Un projet de Wiktionnaire sémantique franco-arabe-berbère pour les sciences humaines et sociales

Hammou FADILI

Laboratoire CEDRIC du Conservatoire National des Arts et Métiers de Paris

Pôle recherche, axe Humanités Numériques de la FMSH

Résumé : Cet article présente une contribution à la construction d'un Wiktionnaire des sciences sociales et humaines (SHS). Cette dernière est basée sur une extension de la plateforme « MediaWiki sémantique » existante afin qu'elle puisse prendre en compte aussi bien les aspects multiculturels et multidisciplinaires des SHS mais aussi pour une représentation de ses entrées selon le standard ISO 1951 relatif aux dictionnaires. Sa construction devrait permettre aux chercheurs des deux rives de la Méditerranée d'échanger et de partager des connaissances dans le domaine des sciences sociales et humaines et cela quels que soient leurs lieux géographiques de travail et/ou de résidence. Après le rappel du contexte du projet et les choix technologiques effectués, nous détaillerons les caractéristiques fonctionnelles et conceptuelles, suivies d'une brève description de l'application développée avant la conclusion.

Mots-clés : Wiktionnaire, Données liées, Web des données, Web sémantique, Web 2.0, Web 3.0, RDF, SPARQL, Wiki sémantique, LOD

1 Introduction

Les sciences sociales et humaines jouent un rôle primordial dans la compréhension et l'interprétation du contexte économique, culturel et social dans lequel vivent et agissent les populations. L'évolution de la recherche dans ce domaine passe inévitablement par l'échange et le partage des connaissances entre les chercheurs. À la demande de la DRIC (Direction des Relations Internationales et de la Coopération du Ministère français de l'Éducation Nationale), un séminaire avait réuni à la Maison des Sciences de l'Homme (MSH) trente chercheurs maghrébins et français. L'objectif était de réfléchir ensemble sur la situation actuelle de la coopération en sciences sociales et

humaines entre la France et les pays du Maghreb, et sur les moyens de la développer. C'est dans ce contexte qu'a été défini et lancé le programme FSP-Maghreb.

Le comité de pilotage de ce programma a confié à la Fondation MSH la responsabilité de définir et coordonner, en liaison avec les équipes des 23 groupes de recherche ayant participé au FSP, un projet de construction d'un dictionnaire et d'un contenu multilingue, multidisciplinaire et multiculturel en collaboration, dont le laboratoire CEDRIC du Conservatoire des Arts et Métiers est l'un des principaux partenaires. Une fois réalisée, la plateforme du dictionnaire permettra de favoriser les échanges entre chercheurs maghrébins et leurs partenaires français en mettant en commun un ensemble de savoirs automatiquement exploitables, en français, arabe, berbère (amazigh) et anglais, sur les deux cultures et les deux sociétés.

Dire que les contenus doivent être automatiquement exploitables, partageables, etc., cela veut dire qu'ils doivent être mis à disposition dans un contexte de données structurées, sémantiquement annotées et liées. Il s'agit, essentiellement dans notre cas, de définir un schéma de description bien adapté et de faire évoluer les fonctionnalités actuellement offertes par les technologies du Web de données qui est une implémentation pragmatique des évolutions du Web 2.0 (réseaux sociaux) et du Web 3.0 (Web sémantique), afin de prendre en charge les problématiques du domaine des sciences humaines et sociales telles que les aspects multilingues, multidisciplinaires et multiculturels des données. Ceci par l'adaptation et l'utilisation de la technologie « WIKI sémantique » et plus particulièrement de la plateforme MediaWiki sémantique de la fondation Wikimédia ; afin prendre en charge les problématiques liées à la normalisation et à l'intégration du dit dictionnaire dans le Web de données ouvertes (LOD).

L'objectif de cet article est de présenter les technologies utilisées et les travaux réalisés pour la mise en œuvre de l'application du Wiktionnaire sémantique des sciences humaines et sociales (WS-SHS). La section 1 rappelle le contexte et les motivations du présent travail. La section 2 décrit les choix technologiques effectués. La section 3 présente quelques éléments de conception et l'architecture du modèle retenu. La section 3 montre une vue d'ensemble du système développé. La dernière section conclut l'article et rappelle quelques perspectives et travaux futurs.

