



CALASS 2026 – Montréal, Québec, Canada

27,28 & 29 août 2026

Prévention en santé et efficience dans l'organisation des soins : défis et innovation dans les systèmes de santé

Association Latine pour l'Analyse des Systèmes de Santé (ALASS)

Actes du congrès

Automatisation du codage clinique de dossiers patients francophones grâce au traitement automatique du langage naturel

Peter Heirman^{1, 2}, Antoine Stasse¹, Stéphanie Leroy², Ashwin Ittoo¹, Philippe Kolh^{1, 2}

¹Université de Liège, ²CHU de Liège

1) Introduction/contexte/objectif

« Plus de 1 100 équivalents temps plein sont affectés au Résumé Hospitalier Minimum (RHM) et les logiciels spécifiques coûtent plus de 11 millions d'euros par an. » (Cour des comptes de Belgique, 2017).

L'enregistrement RHM des pathologies et des procédures au moyen de l'International Classification of Diseases, Tenth Revision, Clinical Modification/Procedure Coding System (ICD-10-CM/PCS) constitue l'un des fondements du financement hospitalier belge, notamment via les All Patient Refined Diagnosis Related Groups (APR-DRG).

Cette communication porte sur le développement d'un outil original d'assistance au codage ICD-10-CM/PCS basé sur l'intelligence artificielle et le traitement automatique du langage naturel (NLP) à partir de dossiers cliniques non structurés, tout en garantissant la traçabilité, l'explicabilité, la conformité réglementaire (RGPD, AI Act) et l'acceptabilité organisationnelle.

2) Méthodes

La revue de littérature met en évidence les limites des approches classiques de classification automatique et des grands modèles de langage utilisés seuls, notamment à cause des risques d'hallucination et du manque de traçabilité.

Notre méthodologie combine :

- Des modèles Transformer pour générer des embeddings sémantiques (DrBERT, CamemBERT, Multilingual-E5-Large et All-MiniLM-L6-v2) ;
- Une base vectorielle Qdrant permettant la recherche de dossiers similaires ;
- Un modèle BioMistral utilisé comme couche de réordonnancement (re-ranking) des suggestions;
- Un pipeline ETL permettant l'extraction, le nettoyage et le découpage sémantique des notes cliniques.

Les expérimentations ont été réalisées sur des données synthétiques ainsi que sur des données cliniques réelles couvrant une période de quatre ans.

3) Résultats

Les résultats montrent que les modèles spécialisés dans le domaine médical obtiennent de meilleures performances que les modèles généralistes. Le modèle DrBERT fine-tuné apparaît comme la solution la plus performante pour la recommandation automatique de codes ICD-10-CM/PCS. Dans les tests synthétiques, il atteint les meilleurs scores en Precision@1, Recall@K, Mean Reciprocal Rank (MRR) et F1-score.

L'étude révèle également que :

- Le fine-tuning spécialisé améliore fortement les performances de DrBERT ;
- Les modèles légers ou généralistes se dégradent lorsqu'ils sont adaptés au domaine médical ;
- La similarité cosinus constitue généralement la meilleure métrique pour la recherche sémantique;
- L'approche RAG offre une meilleure explicabilité et réduit les risques d'erreurs par rapport à une génération directe de codes.

4) Discussion et conclusion

Cette recherche apporte une contribution originale en proposant une architecture d'aide au codage démontrant qu'une combinaison de modèles spécialisés, de recherche vectorielle et de LLM locaux peut fournir des recommandations explicables et compatibles avec les contraintes hospitalières.

Sur le plan pratique, le système pourrait réduire la charge de travail des codeurs et améliorer la qualité du codage. Sur le plan scientifique, il contribue aux recherches sur le NLP médical francophone. Enfin, il met en évidence l'importance des enjeux de gouvernance, de conformité réglementaire et de gestion de projet dans le déploiement de l'IA en milieu hospitalier.