HAL
open science

Intégration d'ontologies médicales : améloriation par association des maladies humaines à leurs plus pertinents signes caratéristiques

Adama Sow, Abdoulaye Guissé, Oumar Niang

To cite this version:

Adama Sow, Abdoulaye Guissé, Oumar Niang. Intégration d'ontologies médicales : améloriation par association des maladies humaines à leurs plus pertinents signes caratéristiques. 29es Journées Francophones d'Ingénierie des Connaissances, IC 2018, AFIA, Jul 2018, Nancy, France. pp.261-263. hal-01839637

HAL Id: hal-01839637

https://hal.science/hal-01839637v1

Submitted on 23 Jul 2018

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.

Intégration d'ontologies médicales : amélioration par association des maladies humaines à leurs plus pertinents signes caratéristiques

Adama Sow, Abdoulaye Guissé, Oumar Niang

LABORATOIRE TRAITEMENT DE L'INFORMATION ET SYSTÈMES INTELLIGENTS (LTISI)
Département du Génie Informatique et Télécommunications
Ecole Polytechnique Thiès (EPT) , THIES, Sénégal
asow@ept.sn, aguisse@ept.sn, oniang@ept.sn

Résumé : Guérir un patient malade nécessite un diagnostic médical avant de proposer un traitement approprié. Avec l'explosion des connaissances médicales, nous nous intéressons à leur exploitation pour aider le médecin à collecter des informations et prendre des décisions lors du processus de diagnostic médical. Le présent article propose une ontologie à partir d'une intégration de plusieurs ontologies et terminologies médicales existantes et ouvertes. Nous constituons une nouvelle ontologie fédératrice couvrant toutes les maladies humaines en incluant des liens avec leurs signes caractéristiques pertinents. Pour prendre en compte cette association, l'ontologie produite est alimentée continuellement par apprentissage des signes à partir d'une base de cas réels de diagnostics cliniques confirmés et validés.

Mots-clés : diagnostic médical, ontologies médicales, intégration d'ontologies, web de données, systèmes e-santé

1 Introduction

Le diagnostic médical, tel que décrit dans le livre de Balogh *et al.* (2015) est une activité cognitive centrée sur le patient dont la compétence quintessentielle appartient au médecin. C'est un procédé qui consiste en une collecte continue des informations médicales qu'effectue le médecin avant de les intégrer et de les interpréter pour la gestion des problèmes de santé de son patient. Cette première étape de collecte est aussi capitale que complexe pour le médecin surtout lorsque cela nécessite de recourir rapidement, en un temps réduit, à des masses de connaissances médicales qui ne cessent d'exploser à l'échelle internationale. C'est dans l'optique d'assister les médecins dans l'exploitation de ces connaissances, que se situe notre recherche. Nous mettons ici le focus sur le nombre important d'ontologies médicales, et l'objectif est d'avoir une ontologie centre qui réportorie toutes les informations pertinentes dans l'élaboration d'un diagnostic médical.

Les ontologies médicales ont été conçues (Hoehndorf *et al.* (2015); Anbarasi *et al.* (2013)) pour mettre en place des vocabulaires médicaux communs reposant sur des concepts partagés qui facilitent l'interopérabilité des documents entre les acteurs du domaine et surtout l'élaboration des connaissances. Nous nous intéressons ici aux ontologies médicales des maladies humaines. La liste est longue et chaque ontologie présente ses propres spécificités. Mais globalement toutes les maladies sont couvertes et renvoient chacune à un concept regroupant ses divers termes nominatifs et leurs synonymes, ses différentes définitions et axiomes textuels et ses signes caractéristiques. Ces derniers indiquent entre autres des signes cliniques et des symptômes (Cox *et al.* (2014)), mais aussi éventuellement l'agent en cause de la maladie, le mode de transmission, et la localisation dans l'anatomie humaine.

Le présent article porte sur une fédération de diverses ontologies. En effet, dans les ontologies existantes nous trouvons d'une part des ontologies de maladies associées à des signes généraux dont l'exhaustivité est à éclaircir, et d'autre part des ontologies qui conceptualisent tous les signes susceptibles d'être identifiés chez un malade mais aucun lien avec les maladies concernées n'est identifié. Nous nous intéressons alors à une intégration de ces ontologies afin de lier chaque maladie à ses plus pertinents signes. Pour prendre en compte cette association, l'ontologie produite est alimentée continuellement par apprentissage des signes à partir d'une base de cas de diagnostics cliniques.