2 Contexte & motivations

Le programme du FSP-Maghreb vise à appuyer la recherche en SHS au Maghreb et à développer les échanges entre les chercheurs et institutions de recherche maghrébins et leurs partenaires français dans une logique de réseaux et d'ouverture à l'international, notamment dans un cadre euro-méditerranéen. Il vise également à mettre en commun, par une meilleure maîtrise partagée de l'arabe, de l'amazigh et du français, et par une politique d'incitation à la traduction, un ensemble de savoirs sur les deux cultures et les deux sociétés, et à en assurer la diffusion internationale le plus large possible. Il vise ainsi à développer le débat d'idées entre les deux rives de la Méditerranée, et à conforter une communauté scientifique franco-maghrébine engagée dans un dialogue continu.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

Trois objectifs principaux ont été retenus pour y parvenir :

- Mettre en place des réseaux de chercheurs des deux rives de la Méditerranée dans le cadre de programmes de recherche conjointe (réseaux intra-maghrébins et Maghreb-France) ;
- Aider au renforcement et au développement de communautés de jeunes chercheurs en Sciences humaines et sociales au Maghreb, mieux intégrées au sein de la communauté scientifique internationale ;
- Développer l'accès à l'information scientifique récente et l'échange de connaissances.

C'est dans ce contexte que 23 projets thématiques ont été engagés, en plus d'un projet transversal commun à tous les projets, pour la mise en place d'une plateforme de communication, de création et d'échanges entre les chercheurs dans le domaine des SHS. D'un point de vue informatique, cette exigence peut se traduire, par le besoin de mettre en place une « sorte » d'ontologie multilingue, multiculturelle et multidisciplinaire du domaine des sciences sociales et humaines (SHS) ; c'est-à-dire d'un espace et d'une base de connaissances du même domaine partagés et reconnus par la communauté des chercheurs français et maghrébins.

Cette demande est justifiée par le fait :

- Qu'il n'existe aucun dictionnaire des sciences sociales et humaines franco-maghrébin reflétant l'état de la coopération scientifique et culturelle entre la France et le Maghreb dont les lexiques et les concepts pourraient évoluer de manière indépendante.
- Qu'il y ait un grand déséquilibre, d'un point de vue de la disponibilité des ressources numériques, entre le français et les langues du Maghreb.
- Que les dictionnaires peu nombreux soient incomplets.
- Que les traducteurs existants (également peu nombreux) simplifient la traduction en faisant du « mot-à-mot », i.e. les traductions ne prennent pas en compte tous les aspects liés aux contextes des définitions des concepts.
- Que les corpus potentiels pouvant constituer des sources de données soient également rares.
- Que les travaux sur les nouvelles technologies et la normalisation des données des langues du Maghreb soient encore à leurs débuts.
- Etc.

Pour y répondre, nous avons fait le choix, d'une solution basée sur un Wiktionnaire sémantique des SHS (WS-SHS), qui est l'objet du présent article.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

3 Choix technologiques

La solution proposée pour répondre aux exigences exprimées dans le programme repose sur des choix technologiques particuliers, que nous proposons de décrire dans ce qui suit.

3.1 Wiki sémantique, une implémentation de type « Web sémantique »

3.1.1 Wiki

En 1995, Ward Cunningham a inventé les Wikis pour faciliter l'écriture collaborative de contenus.

Un Wiki est en général une application ouverte, universelle et multilingue en ligne. C'est une application de type Système de Gestion de Contenu (CMS) de sites Web permettant à des personnes autorisées, d'éditer et de publier facilement et rapidement des contenus en ligne, de les faire évoluer *via* des processus de travail collaboratif par mutualisation de compétences, une gestion complète des versions et des historiques des contenus et enfin une gestion des notifications permettant à des personnes intéressées par des thèmes particuliers d'êtres alertées à chaque fois que des créations, modifications ou suppressions de contenus par rapport à leurs thématiques favorites se reproduisent.

La technologie Wiki est basée sur des solutions légères moins coûteuses, accessibles à tous, où l'utilisation ne nécessite pas de formation particulière ; ce qui peut inciter à la création et encourager la production de nouveaux contenus. Ces solutions peuvent être perçues comme outils de communication de groupes, capables de tisser des liens entre communautés en créant des « réseaux sociaux » autour de centres d'intérêts communs. Ces « sociétés virtuelles » peuvent être un levier pour la création de groupes, augmenter la production de la « littérature grise » et ainsi créer une « intelligence collective ». Les Wikis peuvent être aussi considérés comme des outils permettant de sauvegarder la mémoire d'un groupe, d'un projet, d'une institution, … et d'en assurer une meilleure diffusion du fait qu'ils sont très populaires et largement consultés. Les Wiki ont été adaptés et utilisés dans plusieurs domaines : espaces de rédaction, espaces pédagogiques, édition/publication électronique, bibliothèques numériques, Blogs, veille technologique, etc.

