



Améliorer la qualité d'un thésaurus à l'aide de requêtes SPARQL

Catherine Roussey, Stéphan Bernard

► **To cite this version:**

Catherine Roussey, Stéphan Bernard. Améliorer la qualité d'un thésaurus à l'aide de requêtes SPARQL. 9ème Atelier Recherche d'Information SEmantique,, Jul 2017, Caen, France. 11 p. hal-01597594

HAL Id: hal-01597594

<https://hal.archives-ouvertes.fr/hal-01597594>

Submitted on 28 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Améliorer la qualité d'un thésaurus à l'aide de requêtes SPARQL

Catherine Roussey¹, Stephan Bernard¹

¹ TSCF, Irstea centre de Clermont-Ferrand,
{prénom.nom}@irstea.fr

Résumé : SKOS est un schéma RDF utilisé pour stocker et publier sur le web de données liées des thésaurus ou des taxonomies. Lors du développement d'un thésaurus, il est utile de détecter automatiquement les concepts qui violent les contraintes d'intégrité. Cet article présente le cas d'usage d'un développement manuel d'un thésaurus monolingue. Ce thésaurus porte sur l'usage des cultures en France. Dans cet article, nous avons spécifié diverses contraintes d'intégrité. Pour vérifier ces contraintes un ensemble de requêtes SPARQL a été défini. Ces requêtes permettent à la fois de vérifier des contraintes d'intégrité propres à tous les thésaurus stockés au format SKOS, mais aussi des contraintes propres à notre thésaurus.

Mots-clés : SKOS, requête SPARQL, contrainte d'intégrité, développement de thésaurus, thésaurus agricole.

1 Introduction

SKOS ou Simple Knowledge Organization System (Miles & Bechhofer, 2009) est un schéma RDF utilisé pour publier sur le web de données liées des thésaurus, des taxonomies, des vocabulaires contrôlés. Ainsi les modèles SKOS peuvent être réutilisés pour indexer tout type de sources d'information. Ces modèles deviennent une ressource centrale pour les systèmes de recherche d'information dédiés. Cependant la qualité de ces modèles va impacter directement les résultats des systèmes de recherche d'information qui les utilisent. Comme le montre le Tableau 1, les spécifications SKOS définissent des contraintes d'intégrité que tout modèle SKOS doit vérifier pour être considéré comme cohérent (Miles & Bechhofer, 2009). Les spécifications SKOS définissent les contraintes minimales que doivent remplir les modèles SKOS. En fonction du domaine et de l'usage du modèle SKOS, de nouvelles contraintes peuvent être définies. Par exemple le tableau 2 présente les conventions d'usage proposées dans les spécifications SKOS. Les développeurs de thésaurus peuvent décider ou non de suivre ces conventions.

TABLE 1 – Les contraintes d'intégrité des modèles SKOS issues de (Miles & Bechhofer, 2009).

S9	la classe skos:ConceptScheme est disjointe de la classe skos:Concept.
S13	skos:prefLabel, skos:altLabel et skos:hiddenLabel sont deux à deux des propriétés disjointes.
S14	Une ressource ne peut pas avoir plus d'un skos:prefLabel par langue.
S27	la propriété skos:related est disjointe de la propriété skos:broaderTransitive.
S37	la classe skos:Collection est disjointe des classes skos:Concept et skos:ConceptScheme.
S46	la propriété skos:exactMatch est disjointe des propriétés skos:broadMatch et skos:relatedMatch.
	La propriété skos:memberList est une propriété fonctionnelle, c'est-à-dire

RISE 2017

	qu'elle ne peut pas avoir plus d'une valeur. Cette contrainte essaye d'exprimer qu'il n'y a aucun sens pour une skos:OrderedCollection d'avoir deux fois le même élément. Cette contrainte n'est pas spécifiée totalement par la propriété fonctionnelle.
--	---

TABLE 2 – Les conventions d'usage des modèles SKOS issues de (Miles & Bechhofer, 2009).

Aucune contrainte n'est spécifiée sur l'existence de top concept appartenant à un modèle. Une convention d'usage veut que les top concepts définis dans un modèle donné n'aient pas de concept parent appartenant à ce modèle.
Une convention d'usage veut que l'existence des cycles et de chemins alternatifs dans le graphe hiérarchique des concepts soit évitée.

