



Livrable 4.2.2-doc Document présentant le livrable 4.2.2 : V1 de l'ontologie des métadonnées des images satellite

AUTEUR(S) : H. Arenas, C. Comparot, N. Aussenac-Gilles, C. Trojahn
SOURCE (SOCIETE) : IRIT - UPS
REFERENCE : <Référence>
VERSION : 1
DATE DU DOCUMENT : 29/07/2016
NOMBRE DE PAGES : 10

Table des Matières

1	Résumé	4
2	Présentation du livrable 4.2.2 : V1 de l'ontologie des métadonnées des images satellite	5
2.1	Nature du résultat et accès au fichier	5
2.2	Présentation du contenu du vocabulaire	5
3	Perspectives pour la version V2	10

Liste des modifications

Version	Date	Description	Auteur(s)
V1	29/07/2016	Présentation rapide du vocabulaire utilisé pour décrire en tant que données sémantiques les méta-données associées aux images – correspond à la version v6 du vocabulaire	H. Arenas, C. Comparot, N. Aussenac-Gilles, C. Trojahn

1 Résumé

Ce document documente succinctement le livrable 4.2.2 : V1 des ontologies des métadonnées des images satellite. Ce livrable correspond au fichier `SP4-livrable4_2_2ontmodel_v6.owl` qui peut être ouvert avec un éditeur d'ontologie ou avec toute application capable d'ouvrir un fichier XML. Cette première version de l'ontologie est en fait un vocabulaire en OWL-DRF qui ne possède pas encore toutes les propriétés d'une ontologie, qui réutilise certains des vocabulaires identifiés dans le livrable 4.2.1 et qui est utilisé pour décrire les méta-données des images fournies dans le livrable 4.3.2.

2 Présentation du livrable 4.2.2 : V1 de l'ontologie des métadonnées des images satellite

2.1 Nature du résultat et accès au fichier

2.1.1 Nature du résultat

Dans le domaine du web sémantique, une ontologie est un vocabulaire structuré qui identifie et représente formellement les classes (appelées concepts) des entités présentes dans un domaine d'intérêt, ainsi que les propriétés de ces entités et les relations qu'elles entretiennent entre elles. Une ontologie utilise une représentation formelle sous forme d'un graphe étiqueté dont les nœuds sont les classes et dont les arcs correspondent aux relations ou propriétés. Une ontologie comporte également des axiomes ou contraintes ou règles qui permettent d'exprimer des contraintes sur la valeur de données, sur les relations possibles ou les propriétés nécessaires à définir pour une entité d'une classe donnée. Enfin, les concepts de l'ontologie doivent (normalement) être organisés en fonction de grandes catégories de connaissance : être vivants, rôles sociaux, matières, objets techniques, qualités, etc.

Une ontologie du web sémantique est définie dans un langage formel. Le W3C recommande pour cela d'utiliser le langage OWL. De plus, il est recommandé de réutiliser le plus possible des vocabulaires et des ontologies existantes pour définir des classes. En fin, il est recommandé de rendre publique l'ontologie sur un serveur web.

Le vocabulaire que nous proposons dans cette version ne comporte aucune règle et aucun axiome. De plus, les classes proposées ne sont pas définies par rapport à des catégories générales de connaissance. C'est pour cela qu'on le considère comme un vocabulaire plus que comme une ontologie. De plus, il n'est pas encore disponible en ligne sur un serveur. Cependant, il réutilise le plus possible des vocabulaires existants pour décrire les méta-données.

2.1.2 Accès aux données

Le fichier contenant ce livrable est le fichier SP4-livrable4_2_2ontmodel_v6.owl .

Il qui peut être ouvert avec un éditeur d'ontologie ou avec toute application capable d'ouvrir un fichier XML. Nous suggérons d'utiliser par exemple l'éditeur d'ontologie Protégé qui peut être téléchargé gratuitement en ligne sur le site <http://protege.stanford.edu/>.

2.2 Présentation du contenu du vocabulaire

2.2.1 Liste des vocabulaires réutilisés

Le vocabulaire de cette version réutilise plusieurs vocabulaires qui ont été retenus à l'issue de l'état de l'art fait dans le livrable 4.2.1. :

- GeoSparql, un vocabulaire pour manipuler des données géo-localisées
- Dcat, un vocabulaire pour décrire des catalogues de données
- SSN, un vocabulaire pour décrire des données de capteurs
- FOAF (Friend of a Friend) qui permet de décrire des personnes
- DCT (Dublin core Term) et DCType qui permettent de décrire des documents numériques via des méta-données documentaires (date, auteur, format, provenance, etc.)
- Time, un vocabulaire qui décrit de données temporelles et permet de raisonner avec ces données
- Schema.org, une base de connaissances générales

Plus de détails sur ces vocabulaires sont donnés dans le livrable 4.2.1 .