3.1.2 MediaWiki

MediaWiki est une plateforme Open source, développé à l'origine pour Wikipédia et aujourd'hui utilisé par plusieurs autres projets de l'association à but non lucratif Wikimedia Foundation ainsi que par beaucoup d'autres institutions privées ou publiques. Wikipédia l'un des projets du Wikimedia Foundation créé en 2001, est devenu le site Web le plus visité écrit avec un wiki, développé avec MediaWiki. Néanmoins, le MediaWiki classique et équivalents souffrent de plusieurs limitations, liées surtout à l'interprétation et à l'exploitation de leur contenu par des processus de traitements automatiques. Les contenus de type plein-texte nécessitant des Techniques du Traitement Automatique des Langues (TAL) difficile à mettre en place et dont les résultats de recherche sont encore peu satisfaisants.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

3.1.3 MediaWiki sémantique

Pour remédier à ces limitations, et permettre la création et la gestion de contenus compréhensibles et utilisables au même temps par les machines et par les humains, plusieurs solutions basées sur la notion du Web sémantique ont été mises en place, dont, dans le cas de Wikimedia Foundation, avec l'évolution de MediaWiki vers le MediaWiki sémantique ou Semantic MediaWiki (SMW). Ce dernier a l'avantage de combiner les technologies Wiki et du Web sémantique. Cette extension permet de structurer, rechercher, organiser, naviguer, évaluer et partager des contenus, en les rendant exploitables, aussi bien par les humains que par les machines ; à travers l'utilisation, entre autres, du langage de balises wiki « Wiki-ML » pour les annotations sémantiques, ainsi que beaucoup d'autres extensions annexes.

3.1.4 Semantic forms & pageShema

Semantic forms est une extension du MediaWiki sémantique fournissant des mécanismes de saisie, basés sur des formulaires sémantiques, permettant ainsi la facilité d'édition des données et l'évitement des erreurs de frappes pour la saisie des annotations sémantiques. Cette extension comporte plusieurs utilitaires et outils, comme par exemple, l'introduction de fonctions pour l'auto-complétude des textes, le contrôle du vocabulaire, … La génération de tels formulaires se fait par le biais d'une autre extension appelée « pageSchema » à partir d'un schéma XML facile à éditer et paramétrer *via* un méta-formulaire de définition du schéma.

C'est une solution idéale pour les besoins du projet, car, cela permet la mise en place d'un Wiktionnaire basé sur un schéma dynamique et extensible « à la demande », en fonction des besoins. C'était une réelle exigence au départ du projet. Étant donné, comme décrit précédemment, que nous devons respecter la norme relative aux dictionnaires ISO 1951, très complexe à mettre en place dans sa totalité, notre choix consistait à prendre en compte un sous-ensemble pertinent du schéma, qu'on peut étendre, « à la demande », en fonction des besoins et disposer ainsi d'une implémentation « POTENTIELLEMENT COMPLÈTE » de la totalité de la norme ISO des dictionnaires. Cette souplesse a été décisive pour les choix technologiques retenus.

4 Modélisation du Wiktionnaire

4.1 Aspects fonctionnels

L'analyse des besoins nous a permis de déduire, un premier modèle fonctionnel de l'application. Le Wiktionnaire des SHS doit :

- Permettre une rédaction distribuée selon les localisations géographiques (France et Maghreb) des différentes équipes participant au programme. C'est l'une des plus importantes fonctionnalités du système.

- Être extensible et évolutif, sachant que de nouveaux besoins peuvent apparaître lors de la conception, l'implémentation et l'exploitation du dictionnaire.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

- Être extensible de sorte à englober les langues : anglais, français, arabe et amazigh dans un premier temps puis une extension à toutes les langues du bassin méditerranéen.
- Supporter différents alphabets et graphies des langues retenues.
- Intégrer les différentes disciplines scientifiques du programme FSP, et offrir des passerelles entre celles-ci. Parmi ces disciplines scientifiques, on peut citer l'anthropologie, la sociologie, l'urbanisme, les sciences de l'environnement, etc.
- Être compatible avec les normes dictionnairiques existantes afin de faciliter les échanges.
- Intégrer les fonctions d'import/export pour la compilation des contenus.
- Permettre l'alimentation de la base dictionnaire dans un format structuré suivant la définition d'une entrée et de sa traduction, et vice versa afin de faciliter l'alimentation, les échanges et la génération des contenus.
- Être doté d'interfaces et de techniques lui assurant une grande ergonomie qui facilitera son utilisation par des chercheurs.
- Prendre en charge l'aspect sémantique des définitions des entrées et de leurs traductions afin de permettre des renvois intelligents d'une langue à une autre, ou encore d'un parler à un autre.
- Être libre d'accès en consultation, mais réservé en gestion.

