

## Le traitement du lexique médical par la terminologie computationnelle : vers la réalisation d'une ressource linguistique interopérable

La terminologie computationnelle est un domaine d'étude qui vise à adopter des méthodes computationnelles et quantitatives afin de mener des recherches terminologiques et qualitatives (Bourigault, Jacquemin, L'Homme 2001 ; Drouin et al. 2018). Parmi les activités de recherche les plus répandues, l'extraction automatique de termes (Foo 2012 ; Amjadian et al. 2018 ; Simon, Kešelj 2018) et la création de ressources terminologiques structurées<sup>1</sup> jouent un rôle fondamental pour le traitement et l'analyse des « langues de spécialité ».

Dans cette étude, nous nous proposons d'analyser le langage médical et nous considérons trois différents macro-domaines (communication, traduction spécialisée et recherche d'information) où la terminologie computationnelle peut être appliquée afin de faire face à la complexité du lexique médical. Dans ce contexte, notre objectif est de créer une ressource linguistique intégrée et interopérable qui puisse fournir un large éventail d'informations pour l'exécution de tâches liées au domaine médical.

Dans le domaine de la communication, les difficultés dans le dialogue entre patient et médecin sont souvent abordées (Ley 1988 ; Serianni 2005) et le code de déontologie médicale<sup>2</sup> déclare expressément la nécessité de « transmettre des informations compréhensibles et complètes ». Dans le domaine de la traduction spécialisée aussi, les professionnels se confrontent avec la complexité du lexique médical (Rouleau 2003) en tant qu'ils doivent accomplir la tâche de décoder (comprendre) et ensuite de transcoder (transmettre) les informations médicales dans un texte en langue source vers une langue cible (Jammal 1999). Enfin, dans le contexte de la recherche d'information, de nombreuses études ont été menées pour faciliter les experts en médecine dans le processus de revues systématiques.<sup>3</sup> La tâche consiste dans la récupération des documents pertinents à un sujet médical spécifique et l'approche classique est basée sur la reformulation du besoin d'information à satisfaire en utilisant des variantes terminologiques.

Afin de faire face aux exigences définies ci-dessus, nous nous proposons d'intégrer les données terminologiques d'une ressource multilingue déjà existante<sup>4</sup> qui vise à supporter la communication dans le domaine médical. Notre objectif est d'intégrer, en termes d'interopérabilité, les informations extraites de différentes sources afin de construire un seul point d'accès aux informations médicales et terminologiques. En particulier, nous intégrerons les données extraites :

1. des ressources *Consumer Health Vocabulary*<sup>5</sup> et Glossaire Multilingue européen<sup>6</sup> pour faciliter la communication entre patient et médecin en fournissant le terme technique et son correspondant en langue populaire ;

2. des bases de données terminologiques IATE<sup>7</sup> et TERMIUM Plus<sup>8</sup> pour fournir un outil valide pour les professionnels dans le processus de traduction spécialisée ;

3. des ressources SNOMED CT<sup>9</sup> et MeSH *terms*<sup>10</sup> pour la reformulation de l'information à satisfaire afin de collecter des documents pertinents pour les revues systématiques.

En conclusion, nous présenterons une série de problèmes liés à l'utilisation du standard ISO TBX<sup>11</sup> pour la réalisation de ressources linguistiques normalisées et interopérables.

---

<sup>1</sup> Pour une liste des ressources linguistiques disponibles dans le panorama européen, voir le lien suivant : <http://termcoord.eu/terminology-websites/>

<sup>2</sup> <http://www.ordinemedicilatina.it/wp-content/uploads/2017/05/CodiceDeontologico.pdf>

<sup>3</sup> <https://swiss.cochrane.org/fr/les-revues-systematiques-systematic-reviews>

<sup>4</sup> Pour des raisons d'anonymat, la référence à la ressource n'a pas été insérée.

<sup>5</sup> <http://consumerhealthvocab.chpc.utah.edu/CHVwiki/>

<sup>6</sup> <https://users.ugent.be/~rvdstich/eugloss/welcome.html>

<sup>7</sup> <https://iate.europa.eu/home>

<sup>8</sup> <http://www.btb.termiumplus.gc.ca/tpv2alpha/alpha-eng.html?lang=eng>

<sup>9</sup> <https://www.snomed.org>

<sup>10</sup> <https://meshb.nlm.nih.gov>

<sup>11</sup> <https://www.iso.org/standard/45797.html>

## Références :

- AMJADIAZ E., INKPEN D., PARIBAKHT T.S., FAEZ F. (2018), “Distributed specificity for automatic terminology extraction”. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 24(1), 23-40.
- BOURIGAULT D., JACQUEMIN C., L'HOMME, M.C. (Eds.) (2001), *Recent Advances in Computational Terminology* (Vol. 2), John Benjamins Publishing.
- DROUIN P., GRABAR N., HAMON T., KAGEURA K., TAKEUCHI K. (2018), “Computational terminology and filtering of terminological information”, *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*, 24(1), 1-6.
- FOO J. (2012), *Computational Terminology: Exploring Bilingual and Monolingual Term Extraction*, Doctoral dissertation, Linköping University, Electronic Press.
- JAMMAL A. (1999), « Une méthodologie de la traduction médicale », *Meta: journal des traducteurs/Meta: Translators' Journal*, 44(2), 217–237.
- LEY P. (1988), *Communicating with Patients. Improving Communication, Satisfaction, and Compliance*, London, Croom Helm.
- SERIANNI L. (2005), *Un treno di sintomi. I medici e le parole: percorsi linguistici nel passato e nel presente*, Milano, Garzanti Libri.
- SIMON N. I., KEŠELJ V. (2018), “Automatic term extraction in technical domain using Part-of-Speech and Common-Word features”, *Proceedings of the ACM Symposium on Document Engineering 2018* (p. 51), ACM.
- ROULEAU M. (2003), « La terminologie médicale et ses problèmes », *Tribuna*, Vol. IV, n° 12, [http://www.medtrad.org/panacea/IndiceGeneral/n12\\_tribuna\\_Rouleau.pdf](http://www.medtrad.org/panacea/IndiceGeneral/n12_tribuna_Rouleau.pdf)