D'autres vocabulaires sont nécessaires pour la représentation en OWL de ce vocabulaire : RDF, RDFs, OWL, SKOS, XSD

Enfin, des vocabulaires aident à mieux structurer l'ontologie, comme le vocabulaire permettant de définir des patrons de bonne structuration ou Design patterns. C'est le vocabulaire DUL.

Chacun des vocabulaires réutilisés est mentionné dans l'entête du fichier, et les classes ou propriétés empruntées à ces vocabulaires sont préfixés par le préfixe associé. La liste en est donnée dans la figure 1.

Ontology imports		Ontology Prefixes		General class axioms	
Ontology prefixes:					
Prefix					
					http://www.semanticweb.org/linux/ontologies/2016/5/untitled-ontology-29#
DUL					http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#
dcat					http://www.w3.org/ns/dcat#
dct					http://purl.org/dc/terms/
dctype					http://purl.org/dc/dcmitype/
foaf					http://xmlns.com/foaf/0.1/
geo					http://www.opengis.net/ont/geosparql#
owl					http://www.w3.org/2002/07/owl#
rdf					http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfs					http://www.w3.org/2000/01/rdf-schema#
schema					http://schema.org/
skos					http://www.w3.org/2004/skos/core#
sparkindata					http://www.irit.fr/sparkindata#
ssn					http://purl.oclc.org/NET/ssnx/ssn#
time					http://www.w3.org/2006/time#
untitled-ontology-29					http://www.semanticweb.org/linux/ontologies/2016/5/untitled-ontology-29#
vann					http://purl.org/vocab/vann/
vcard					http://www.w3.org/2006/vcard/ns#
xsd					http://www.w3.org/2001/XMLSchema#

Figure 1 : Liste des vocabulaires réutilisés et des préfixes associés

2.2.2 Classes de l'ontologie

Les principales classes de ce vocabulaire sont listées sur la figure 2. Les méta-données plus spécifiques sont des valeurs prises par ces classes pour une image donnée. Elles sont des instances de ces classes et ne sont pas représentées sur cette figure.

The screenshot displays an ontology editor interface. On the left, a 'Class hierarchy' tree shows a tree structure starting from 'Thing'. Under 'Dataset', 'SatelliteImgMetadataRecord' is highlighted. Other classes include 'Amount', 'Catalog', 'Concept', 'ConceptScheme', 'Distribution', 'Document', 'FeatureOfInterest', 'Kind', 'MediaTypeOrExtent', 'PeriodOfTime', 'PhysicalObject', 'SensingDevice', 'Platform', 'ProcessingLevel', 'Resource', 'SatelliteProduct', 'SensorMode', 'SensorOutput', 'Situation', 'Observation', 'SpatialObject', 'Feature', 'Geometry', 'UnitOfMeasure', and 'Vocabulary'. On the right, the 'Usage' tab for 'SatelliteImgMetadataRecord' shows 7 uses: 'hasPlatform', 'hasProcessingLevel', 'hasSensorMode', 'hasSpatialRepresentation', 'isDescribedBy', and 'SatelliteImgMetadataRecord' (Class and SubClassOf Dataset). The 'Description' tab shows 'Equivalent To', 'SubClass Of' (Dataset), and 'SubClass Of (Anonymous Ancestor)'. Other tabs include 'Active Ontology', 'Entities', 'Classes', 'Object Properties', 'Data Properties', 'Annotation Properties', 'Individuals', 'OWL Viz', 'DL Query', 'OntoGraf', and 'SPARQL Query'.

Figure 2 : Extrait du vocabulaire : liste des classes présentes dans ce vocabulaire

Le choix a été fait de représenter les méta-données et les données externes amenées à devenir aussi des méta-données (comme les données météo) comme des caractéristiques géo-spatiales (classe geo :Feature dans l'ontologie geo) en utilisant une relation explicite « hasSpatialRepresentation » ce qui permet de distinguer la donnée de sa localisation.