Lors du développement d'un modèle SKOS, il est utile de détecter automatiquement les concepts SKOS qui violent les contraintes d'intégrité. Cet article présente le cas d'usage d'un développement manuel d'un thésaurus monolingue. Ce thésaurus porte sur l'usage des cultures en France. Dans ce cas d'usage, nous avons défini un ensemble de contraintes d'intégrité que doivent vérifier les concepts de ce thésaurus. Nous avons formalisé ces contraintes sous forme de requêtes SPARQL pour détecter les concepts qui violent ces contraintes.

2 Contexte

Les thésaurus agricoles sont utilisés principalement pour indexer des ressources produites ou utilisées par de grandes institutions travaillant dans le domaine de l'agriculture (INRA, Irstea, FAO). La FAO en particulier développe son thésaurus multilingue Agrovoc. Ce thésaurus, mondialement connu, est publié sur le web de données liées au format SKOS. Ce thésaurus couvre un domaine large et répond à des contraintes internationales. Les spécificités nationales n'ont pas vocation à être décrites dans ce thésaurus.

Pour les besoins d'indexation d'un corpus d'alerte agricole, nous recherchons un thésaurus capable de lister les cultures françaises mais aussi d'indiquer une parentalité d'usage dans ces cultures. À notre connaissance, il n'existait pas de ressource structurée française permettant de décrire les cultures par leurs usages ou leur destination. Les grandes classes d'usage de l'agriculture sont l'alimentation humaine ou l'alimentation animale. Certaines productions sont destinées à être transformées pour faciliter leur consommation. Par exemple, la production de houblon est destinée à la fabrication de la bière. Très peu de productions agricoles sont destinées à l'industrie sans avoir un but alimentaire. Nous pouvons citer par exemple le chanvre, qui est utilisé pour la fabrication de textile.

Notre but était de construire une hiérarchie des cultures en fonction de leur usage. Les liens hiérarchiques représentaient des relations de généralisation/spécialisation entre cultures (céréale/blé). Pour construire notre thésaurus intitulé FrenchCropUsage, nous avons étudié les termes contenus dans des documents disponibles librement. Les documents étudiés sont :

- Les statistiques agricoles annuelles publiées sur le site de l'Agreste. Le document intitulé "la statistique agricole annuelle : présentation générale" décrit la hiérarchie des cultures pour répertorier l'ensemble de la production agricole (Agreste).
- Les métadonnées du registre parcellaire graphique présentent une nomenclature des cultures ou groupes de culture (Registre Parcellaire).
- Les listes des rubriques utilisées pour organiser les bulletins d'alerte sur chacun des sites web des chambres d'agriculture (une liste contient les rubriques "Arboriculture", "Grandes cultures", ...).

- Le classement des cultures par groupe d'usage proposé par Wikipédia (Wikipédia France Culture).
- Pour compléter avec des définitions chacun des concepts de la hiérarchie, nous avons recherché les définitions dans le Larousse Agricole (Larousse Agricole).
- En cas d'absence d'information dans le Larousse Agricole, nous avons utilisé le portail français de l'agriculture de Wikipédia (Portail Agricole). Des absences de définition sont à noter surtout pour tous les fruits tropicaux.

3 Modélisation du thésaurus

L'ensemble de la hiérarchie a été modélisée à l'aide du vocabulaire SKOS proposé par le W3C (Miles & Bechhofer, 2009). Notre vocabulaire de types de cultures est disponible sur le web de données liées¹. Il contient 272 concepts. La profondeur maximale de la hiérarchie est de 6 niveaux. Comme le montre la figure 1, chaque concept est défini par les propriétés suivantes:

- `skos:prefLabel` contient le terme préféré utilisé comme étiquette du concept en français. En général, le terme est le nom usuel de la plante cultivée suivi de son usage.
- `skos:altLabel` contient les autres termes qui peuvent être utilisés comme étiquettes du concept.
- `skos:definition` contient la définition en français du concept justifiant sa position dans la hiérarchie.
- `skos:inScheme` exprime l'appartenance du concept au thésaurus.
- `rdfs:seeAlso` contient un lien web vers une définition retenue lors de la construction du thésaurus, comme par exemple les définitions du Larousse Agricole
- `skos:note` contient au moins une définition trouvée dans une autre source comme l'agreste ou Wikipédia.
- `skos:editorialNote` contient la définition du Larousse Agricole. Pour des raisons de propriété intellectuelle cette propriété est supprimée dans les versions en ligne.
- `skos:broader` indique le lien vers le concept plus générique.
- `skos:narrower` indique le lien vers le concept plus spécifique

Les grandes classes de cultures représentées dans notre thésaurus sont : culture légumière (l'alimentation humaine), culture fruitière (l'alimentation humaine), grande culture (industrie, alimentation humaine et animale), culture fourragère (l'alimentation animale), horticulture ornementale, zone non agricole. Toute culture doit être un descendant d'un de ces concepts (top concepts dans le modèle skos).

Il existe des plantes qui ont plusieurs usages : c'est le cas par exemple des céréales fourragères qui sont à la fois destinées à l'alimentation humaine (céréale, contenue dans l'épi) et à l'alimentation animale (paille, parfois utilisée comme fourrage). Quand un nom de plante est suffisamment ambigu pour qu'aucun usage ne puisse être déduit de ce nom, il faut pouvoir modéliser cette ambiguïté. Pour ce faire, nous avons défini la notion de concepts agricoles ambigus.

Concept agricole ambigu: correspond à un nom de plante cultivée pour des usages différents amenant une indécision dans le choix de l'usage de la production de la culture. Par exemple les céréales fourragères, par définition, sont destinées à l'alimentation humaine pour leur grain et à l'alimentation animale pour leur tige. Le concept représentant les céréales fourragères n'est pas ambigu. Par compte, le terme « pomme de terre » est ambigu car la

¹ FrenchCropUsage est disponible sur la plateforme AgroPortal du LIRMM et aussi sur le site <http://ontology.irstea.fr>

RISE 2017

potage est destinée à l'alimentation humaine directe et la pomme de terre cultivée en grande culture est destinée à l'industrie pour la consommation humaine indirecte ou pour la production de matériaux. Le terme « pomme de terre » ne permet pas de déterminer de manière certaine quelle est la destination de cette production.

Ces concepts ambigus ne peuvent pas être utilisés pour représenter une culture. Un de ses fils doit être choisi en remplacement. Pour la pomme de terre par exemple, on utilisera donc «pomme de terre féculière» ou «pomme de terre potagère» pour lever l'ambiguïté.