2 Méthodologie de fédération

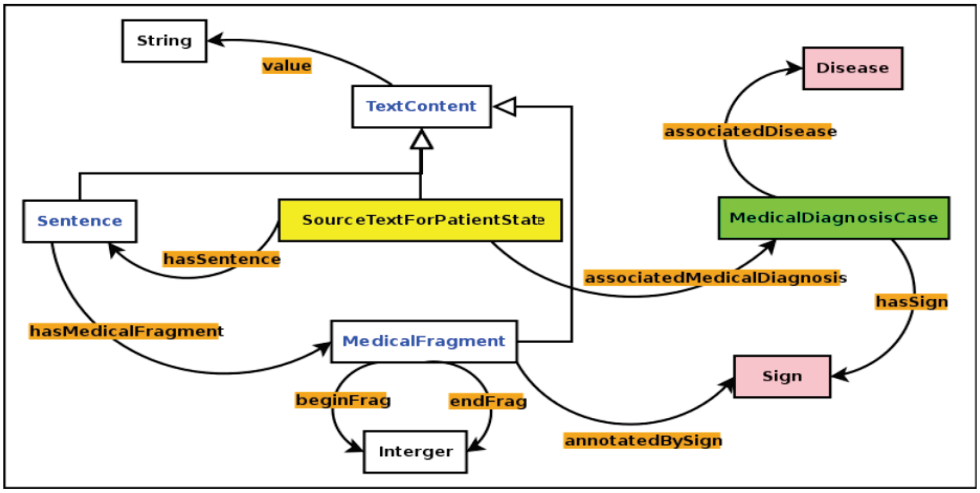

FIGURE *1 – Structure de données d'ensemble*

La constitution d'une ontologie de maladies et de signes consiste à une fédération (Figure 1) d'un ensemble d'ontologies autour d'une structure unifiant toutes les maladies humaines ainsi que leurs signes caractéristiques. Les maladies correspondent aux diagnostics possibles. Les signes sont ceux susceptibles d'être identifiés sur un patient afin de conclure sur un diagnostic précis qui lui peut renvoyer à une ou plusieurs maladies.

Les maladies sont organisées de façon hiérarchisée ; elles et leurs formes dérivées sont regroupées par catégories, qui peuvent elles-mêmes être composées de sous-catégories de maladies. Les maladies sont lexicalisées afin d'avoir pour chaque maladie l'ensemble des termes nominatifs les plus connus et leurs synonymes. Pour chaque maladie, il sera important de conserver les définitions afin de contrôler la sémantique la mieux partagée. La plupart des signes connus de chaque maladie sont listés formellement à partir de ceux disponibles dans les ontologies médicales cibles.

Nous analysons ici des ontologies médicales mises à la disposition du public via la plateforme BioPortal. Notre choix s'est porté sur la DOID, la MESH, la SNOMED comme ontologies de maladies, ainsi que la SYMP, et la CSSO comme ontologies de signes. L'ontologie est chargée (voire tableau 1 de la Figure 2) par interrogation de ces différentes ressources ontologiques cibles avec le langage de requêtes SPARQL à partir de la plateforme BioPortal. Nous sélectionnons ainsi toutes les maladies qui constituent les feuilles des classes à partir de la DOID, ainsi que leurs définitions à partir de MESH ; toutes les catégories de maladies à partir de la DOID où nous sélectionnons leur nom, leur description, et leurs catégories mères ; tous les termes nominatifs synonymes des maladies à partir de la DOID, mais surtout à partir de MESH, soient le label préférentiel, ainsi que les labels alternatifs pour chaque maladie ; tous les signes caractéristiques de base pour chaque maladie à partir des descriptions semi-formalisées de la DOID ; et enfin tous les termes nominatifs synonymes des signes : les labels préférentiels sont extraits de SYMP, les labels alternatifs sont quant à eux extraits des ontologies CSSO, et SNOMED.

