

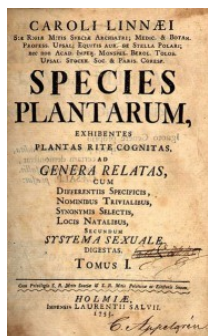
Thomas Francart

"Data artisanal" : semantic web, open data, linked data, big data...

- Accueil
- Linked Data
- Open Data
- RDF
- Sesame
- Divers
- sparna.fr

Ontologie, Thesaurus et Taxonomie sur le web de données

07/12/2013



Trois espèces d'animaux différentes se rencontrent principalement dans les steppes des langages de représentation des connaissances du web de données. Les ontologies, les thesaurus et les taxonomies, si elles partagent l'organisation hiérarchiques de concepts entre eux, n'ont pas les mêmes usages, ni les mêmes objectifs. L'ontologie est faite pour décrire le monde tel qu'il est; le thesaurus est fait pour faciliter l'accès à des contenus; la taxonomie est fait pour classer des ressources dans des dossiers, des catégories. Les systèmes d'accès aux contenus, aux données ou aux connaissances, combinent et articulent ces 3 systèmes d'organisation pour décrire le monde, indexer, et catégoriser les contenus.

	Ontologie	Thesaurus	Taxonomie
Définition	La définition couramment citée, « une ontologie est la spécification formelle d'une conceptualisation partagée », n'éclairera guère le profane. De façon plus opérationnelle, l'ontologie cherche à décrire de façon formelle un domaine de connaissance, en identifiant les types d'objets de ce domaine, leurs propriétés et leurs relations.	»Liste organisée de termes contrôlés et normalisés (descripteurs et non-descripteurs) servant à l'indexation des documents et des questions dans un système documentaire » (D. Degez, D. Ménillet, Thesauroglossaire des langages documentaires)	« Taxonomy » est très utilisé par les anglo-saxons pour désigner la « science de la classification » et par extension tout système de classification/catégorisation.
Contient...	Des classes, des propriétés, et des règles logiques formelles. Eventuellement des instances de classe.	Des concepts et des termes, organisés entre eux, avec leurs libellés, leurs traductions, leurs synonymes, et leurs descriptions/définitions.	des catégories organisés hiérarchiquement
Utilisation	Sert à instancier et à raisonner	Sert à indexer des contenus ou des ressources avec des mots-clés et à les rechercher (avec les mêmes mots-clés)	Sert à classifier , à ranger des contenus ou des ressources
Niveau de formalisme logique	Très formel (formalisme mathématique)	Peu formel	Moyennement formel (peut être plus formel qu'un thesaurus, dans le sens où la hiérarchie des entrées peut suivre certaines contraintes.)
Niveau de proximité avec la langue naturelle	Très éloigné de la langue naturelle (utilise des identifiants techniques pour s'abstraire du langage naturel)	Proche de la langue naturelle (donne des équivalents linguistiques de chaque entrée, des traductions dans d'autres langues)	Pas particulièrement en lien avec la langue naturelle.
Types de relations utilisées	<ul style="list-style-type: none"> Inclusion (classe / sous-classe); Opérations ensemblistes : union, intersection, exclusion; Caractéristiques des propriétés : domaine, ensemble d'arrivée, transitivité, propriétés inverses, etc. 	Hiérarchiques et associatives; éventuellement relations d'alignement	Hiérarchiques seulement

QUI SUIS-JE ?

Je suis consultant indépendant spécialisé sur le web de données, l'open data, l'enrichissement sémantique des moteurs de recherche, l'ingénierie documentaire et l'organisation des connaissances. Vous pouvez en apprendre plus sur moi sur le site de [Sparna](#), ma société de conseil, en particulier la page qui décrit qui je suis.

SUIVEZ CE BLOG

[RSS - Articles](#)

ABONNEZ-VOUS À CE BLOG PAR EMAIL

Recevez une notification de chaque nouvel article par email.

CATÉGORIES

ARTICLES RÉCENTS

- [Ontologie, Thesaurus et Taxonomie sur le web de données](#)
- [Vidéo parodique sur la disponibilité de DBPedia – et les services SPARQL](#)
- [Traduction française de SKOS](#)
- [Ginco – Un éditeur SKOS Open Source](#)
- [Installation du triplestore OpenLink Virtuoso sous Linux](#)

Exemple	« Voiture » est un type particulier de « Véhicule terrestre », lui-même un type particulier de « Véhicule ». La propriété « milieu de déplacement du véhicule » (terre, air, mer, espace) s'applique aux « Véhicules »; les « Véhicules terrestres » sont tous les « Véhicules » dont « milieu de déplacement » vaut « terre ».	Le terme « Voiture » est un terme spécifique de « Véhicule terrestre », lui-même spécifique de « Véhicule ». Le terme « Voiture » peut aussi se dire « Automobile », ou « Bagnole ». En anglais on dira « Car ». Si on s'intéresse à « Voiture » on peut également aller voir le terme « Autoroute ».	La catégorie « Véhicules terrestre » est une sous-catégorie des « Véhicules ». « Voiture » est rangée dans la catégorie des « Véhicules terrestre ».
Un exemple	L'ontologie FOAF définit des classes et des propriétés pour décrire des personnes.	Le thesaurus GEMET (General Multilingual Environmental Thesaurus) est un thesaurus multilingue (+ de 30 langues) développé par l'agence européenne de l'énergie et disponible gratuitement.	Les taxonomies publiées par l'IPTC servent à catégoriser les articles de news. Les catégories de Wikipedia servent à organiser les entrées de Wikipedia.
Standard de représentation sur le web de données	OWL	SKOS	SKOS
A utiliser si...	Si vous avez besoin de décrire les choses telles qu'elles sont, et pas simplement d'indexer des contenus, et que vous avez besoin de décrire précisément les caractéristiques de chaque chose, alors c'est une ontologie qu'il vous faut. (Typiquement, projets de capitalisation de connaissance)	Si vous voulez mettre des mots-clés sur des contenus, et pouvoir rechercher avec ces mêmes mots-clés, vous avez besoin d'un thesaurus. Vous pouvez vouloir utiliser un thesaurus en combinaison avec un moteur de recherche plein-texte pour améliorer sa pertinence.	Si vous avez besoin de ranger des documents dans des dossiers, de mettre au point une recherche à facettes, alors c'est d'une taxonomie dont vous avez besoin.
Mais encore...	Une ontologie n'est pas une anthologie.	Un thesaurus n'est pas un dinosaure.	Une taxonomie

Partager :

 0
  Share
  J'aime 0
  Submit
  Plus

Tags: [ontologie](#), [owl](#), [SKOS](#), [taxonomie](#), [thesaurus](#)

Poster un commentaire