4.2 Aspects conceptuels

La conception du Wiktionnaire en SHS repose sur des correspondances entre les éléments de départ dans leurs contextes pour la langue source et les éléments d'arrivée dans leurs contextes pour la/les langue(s) cible(s) selon un sous-ensemble du schéma de la norme ISO1951.

On doit prendre en compte le fait qu'une entrée A_k dans une langue source peut avoir plusieurs sens et donc plusieurs traductions B_1, …, B_j, …B_m dans la langue cible. Cette même entrée A_k peut être définie avec plusieurs éléments A_1, …, A_i, …A_n du schéma du dictionnaire (synonyme, antonyme, étymologie, hyperonyme, hyponyme, etc.) qui peuvent êtres à leurs tours des entrées dans la même langue source et par conséquent, peuvent avoir plusieurs sens dans cette même langue source et plusieurs traductions dans la langue cible (voir figure 1). Notons que selon le sens de la traduction une langue source peut devenir cible et réciproquement.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

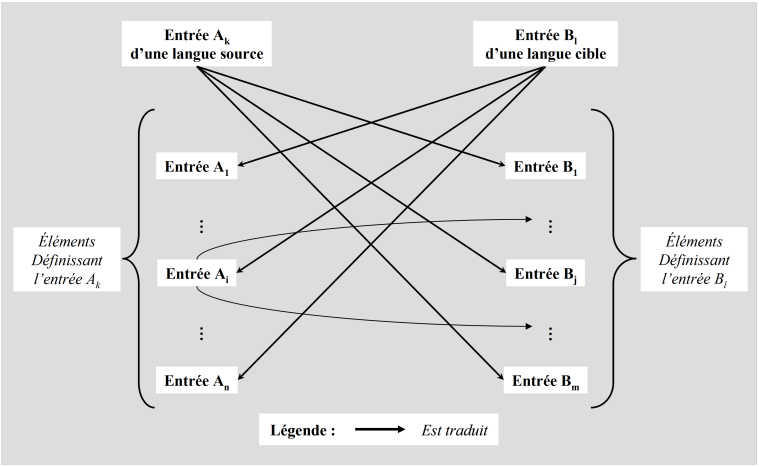


Fig. 1 : Extrait du schéma du dictionnaire des SHS

Ce schéma, récapitulant ce qui a été décrit précédemment, montre la complexité des renvois entre la langue source et la langue cible. Elle nécessite la définition d'un schéma spécifique au Wiktionnaire des SHS dans chaque langue ainsi que la mise en place d'un système de gestion des correspondances d'une manière automatique. Pour ce faire, nous pouvons procéder tout d'abord par une première simplification du problème de départ, qui consiste à associer à une entrée source (mot, locution, etc.) un ou plusieurs sens (définitions) qui renvoient à une ou plusieurs entrées cibles ; puis revenir du terme traduit, pris cette fois-ci comme entrée source. Ce procédé doit être pris en charge par la mise en place d'un système de guidage d'aide à la génération et la définition des correspondances entre les entrées, leurs définitions et leurs traductions dans les différentes langues et que l'utilisateur peut compléter à partir des suggestions du système suivant les critères suivants :

- La définition d'une entrée se fait par la description d'une fiche suivant un format structuré déterminé par le schéma du Wiktionnaire sémantique des SHS et composé des éléments suivant : contexte, définitions, relations sémantiques, traductions, indications grammaticales, parlers, etc.
- La signification attribuée à une entrée dépend du contexte de définition. Ce dernier est décrit par un ensemble fini et connu de paramètres contextuels des aspects : temporels, géographiques, disciplinaires, culturels, linguistiques, etc.
- Les relations entre les termes se font par le biais de relations sémantiques telles que : la synonymie, l'antonymie, l'hyperonymie, l'hyponymie, l'isonymie, la conversion, etc.
- L'ensemble des éléments servant à décrire une entrée fait partie de la norme ISO 1951 [ISO 1951].