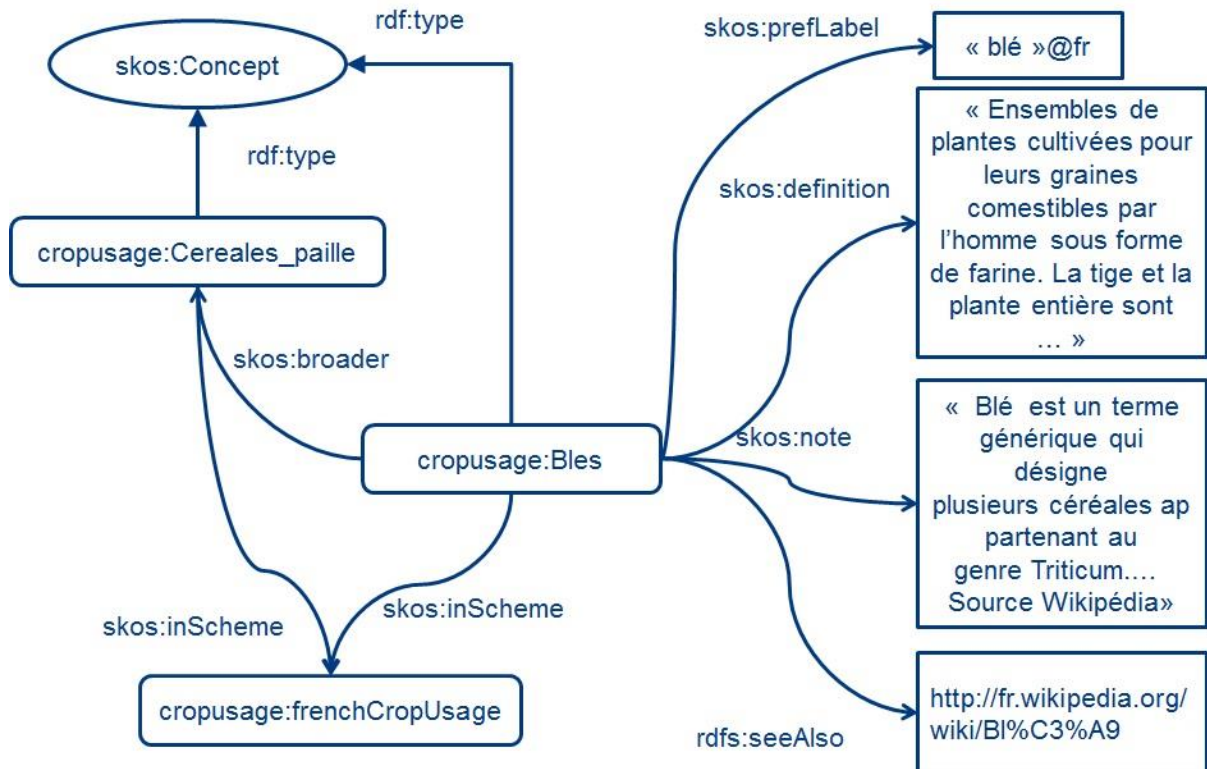


FIGURE 1 : extrait du thésaurus FrenchCropUsage

4 Liste des contraintes

Cette section liste les contraintes que nous avons utilisées pour le développement de notre thésaurus. Une contrainte est décrite par son identifiant, son niveau de priorité et une description de la requête associée. Le niveau de priorité varie entre contrainte forte > contrainte faible > indication. Pour des raisons de lisibilité nous ne présentons que le code des requêtes les plus courtes. L'ensemble des requêtes SPARQL associées est disponible sur le site ontology.irstea.fr.

4.1 Contraintes sur les labels

LabLang

Contrainte forte : tous les labels doivent être associés à une langue, le français dans notre cas.

Améliorer la qualité d'un thésaurus

Requête : Afficher les labels (préfééré, alternatif, caché) qui n'ont pas la langue française associée.

ConcLabPref

Contrainte forte : dans notre thésaurus, un concept skos doit avoir un label préfééré.

Requête : afficher les concepts qui n'ont pas de label préfééré en français.

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?concept WHERE {
  ?concept rdf:type skos:Concept.
  FILTER NOT EXISTS {
    ?concept skos:prefLabel ?label.
    BIND ( LANG(?label) AS ?LANG)
    VALUES ?LANG {"fr"}
  }
}
```

UniLabPref

Contrainte forte : un label préfééré doit être associé à un unique concept.

Requête : afficher les labels préféérés qui sont associés à plusieurs concepts.

```
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?label (COUNT(?concept) AS ?nb) WHERE {
  ?concept rdf:type skos:Concept.
  ?concept skos:prefLabel ?label.
}
GROUP BY ?label
HAVING (?nb > 1)
```

LabDiff

Contrainte faible : un label préfééré doit être distinct d'un label alternatif pour un concept donné.

Requête : trouver les concepts dont un label alternatif est équivalent à un label préfééré.

```
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>

SELECT ?label1 ?concept
WHERE{
  ?concept rdf:type skos:Concept.
  ?concept skos:prefLabel ?label.
  ?concept skos:altLabel ?label.
}
```

Polysem

Indication : éviter autant que possible les répétitions entre les différents types de labels, sauf en cas de polysémie où un label est associé à deux concepts.

Requête : trouver les concepts qui partagent des labels (préféérés ou alternatifs) en commun.

