



HAL
open science

LinkedVocabularyEditor: une extension MediaWiki pour l'édition collaborative et la publication de vocabulaires liés

Maxime Lefrançois, Antoine Zimmermann

► **To cite this version:**

Maxime Lefrançois, Antoine Zimmermann. LinkedVocabularyEditor: une extension MediaWiki pour l'édition collaborative et la publication de vocabulaires liés. 26e Journées Francophones d'Ingénierie des Connaissances, IC, Jul 2015, Rennes, France. emse-01158410

HAL Id: emse-01158410

<https://hal-emse.ccsd.cnrs.fr/emse-01158410v1>

Submitted on 1 Jun 2015

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.

LinkedVocabularyEditor: une extension MediaWiki pour l'édition collaborative et la publication de vocabulaires liés

Maxime Lefrançois¹, Antoine Zimmermann¹

École Nationale Supérieure des Mines, FAYOL-ENSMSE, Laboratoire Hubert Curien, F-42023 Saint-Étienne, France
{maxime.lefrancois, antoine.zimmermann}@emse.fr

Résumé : Nous présentons le prototype d'une extension pour MediaWiki qui permet l'édition collaborative d'ontologies par les experts de domaines, ainsi que la publication des ontologies conformément aux bonnes pratiques du W3C.

Mots-clés : Ontologies, Vocabulaires liés, Edition collaborative, MediaWiki.

Introduction. Nous présentons LINKEDVOCABULARYEDITOR¹, le prototype d'une extension pour MediaWiki qui en transforme une instance en une plateforme (i) d'édition collaborative d'ontologies par les experts de domaines, et (ii) de publication des ontologies conformément aux bonnes pratiques du W3C.

Besoins identifiés. Plus précisément, le but de LINKEDVOCABULARYEDITOR est de répondre à l'ensemble des besoins suivants :

- l'outil doit offrir aux experts du domaine des interfaces de visualisation et d'édition simplifiées d'une ontologies et de ses différents modules ;
- l'outil doit permettre aux ingénieurs de connaissances de surveiller l'évolution du vocabulaire : les notifier des changements, permettre le contrôle de versions ;
- l'outil doit parallèlement servir de plateforme de publication des ontologies, suivant les bonnes pratiques énoncées par W3C ;
- l'outil doit s'intégrer dans un outil de travail collaboratif open-source courant, afin d'encourager une communauté de développeurs à l'améliorer.

Les outils existants ne répondent pas à l'ensemble de ces besoins. Par exemple, Protégé et sa version web (Noy *et al.*, 2001; Tudorache *et al.*, 2008) ciblent les ingénieurs des connaissances. La plateforme Neologism (Basca *et al.*, 2008) ne permet pas la gestion de versions. Semantic MediaWiki (Krötzsch *et al.*, 2006) et WikiData (Vrandečić, 2012) se focalisent respectivement sur l'annotation des pages et sur l'édition de données structurées, plutôt que sur l'édition et la publication des vocabulaires qui structurent ces annotations ou ces données.

Principe de l'extension. L'extension LINKEDVOCABULARYEDITOR définit un espace de nom² spécial nommé *Resource:*, synchronisé avec un triple store ARC2³. Chaque page y représente la description d'une ressource, comme suit : si le préfixe *seas:* représente `http://`

1. Le code de LinkedVocabularyEditor est disponible sur GitHub - <https://github.com/maximelefrancois86/LinkedVocabularyEditor>; une instance est testable en ligne avec le compte : login : **IC2015**, mot de passe : **lve** - <http://wiki-lve.maxime-lefrancois.info>,

2. Espace de nom MediaWiki - http://fr.wikipedia.org/wiki/Aide: Espace_de_noms

3. ARC2 - PHP&MySQL-based triplestore - <https://github.com/semsol/arc2/>

`purl.org/NET/seas#`, alors la page *Resource:Seas:Site* décrit la ressource `http://purl.org/NET/seas#Site`. Lorsqu'on accède à cette page, une négociation de contenu est opérée, et le wiki sert une représentation HTML, RDF/XML, turtle, ou N-triples suivant l'option HTTP Accept de la requête.

En interne, le contenu d'une page est une représentation en JSON-LD de la ressource décrite, dans chaque graphe nommé du triple store où cette ressource apparaît. Le contenu de la page HTML est généré non seulement à partir de ce JSON-LD, mais également à l'aide d'appels au triple store. L'interface d'édition de la page utilise principalement *angularjs*⁴, et des appels au Endpoint SPARQL de ARC2 pour la recherche de ressources dans le wiki.

Utilisation. Lorsqu'un expert du domaine crée une nouvelle page ou édite une page existante, un formulaire simple lui est proposé, et la validation de ce formulaire met à jour le triple store. L'extension permet déjà d'éditer des ontologies avec le profile OWL LD (Glimm *et al.*, 2012). De son côté, l'ingénieur des connaissances bénéficie des fonctionnalités de notification, modération, et de contrôle de version, déjà implémentées dans MediaWiki. De plus, il a accès à deux pages spéciales. L'une lui permet d'éditer les espaces de noms RDF (pour définir l'espace de nom d'un module d'ontologie par exemple), l'autre lui permet d'importer un vocabulaire existant (elle lui permet d'exporter le vocabulaire, de l'éditer à la main en dehors du wiki, puis de l'importer à nouveau).

Limitations actuelles et évolutions possibles. LINKEDVOCABULARYEDITOR implémente un petit ensemble de fonctionnalités essentielles, qui forme une base intéressante pour le développement et l'évaluation de nouvelles techniques de visualisation et d'édition collaborative d'ontologies. D'autres fonctionnalités manquent, comme le renommage ou la suppression des ressources. Par ailleurs, certaines fonctionnalités de MediaWiki méritent d'être utilisées, comme les hiérarchies de catégories pour représenter les hiérarchies des classes et de propriétés. Enfin, chaque page contient seulement les triplets qui ont la ressource représentée comme sujet. Ce choix temporaire supprime les inter-dépendances entre contenus de pages, et des questions d'ingénierie et de recherche non triviales se poseront lorsque l'on supprimera cette limitation.

Remerciements. Ce travail a été développée dans le cadre du projet ITEA2 14002 SEAS - Semantic Energy Aware Systems - <https://itea3.org/project/seas.html>

Références

- BASCA C., CORLOSQUET S. & CYGANIAK R. (2008). Neologism : Easy vocabulary publishing. In *4th Workshop on Scripting for the Semantic Web*.
- GLIMM B., HOGAN A., KRÖTZSCH M. & POLLERES A. (2012). OWL : Yet to arrive on the Web of Data ? *LDOW'2012*.
- KRÖTZSCH M., VRANDEČIĆ D. & VÖLKEL M. (2006). Semantic mediawiki. In *The Semantic Web- ISWC 2006*, p. 935–942 : Springer Berlin Heidelberg.
- NOY N., SINTEK M. & DECKER S. (2001). Creating semantic web contents with protege-2000. *IEEE intelligent systems*, **16**(2), 60–71.
- TUDORACHE T., VENDETTI J. & NOY N. (2008). Web-Protege : A Lightweight OWL Ontology Editor for the Web. *OWLED'2008*, **432**.
- VRANDEČIĆ D. (2012). Wikidata : A new platform for collaborative data collection. In *WWW'2012*, p. 1063–1064 : ACM.

4. Angular.js - <https://angularjs.org/>