Après avoir constitué notre fédération d'ontologies, nous enrichissons l'ontologie produite à partir de l'analyse de rapports médicaux de cas de diagnostics ayant déjà été validés par des médecins. Le tableau 2 (voire Figure 2) montre les symptômes et les signes cliniques trouvés sur une dizaine de cas réels de patients ayant déjà été diagnostiqués. Les exemples choisis portent sur des maladies tropicales que nous trouvons au Sénégal. Nous pouvons alors remarquer que pour chaque maladie, il y a un nombre précis de symptômes généraux indiqués par notre ontologie de maladies mais la totalité d'entre eux ne sont pas présents chez les

patients. De nouveaux symptômes non répertoriés dans l'ontologie font leur apparition ainsi que les signes cliniques dont les valeurs sont spécifiques à chaque patient.

Types Eléments	Objet ontologique	Ontologies d'origines	Nbr. d'individus
Diagnostics			
Maladies	Disease Class	DOID	6442
Catégories	SetOfDiseases Class	DOID	3947
Termes Synonymes	AnnotationProperty (prefLabel, altLabel, hiddenLabel)	DOID, MESH	27586
Signes			
Symptômes et Signes Cliniques	Symptom Class et CinicalSign Class - subClassOf Sign Class	SYMP	942
Autres signes	PhysicalAgent Class, ChimicalAgent Class, TopographicalLocate Class, MedicalProcedure Class : subClassOf Sign Class	DOID, SNOMED	6020
Termes Synonymes	AnnotationProperty (prefLabel, altLabel)	CSSO, SNOMED	1346

TABLE *1 – Description de l'état actuel de notre ontologie de maladies et de signes*

Maladie	Symptômes indiqués par l'ontologie	Symptômes de l'ontologie présents sur le cas	Symptômes nouveaux	Signes cliniques
Hepatitis A	9	7	7	9
Cholera	5	3	10	2
Rougeaole	6	4	11	3
Dengue	10	5	11	20
Tétanos	4	3	12	4
Paludisme	6	4	8	24
Syphilis	5	2	12	14
Chikungunya	9	5	7	29
FièvreTyphoide	8	5	8	3
Meningite	9	4	5	7

TABLE *2 – Nombre de Simptômes et de Signes cliniques trouvés sur des cas réels*

FIGURE *2 – L'ontologie résultantes en chiffres*

3 Conclusion

Dans cet article, la problématique porte sur la mise en place d'un système d'exploitation des ressources ontologies ouvertes et partagées. Il est question ici de constitution d'une ontologie centrale fédérant un ensemble d'ontologies et terminologies médicales cibles, qui répondent au besoin en informations afin de faciliter la tâche du médecin dans l'identification des diagnostics potentiels, parmi lesquels il aura la latitude de choisir ou de valider le plus fiable en connaissance de cause. Ce type de système ne se substitue donc nullement au médecin. Nous avons donc proposé une méthodologie d'intégration autour d'une structure de graphe RDF facilitant la récupération des maladies humaines et de leurs plus pertinents signes caractéristiques, à partir des ontologies cibles et d'une analyse de cas réels de diagnostics confirmés par des médecins. Au final, nous disposons d'une ontologie de maladies et de signes qui sert de base de connaissances pour l'aide au diagnostic médical.

Références

ANBARASI M., NAVEEN P., SELVAGANAPATHI S. & NOWSATH M. (2013). Ontology based medical diagnosis decision support system. In *Actes de International Journal of Engineering Research and Technology (Traitement automatique des langues naturelles)*.

BALOGH E. P., MILLER B. T. & BALL J. R. (2015). Improving diagnosis in health care. In *Actes de National Academies of Sciences, Engineering, and Medicine. The National Academies Press, Washington, DC (Traitement automatique des langues naturelles)*.

COX A. P., RAY P. L., JENSEN M. & DIEHL A. D. (2014). Defining 'sign' and 'symptom'. In *Actes de IWOOD Workshop, In ICBO, Houston, TX, USA, October 6-7, 2014 (Traitement automatique des langues naturelles)*, p. 101–110.

HOEHNDORF R., SCHOFIELD P. & GKOUTOS G. (2015). The role of ontologies in biological and biomedical research : a functional perspective. In *Actes de Briefings in Bioinformatics Journal, 2015 (Traitement automatique des langues naturelles)*.