```
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?label1 (COUNT (DISTINCT ?concept1) AS ?nbc)
```

RISE 2017

```

WHERE {
  ?concept1 rdf:type skos:Concept.
  ?concept1 ?p ?label1.
  {
    SELECT ?label ?concept
    WHERE {
      {
        ?concept skos:prefLabel ?label.
      }
      UNION
      {
        ?concept skos:altLabel ?label.
      }
    }
  }
  FILTER (?label1=?label && ?concept1=?concept).
}
GROUP BY (?label1)
HAVING (?nbc > 1)

```

4.2 Contraintes sur les notes, définitions*NoteLang*

Contrainte forte : une définition, une note, une note éditoriale doit être associée à une langue, le français dans notre cas.

Requête : trouver les concepts dont la définition, la note ou la note éditoriale ne sont pas associées à la langue française.

ConcDef

Contrainte forte : un concept doit avoir une définition.

Requête : trouver les concepts qui ne sont pas associés à une définition.

```

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>

SELECT ?concept WHERE {
  ?concept rdf:type skos:Concept.
  FILTER NOT EXISTS {
    ?concept skos:definition ?def.
  }
}

```

SupNoteEdition

Contrainte forte propre à notre thésaurus : pour la publication sur le web, les notes éditoriales doivent être supprimées.

Requête : trouver les concepts qui ont des notes éditoriales.

```

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?concept WHERE {
  ?concept rdf:type skos:Concept.
  FILTER EXISTS {

```

```

        ?concept skos:editorialNote ?note.
    }
}

```

4.3 Contraintes sur les schemas

HasSchema

Contrainte forte : un concept doit être associé à un schéma.

Requête : trouver les concepts qui ne sont pas associés à un schéma.

```

PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
SELECT ?concept WHERE {
    ?concept a skos:Concept.
    FILTER NOT EXISTS
        {?concept skos:inScheme ?aScheme.}
}

```

HasTopConc

Contrainte forte : un schéma doit avoir des top concepts.

Requête : trouver les schémas qui n'ont pas de top concepts.

```

PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
SELECT ?schema WHERE {
    ?schema a skos:ConceptScheme.
    FILTER NOT EXISTS
        {?schema skos:hasTopConcept ?concept.}
}

```

4.4 Contraintes sur la hiérarchie

CheckRoot

Contrainte forte : tout concept doit être rattaché à la racine par un chemin de propriétés skos:broader.

Requête : recherche les concepts qui ne sont pas rattachés à la racine par un chemin de propriétés skos:broader. Le respect de la règle hasTopConc est une précondition à cette requête.

```

PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
SELECT ?concept WHERE {
    ?concept rdf:type skos:Concept.
    FILTER NOT EXISTS {
        ?concept skos:broader* ?parent.
        ?parent a skos:Concept.
        ?parent skos:topConceptOf ?schema.
        ?schema a skos:ConceptScheme .
    }
}

```

TopOrphelin

Contrainte forte : les top concepts du schéma ne doivent pas avoir de père.

RISE 2017

Requête : trouver les top concepts qui ont un père.

```
PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
SELECT ?concept WHERE {
    ?concept rdf:type skos:Concept.
    ?schema skos:hasTopConcept ?concept.
    FILTER EXISTS {
        ?concept skos:broader ?parent.
    }
}
```

CheckInverse

Contrainte forte : si un lien skos:broader existe le lien inverse skos:narrower doit aussi exister (et inversement).

Requête : trouver un lien skos:narrower qui n'a pas son lien inverse skos:broader.

Requête : trouver un lien skos:broader qui n'a pas son lien inverse skos:narrower.

```
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?parent ?enfant WHERE {
    ?parent a skos:Concept.
    ?enfant a skos:Concept.
    ?parent skos:narrower ?enfant.
    FILTER NOT EXISTS{ ?enfant skos:broader ?parent.}
}
```

CheckBi

Contrainte forte : les liens skos:related doivent être bidirectionnels.

Requête : trouver un lien skos:related qui n'a pas son inverse.

```
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?source ?cible WHERE {
    ?source a skos:Concept.
    ?cible a skos:Concept.
    ?source skos:related ?cible.
    FILTER NOT EXISTS{ ?cible skos:related ?source.}
}
```

CheckTrans

Indication : rechercher des liens transitifs dans la hiérarchie des skos:broader ou skos:narrower. Ces propriétés ne doivent indiquer que des liens directs.

Requête : trouver un lien skos:broader (ou skos:narrower) qui s'exprime aussi par un chemin de liens skos:broader (ou skos:narrower).