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

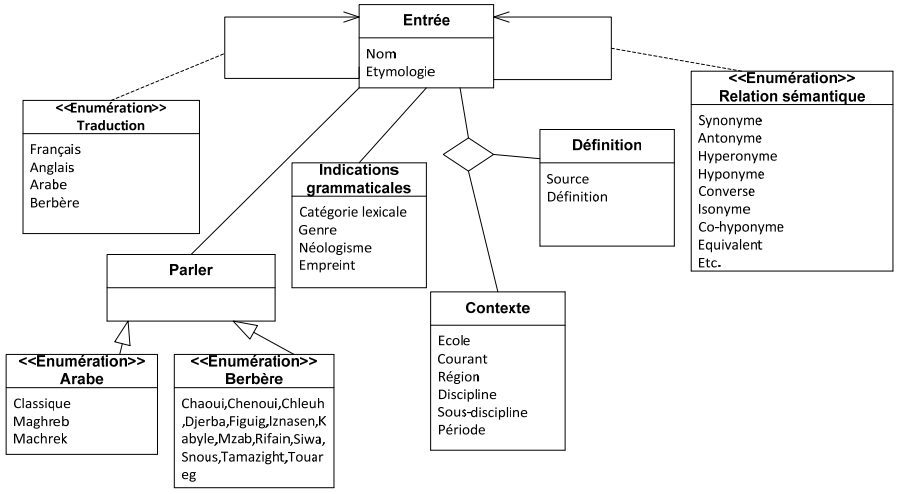


Fig. 2 : modèle du schéma du Wiktionnaire

Pour la mise en place du dictionnaire, nous nous sommes basés sur un sous-ensemble de la norme ISO mentionnée précédemment pour retenir un schéma de base, à partir duquel, on a défini ou traduit des versions pour chacune des langues. Concrètement, nous avons créé deux versions pour chaque formulaire, une en Français et une traduite dans la langue qu'il représente : français, arabe, amazigh ou anglais. Ceci, permettra, à chacun de renseigner le Wiktionnaire dans sa langue et aux non arabophones et aux non amazighophones de le renseigner en utilisant les versions françaises de tous les formulaires.

5 Vue d'ensemble du système développé

Le paragraphe suivant décrit une vue d'ensemble de la plateforme mise en place, à travers 3 sous-paragraphes dédiés respectivement, aux fonctionnalités principales offertes par le système, à nos contributions pour sa mise en place et à un aperçu de l'application à travers quelques copies d'écrans.

5.1 Principales fonctionnalités

L'analyse des besoins nous a permis de déduire un certain ensemble de fonctionnalités importantes que nous devons retenir et implémenter dans la plateforme. À rappeler que l'application été conçue et développée dans le but d'être compatible et intégrable dans le Web de données et dont les données doivent être gérées dans un esprit de partage et coopération. Ci-après une liste non exhaustive des principales fonctionnalités et services implémentés et mis à disposition par la plateforme.

Édition et gestion :

- Gestion ergonomique, collaborative et orientée communautés
- Sécurisation des accès et des contenus

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

- Saisie et annotation des données *via* des formulaires
- Gestion automatique de liens
- Génération automatique de listes
- Complétion automatique de la saisie

Structuration :

- Catégorisation et hiérarchisation des données
- Instanciation des catégories et création des articles
- Annotations et création des propriétés et relations des articles
- Génération automatique « à la demande » des versions multilingues du schéma
- Extension automatique « à la demande » du schéma à d'autres éléments de la norme

Interrogation/navigation :

- Interrogation sémantique les contenus
- Navigation par ontologie
- Recherche à facettes
- Export/import RDF
- Exposition d'un point de terminaison SAPRQL
- Intégration dans le Web de données (triplets RDF)
- Navigation riche entre les langues

Diffusion et enrichissement automatique du dictionnaire :

- Meilleure diffusion : l'intégration du dictionnaire dans le Web de données permet bien sûr une meilleure diffusion ; les sources de données qui lui sont liées sont dans la plupart des cas très connues et populaires, comme par exemple : le dbpedia.
- Enrichissement automatique : nous avons créé pour cela des requêtes paramétrées SPARQL permettant, suivant le contexte, de générer des contenus bien spécifiques, soit pour créer de nouvelles pages, soit pour enrichir d'une manière automatique la définition d'un concept SHS du dictionnaire. Le Wiktionnaire peut se servir potentiellement de la totalité du Web de données compatibles pour enrichir ses contenus.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

5.2 Contribution et améliorations

L'implémentation de l'application a été réalisée à partir de la version de base de la plateforme MediaWiki sémantique et de ses extensions. Nous avons procédé à des développements, améliorations et adaptations des différents outils, que l'on peut résumer comme suit :

- Développement et adaptation du module associé au méta-formulaire basé sur « pageSchema », pour la génération automatique des formulaires du module « semantic forms », et la prise en charge de toutes les langues dont l'arabe. Les versions arabes de ces modules n'existent pas
- Développement et adaptation du module « semantic forms » pour la génération de modèles d'affichage multilingues dont le cas de l'arabe
- Personnalisation des feuilles de styles en fonction de la langue affichée
- Définition d'un modèle pertinent du schéma du dictionnaire
  - Basé sur des catégories, propriétés, relations, …
  - Définition des versions spécifiques du schéma de base, et ses versions légèrement différentes pour l'Arabe et l'Amazigh
- Définition et génération de tous les formulaires

