw Response Pivot Table Google Chart

Download as

Filter query results Showing results from 1 to 4 of 4. Query took 0.1s, moments ago.

	sensor	platformName	sensorName
1	http://w3id.org/eoi#INST/OLI%2FTIRS	"Landsat8"	"OLI/TIRS"
2	http://w3id.org/eoi#INST/TM	"Landsat4-5"	"TM"
3	http://w3id.org/eoi#INST/ETM%2B	"Landsat7"	"ETM+"
4	http://w3id.org/eoi#INST/PAN	"IRS1D"	"PAN"

Figure D.9: CQ7 part 2 as displayed on GraphDB.

```

1 PREFIX eoi: <http://w3id.org/eoi/>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX owl: <http://www.w3.org/2002/07/owl#>
4 PREFIX eo: <http://example.org/voc#>
5 PREFIX eop: <http://www.opengis.net/ont/eo-geojson/1.0#>
6 select distinct ?dataset ?datasetName
7 where {
8   ?dataset a eop:EarthObservation;
9           eop:properties ?prop.
10  ?prop eop:acquisitionInformation ?aqc.
11  ?aqc eop:platform ?inst.
12  ?inst eop:platformShortName "Sentinel2B".
13  ?dataset rdfs:label ?datasetName.
14 }
    
```

Run

Press Alt+Enter keyboard shortcuts

Table Raw Response Pivot Table Google Chart Download as

Filter query results Showing results from 1 to 2 of 2. Query took 0.1s, moments ago.

	dataset	datasetName
1	http://w3id.org/eoi#PRODUCT/2	"Netherlands_20210609_Sentinel2B"
2	http://w3id.org/eoi#PRODUCT/4	"Sanfrancisco_202210613_Sentinel2B"

Figure D.10: CQ8 as displayed on GraphDB.

```

1 PREFIX geof: <http://www.opengis.net/def/function/geosparql/>
2 PREFIX eoi: <http://w3id.org/eoi/>
3 PREFIX eop: <http://www.opengis.net/ont/eo-geojson/1.0#>
4 PREFIX gsp: <http://www.opengis.net/ont/geosparql#>
5 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
6 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
7 PREFIX eoip: <http://w3id.org/eoi#PRODUCT/>
8 SELECT ?data ?dataName ?WKT1
9 WHERE {
10 VALUES ?data { eoip:2 eoip:4 }
11 ?data a eop:EarthObservation;
12 rdfs:label ?dataName;
13 gsp:hasGeometry ?geom1.
14 ?geom1 gsp:asWKT ?WKT1.
15 FILTER( geof:sfIntersects(?WKT1, ""POLYGON((
16 6.3967 53.0407, 6.7592 53.0407,
17 6.7592 52.923, 6.3967 52.923,
18 6.3967 53.0407))""^^gsp:wktLiteral)).
19 }

```

Press Alt+Enter keyboard shortcuts

Run

Table Raw Response Pivot Table Google Chart Download as

Filter query results Showing results from 1 to 1 of 1. Query took 0.1s, moments ago.

	data	dataName	WKT1
1	http://w3id.org/eoi#PRODUCT/2	"Netherlands_20210609_Sentinel2B"	"Polygon((6.0047609 53.2119778, 7.6483415 53.2419635, 7.6785441 52.2551507, 6.071595 1 52.2262109, 6.0047609 53.2119778))"^^<http://www.opengis.net/ont/geosparql#wktLiteral>

Figure D.11: CQ9 as displayed on GraphDB.