```
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?grandparent ?parent ?enfant WHERE {
    ?grandparent a skos:Concept.
    ?parent a skos:Concept.
    ?enfant a skos:Concept.
    ?grandparent skos:narrower+ ?parent.
    ?parent skos:narrower+ ?enfant.
    FILTER EXISTS{ ?grandparent skos:narrower ?enfant.}
}
```

CheckCycle

Contrainte forte : il ne doit pas y avoir de cycle entre les liens skos:narrower (idem pour les liens skos:broader).

Requête : recherche des cycles.

```
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?a ?b WHERE {
  ?a a skos:Concept.
  ?b a skos:Concept.
  ?a skos:broader+ ?b.
FILTER EXISTS{ ?b skos:broader+ ?a.}
}
```

4.5 Contraintes de domaine propre à notre thésaurus*CheckCoherence*

Indication, contrainte propre à notre thésaurus. Il faut vérifier que le parent est compatible avec son enfant. Nous avons 6 grandes catégories de cultures. Un concept parent qui est un descendant d'une catégorie doit avoir généralement sa descendance sous cette catégorie uniquement.

Requête : rechercher les descendants d'un concept qui ne sont pas descendants de la même catégorie de culture. Les résultats de cette requête doivent être validés par un expert. Les céréales sont par exemple classées dans « grande culture » et dans « culture fourragère » donc l'association « grande culture » et « culture fourragère » est acceptée pour toutes les céréales.

CheckAmbiguity

Indication : contrainte propre à notre thésaurus. Il faut vérifier qu'un concept ambigu a bien des enfants qui sont catégorisés dans des cultures différentes (parmi les 6 grandes catégories de culture).

Requête : trouver parmi les concepts ambigus (les fils directs de la racine qui ne sont pas dans les 6 grandes catégories) ceux dont les enfants sont tous classés dans la même catégorie.

CheckPatternCropMultipleUsage

Indication : il faut valider auprès d'experts les cultures qui ont plusieurs usages (cas des céréales fourragères).

Requête : trouver les concepts de culture qui ont plusieurs parents directs.