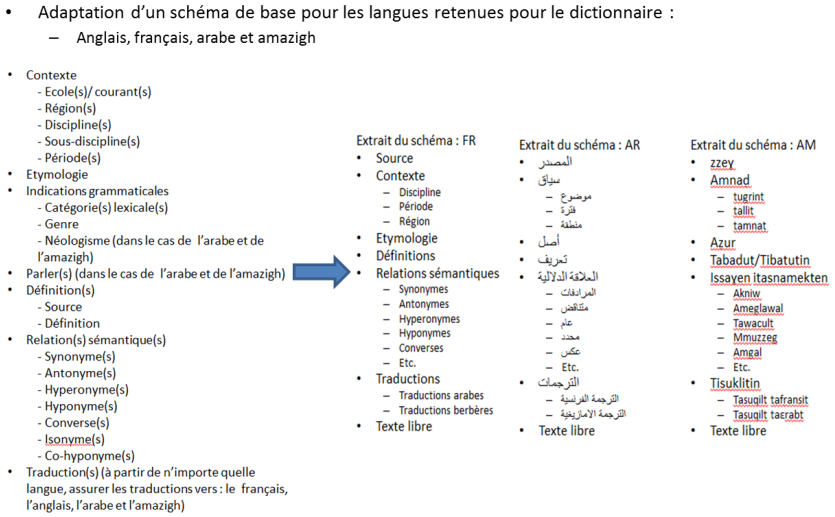


Fig. 3 : Adaptation d'un extrait du schéma du dictionnaire

5.3 L'application

L'application est issue d'une adaptation de la technologie du MédiaWiki sémantique et de ses extensions pour prendre en charge la totalité des besoins de gestion du sous-ensemble du schéma du Wiktionnaire retenu dans les différentes langues. Ce dernier

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

doit être facilement et rapidement extensible « à la demande », et doit couvrir, potentiellement, la totalité de la norme ISO1951. Nous avons pris en compte les spécificités de chaque langue, en particulier de l'arabe et son intégration dans le modèle de saisie orienté formulaire, comme par exemple l'écriture de la droite vers la gauche. La solution a l'avantage de nous offrir la possibilité de réaliser, dans un premier temps, un Wiktionnaire anglais-français-arabe-amazigh, facilement extensible à d'autres langues pratiquées dans le bassin méditerranéen, contrainte majeure d'implémentation spécifiée dans les besoins initiaux du projet.

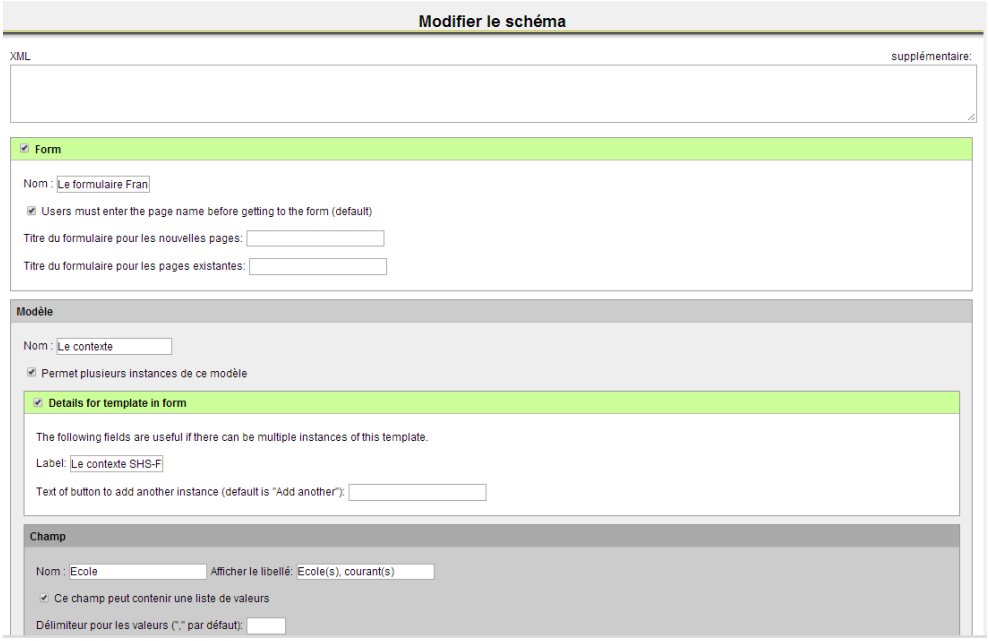


Fig. 4 : Formulaire de gestion des schémas

Le méta-formulaire de création du schéma du Wiktionnaire permettant la génération automatique des formulaires de saisie ou d'annotation sémantique.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