Les différentes classes sont reliées par des propriétés et forment le graphe présenté sur la figure 3.

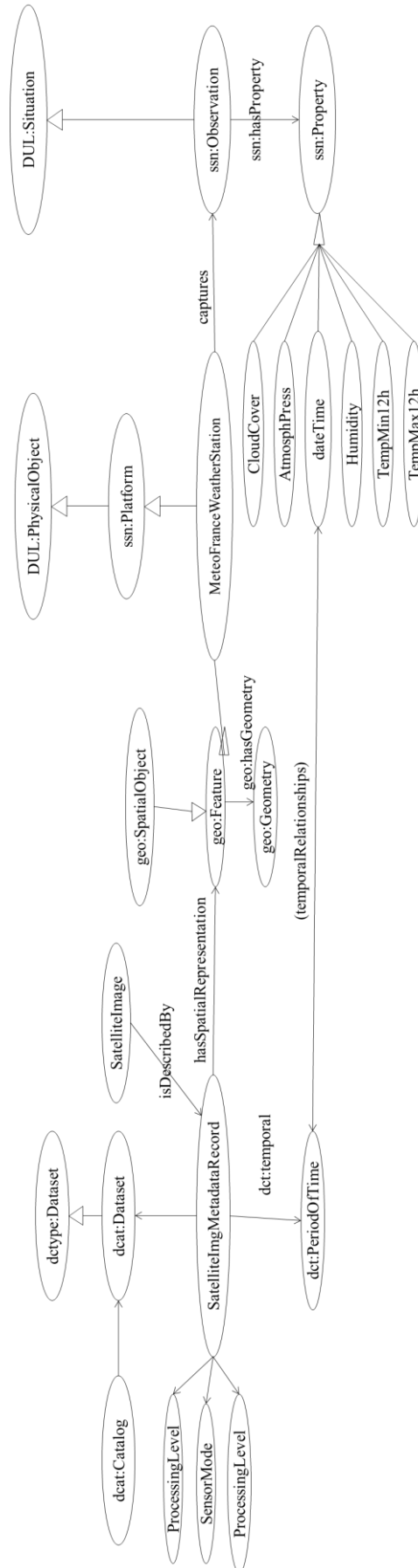


Figure 3 : graphe des classes de l'ontologie

Afin de décrire les différents types de capteur utilisé pour observer le sol, des instances de la classe « SensorMode » ont été définies et renvoient aux différentes techniques présentes sur les différentes séries de satellites Sentinel. La Figure 4 montre la définition du mode EW (Extra Wide Swath). Chaque modalité d'observation est une instance de la classe Sensor Mode.

The screenshot displays a software interface with two main panels. The left panel, titled 'Class hierarchy', shows a tree structure of classes. Under 'PhysicalObject', 'SensorMode' is listed. The right panel is split into two sections: 'Annotations: EW' and 'Description: EW'. The 'Annotations' section contains a comment in English: 'This is an operation mode for Sentinel 1.' and a detailed description of Wave Mode (WV), Interferometric Wide Swath (IW), and Extra Wide Swath (EW) modes. The 'Description' section shows 'SensorMode' as a type and lists 'Same Individual As' and 'Different Individuals' options.

Figure 4 : Exemple de type particulier de Sensor-Mode : EW

3 Perspectives pour la version V2

Plusieurs évolutions sont nécessaires pour faire de ce vocabulaire une ontologie qui permette à la fois de décrire les méta-données déjà associées aux images satellite et les nouvelles données tirées d'autres sources de données ouvertes et qui viennent enrichir les descriptions (comme des données météorologiques, des données sur la nature des sols ou encore des informations sur les cultures présentes sur des parcelles situées sur ces images). Ces évolutions sont de plusieurs ordres :

- Ajouter des classes de haut niveau, par exemple celles de l'ontologie de haut niveau DOLCE, pour mieux organiser les concepts
- Ajouter des axiomes et des contraintes sur les propriétés, en particulier pour exprimer les règles que doivent satisfaire les valeurs de certaines classes ou les relations entre entités
- Ajouter des vocabulaires pour décrire les données propres au cas d'usage étudiés. Nous avons commencé ce travail avec les données météorologiques. Un travail est en cours pour rajouter des classes afin de décrire les parcelles cultivées : la nature des sols, la nature des cultures et des intrants, ainsi que des parasites et maladies pouvant les affecter.