```
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT ?cropLabel (COUNT(DISTINCT ?parentLabel) AS ?nb) WHERE {
  ?crop a skos:Concept.
  ?crop skos:prefLabel ?cropLabel.
  ?crop skos:broader ?parent.
  ?parent skos:prefLabel ?parentLabel.
}
GROUP BY ?cropLabel
HAVING (?nb > 2)
```

5 Travaux connexes

RISE 2017

Le W3C met à disposition une liste d'outils pour valider les modèles SKOS. Les plus connus sont Poolparty Thesaurus Consistency Checker et Skosify (Suominen & Hyvönen, 2012). Ces outils vérifient non seulement les contraintes d'intégrité spécifiées dans (Miles & Bechhofer, 2009) mais aussi d'autres contraintes structurelles propres au thésaurus multilingue, comme la couverture homogène des langues dans le thésaurus, la détection de concepts isolés etc... Comme le montre le tableau 3, certaines de nos requêtes sont une spécialisation de ces contraintes d'intégrité génériques au thésaurus.

L'outil qSKOS (Mader et al, 2012) propose 15 contraintes que les modèles SKOS doivent vérifier. qSKOS met à disposition les fonctions pour vérifier ces contraintes. Parmi ces contraintes on retrouve celles testées dans les deux outils précédents, plus de nouvelles contraintes liées à la publication sur le web de données liées (test des liens entrants, des liens sortants, de la validité des URI etc...). Les travaux de (Lacasta et al, 2016) vérifient que les thésaurus suivent la norme ISO 25964. Ils vérifient en plus que les liens entre concepts soient informatifs. Par exemple un lien skos:related ne doit pas relier deux concepts liés par des liens hiérarchiques (skos:broder ou skos:narrower). Pour valider les liens hiérarchiques, ces travaux projettent le thésaurus sur une ressource générique comme Wordnet ou dolce. Ainsi la hiérarchie du thésaurus est comparée à la hiérarchie extraite de ces ressources.

TABLE 3 : Comparaison entre 4 outils de validation de modèle SKOS

Nom de contraintes	Poolparty checker	Skosify (Suominen & Hyvönen, 2012)	qSKOS (Mader et al, 2012)	(Lacasta et al. 2016)
LabLang	X	X	X	X
ConcLabPref	X	X	X	X
UniLabPref	X	X	X	X
LabDiff	X	X	X	X
Polysem	X	X	X	X
NoteLang	X	X	X	X
ConcDef			X	X
hasSchema				
hasTopConc			X	
CheckRoot		X		
topOrphelin			X	
CheckInverse				
CheckBi				
CheckTrans				
CheckCycle		X	X	X

Nos travaux sont une implémentation en SPARQL de la vérification de ces contraintes mais en utilisant uniquement l'interface d'interrogation d'un Sparql-endpoint. Nous pouvons ainsi tester de nouvelles contraintes propres à ce thésaurus. Certains des travaux précédents ont eu pour but d'évaluer les thésaurus en proposant des métriques. Le but est de choisir le meilleur thésaurus pour une tâche donnée. Notre proposition a uniquement pour but de détecter les erreurs pour aider le développeur à modifier son thésaurus. Nos requêtes sont des tests unitaires faciles à mettre en œuvre qui s'appliqueront à chaque modification du thésaurus. Nous espérons que cette base de requêtes mise à disposition pourra être réutilisée et adaptée au cours du développement de notre thésaurus et d'autres. Les outils précédents peuvent nous permettre de proposer de nouvelles requêtes. En effet notre thésaurus ne contient pas actuellement de liens skos:related ou de liens de correspondance (skos:exactMatch ou skos:broadMatch).

6 Conclusion

Améliorer la qualité d'un thésaurus

SKOS est un schéma RDF utilisé pour stocker et publier sur le web de données liées les thésaurus ou taxonomies. Lors du développement d'un thésaurus, il est utile de détecter automatiquement les concepts qui violent les contraintes d'intégrité. Cet article présente le cas d'usage d'un développement manuel d'un thésaurus monolingue. Ce thésaurus porte sur l'usage des cultures en France.

Dans cet article, nous avons spécifié des contraintes propres à un thésaurus. Pour vérifier ces contraintes un ensemble de requêtes SPARQL a été défini. Ces requêtes permettent à la fois de vérifier des contraintes d'intégrité propres à tous les thésaurus stockés au format SKOS mais aussi des contraintes propres à ce thésaurus agricole.

Références

la statistique agricole annuelle présentation générale. Disponible à l'url

http://www.agreste.agriculture.gouv.fr/IMG/pdf_methosaa.pdf

LACASTA J., FALQUET G., F. ZARAZAGA-SORIA J., NOGUERAS-ISO J (2016). An automatic method for reporting the quality of thesauri. *Data & Knowledge Engineering* Vol 104, Pages 1–14

<https://doi.org/10.1016/j.datak.2016.05.002>

Larousse agricole Édition 2002. Disponible à l'url <http://www.larousse.fr/archives/agricole/>

MADER C., HASLHOFER B., ISAAC A. (2012) Finding quality issues in SKOS vocabularies. *Theory and Practice of Digital Libraries*, 2012, p. 222-233.

MILES A. & BECHHOFER S. (2009) SKOS simple knowledge organization system reference.

Portail: Agriculture et agronomie. Disponible à l'url

https://fr.wikipedia.org/wiki/Portail:Agriculture_et_agronomie

Description de la couche Registre parcellaire graphique 2012 (îlots PAC) Métadonnée du 24/09/2013. Disponible à l'url

http://piece-jointe-carto.developpement-durable.gouv.fr/DEPT063A/METADONNEES/N_RPG_2012_S_063_metadonnees.pdf

SUOMINEN O. & HYVÖNEN E.(2012) Improving the quality of SKOS vocabularies with Skosify.In : *International Conference on Knowledge Engineering and Knowledge Management*. Springer Berlin Heidelberg, p. 383-397.

page de Wikipédia sur les classements des cultures disponible à l'url:
https://fr.wikipedia.org/wiki/Classement_en_France_des_cultures_par_groupes_d'usage