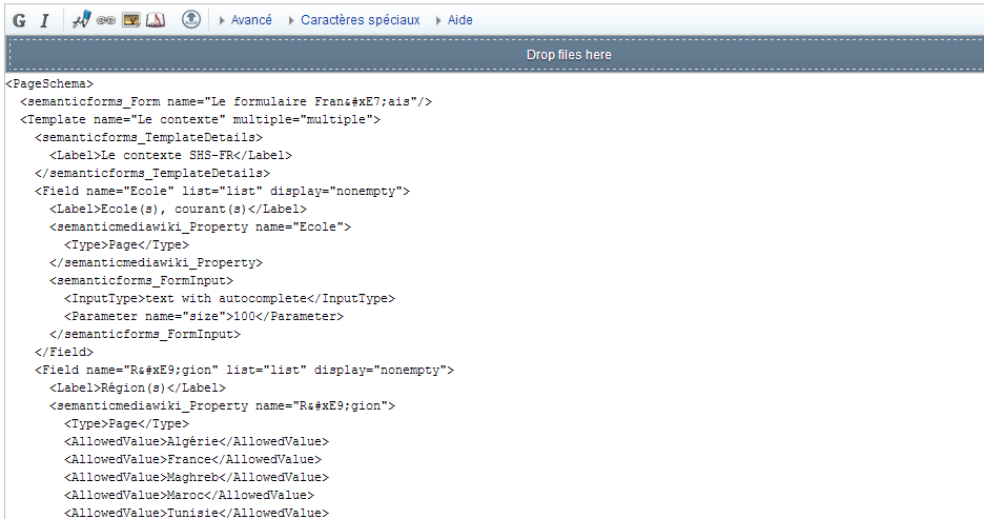


Fig. 5 : Version XML générée du schéma

Version XML du schéma généré par le méta-formulaire de gestion de schémas précédent.

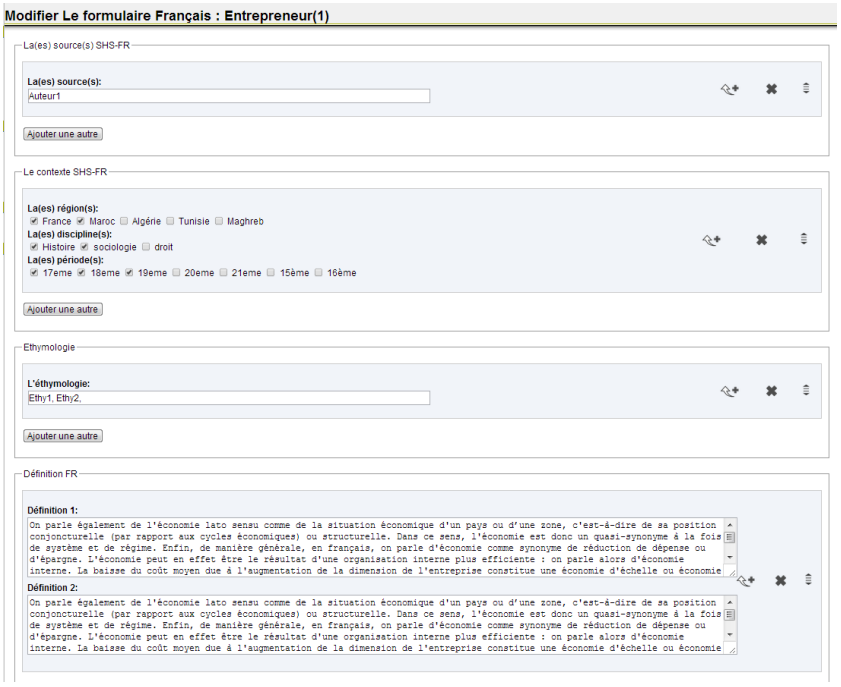


Fig. 6 : Formulaire d'édition.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

Formulaire d'édition implémentant le schéma du Wiktionnaire retenu à un moment donné pour une langue donnée.



Fig. 7 : Interface de navigation

Interface de lecture des entrées associées à des articles annotés suivant les modèles de description associés à chaque langue.

