Actes de CORIA-TALN 2026
Nantes, France, 29 juin au 3 juillet 2026

Utilisation de modèles de langue de grande taille pour l'extraction de relations temporelles dans les comptes rendus cliniques pédiatriques

Judith Jeyafreeda Andrew[1,2] Juliette Potier[1] Nicolas Garcelon[1] Anita Burgun[1,2,3] Marc Vincent[1]

(1) Clinical Bioinformatics Laboratory, INSERM UMR1163, Imagine Institute, Université Paris Cité, Paris, France

(2) PRAIRIE, PaRis Artificial Intelligence Research InstitutE, Paris, France

(3) Necker Enfants Malades Hospital, AP-HP, Paris, France

judithjeyafreeda@gmail.com

RÉSUMÉ

Objectifs — Évaluer l'utilisation de modèles de langue de grande taille (LLM) pour extraire des relations temporelles à partir de comptes rendus cliniques pédiatriques dans le domaine des maladies rares, en vue d'automatiser la construction de frises chronologiques patient.

Méthodes — Nous avons développé un cadre d'extraction de relations temporelles appliqué à 25 comptes rendus cliniques issus d'un hôpital pédiatrique spécialisé. Trois LLM ont été testés en environnement sécurisé, en utilisant un schéma de *few-shot prompting*.

Résultats — Les approches binaires surpassent nettement les approches multi-classes, avec des F1 atteignant 0,70 pour les relations les plus simples, tandis que les relations complexes restent difficiles (F1 : 0,03–0,40). Mistral 22B obtient les meilleures performances globales, bien que la supériorité varie selon le type de relation.

Discussion — La forte amélioration observée lors du passage à une classification binaire montre que la formulation de la tâche influence fortement l'efficacité des LLM dans les domaines cliniques spécialisés. Notre approche *few-shot* permet d'extraire des relations temporelles à partir de textes cliniques pédiatriques en français tout en respectant les contraintes de confidentialité grâce à un déploiement local.

Conclusion — L'approche proposée montre des performances prometteuses en environnement sécurisé. Elle permet de partager des techniques reproductibles sans exposer de données sensibles, ouvrant la voie à de nouvelles avancées en TAL appliqué aux textes cliniques sous fortes contraintes de gouvernance des données.

ABSTRACT

Using large language models for temporal relation extraction from pediatric clinical reports

Objectives : To evaluate large language models (LLMs) for extracting temporal relations from pediatric rare disease clinical reports to enable automated patient timeline creation.

Materials and Methods : We developed a temporal relation extraction framework for electronic health records, using 25 clinical reports from a pediatric rare disease hospital. We implemented few-shot prompting with 3 different LLMs in secure environments.

Results : Our findings reveal that binary classification significantly outperforms multi-class approaches for temporal relation extraction, with best F1 scores reaching 0.70 for simpler relations

Actes des 33ème Conférence sur le Traitement Automatique des Langues Naturelles.
Volume 2 : articles déjà publiés, pages 56–57.

Cette œuvre est mise à disposition sous licence Attribution 4.0 International.

while more complex relations remain challenging (F1 : 0.03–0.40). Mistral 22B emerged as the strongest overall performer, though model superiority varied by relation type.
Discussion : The dramatic performance improvement from reducing cognitive load (binary vs multi-class classification) demonstrates that task formulation critically impacts LLM effectiveness in specialized clinical domains. Our few-shot approach successfully enables temporal relation extraction from French pediatric texts while maintaining data privacy through local deployment, offering a viable methodology for healthcare institutions with strict data governance requirements.
Conclusion : Our few-shot prompting approach demonstrated promising results in secure environments. This methodology allows technique sharing without exposing sensitive data, advancing research possibilities for clinical natural language processing in restricted settings.

MOTS-CLÉS : extraction de relations temporelles, modèles de langue de grande taille, TAL clinique, maladies rares pédiatriques, few-shot prompting, rapports cliniques en français.

KEYWORDS: temporal relation extraction, large language models, clinical NLP, pediatric rare diseases, few-shot prompting, French clinical reports.

ARTICLE ACCEPTÉ À : JAMIA Open.
URL : https://doi.org/10.1093/jamiaopen/ooaf121