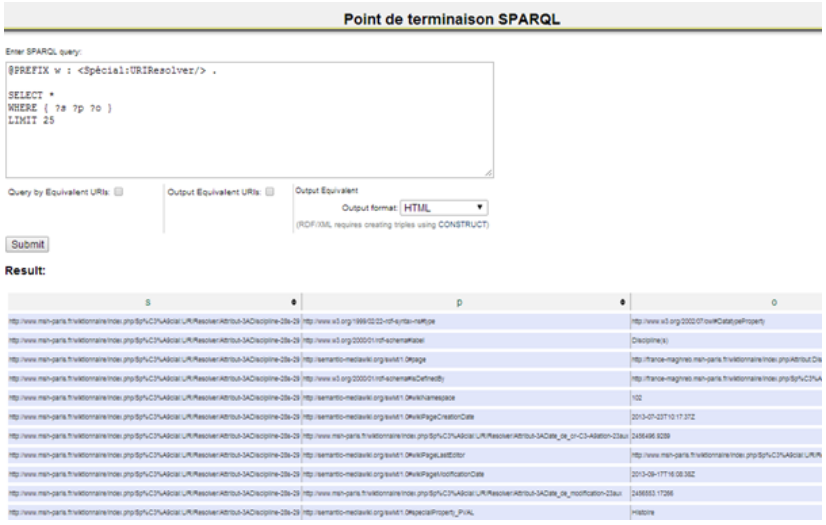


Fig. 8 : Point de terminaison et interrogation SPARQL

Interface de terminaison et d'interrogation SPARQL.

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

6 Conclusion et travaux futurs

Nous avons présenté dans cet article une approche de construction d'un dictionnaire multilingue, multidisciplinaire et multiculturel franco-maghrébin pour les sciences humaines et sociales. Ce projet a été initié par la Fondation Maison des Sciences de l'Homme.

L'architecture mise en place repose sur une conception générique représentant le choix du schéma et des langues retenus à un moment donné, mais facilement extensible à d'autres schémas et généralisable à d'autres langues. La plateforme développée est une extension de la plateforme MediaWiki et plus particulièrement de la plateforme MediaWiki sémantique, garantissant une utilisation ergonomique : saisie, annotations, consultation, navigation et recherche sémantiques. Elle est en phase de production après plusieurs mois de développements-tests-améliorations et une validation par un comité d'experts/chercheurs du domaine.

Comme perspective de recherche à ce travail, on prévoit, l'intégration du Wiktionnaire dans le Web de données, puis l'utilisation des données, structurées, sémantiquement annotées et liées (intégrées dans le Web de données) pour enrichir un travail de recherche parallèle, concernant une nouvelle approche de détection automatique et contextuelle du sens dans le texte : Basé sur les relations lexico-sémantiques et la notion de contexte.

Bibliographie

Barbara Cassin (éd.), *Vocabulaire européen des philosophies. Dictionnaire des intraduisibles.* Le Seuil/ Le Robert, Paris, 2004.

Michel Buffa, Gaël Crova, Fabien Gandon, Claire Lecompte, Jeremy Passeron, SweetWiki : A semantic wiki. Special Issue of the Journal of Web Semantics on Semantic Web and Web 2.0, Volume 6, Issue 1, Pages 84-89, February 2008.

ISO TC 37/SC 2/N 323. « Presentation/Representation Of Entries In Dictionaries ». 2006-03-28

Kensaku Kawamoto, Yasuhiko Kitamura and Yuri Tijerino, KawaWiki : A SemanticWiki Based on RDF Templates. Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT 2006 Workshops WI-IATW'06), 2006.

Malte Kiesel : Kaukolu : Hub of the Semantic Corporate Intranet. Workshop From Wiki to Semantics, ESWC 2006.

Markus Krötzsch, Denny Vrandecic and Max Völkel : Semantic MediaWiki. The 5th International Semantic Web Conference, ISWC2006, 2006

Sebastian Schaffert, IkeWiki : A Semantic Wiki for Collaborative Knowledge Management. 15th IEEE International Workshops on STICA06, Manchester, UK, June 2006.

Webographie

1. http://www.w3.org/DesignIssues/LinkedData.html

2. http://linkeddata.org/

3. http://linguistics.okfn.org/2011/05/20/the-open-linguistics-working-group/

4. http://ldl2012.lod2.eu/

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260

5. http://inkdroid.org/lod-graph/
6. http://www.MediaWiki.org/wiki/Special:ExtensionDistributor
7. https://semantic-MediaWiki.org/

Published in : JAMBOROVA LEMAY, Diana ; PEREIRA, Christophe ; ŠMILAUER, Ivan (dir.), *Des langues calculables à l'homme incalculable. Hommage à Patrice Pognan*, Editions des archives contemporaines, ISBN : 9782813004260, doi : 10.17184/eac.9782813004260