

---

# Extraction de relations dans des textes complexes – Le cas des documents de procès de l’inquisition espagnole

Diego Ortiz<sup>\*1</sup>, Josiane Mothe<sup>\*†1</sup>, David Panzoli<sup>\*‡2</sup>, David Kahn<sup>\*§3</sup>, Michel Boeglin<sup>\*¶4</sup>,  
and Hector Lopez Hidalgo<sup>||5</sup>

<sup>1</sup>Systèmes d’Informations Généralisées (IRIT-SIG) – Université Paul Sabatier - Toulouse III, Université  
Toulouse le Mirail - Toulouse II, CNRS – IRIT118 Route de Narbonne31062 Toulouse Cedex 9, France

<sup>2</sup>Real Expression Artificial Life (IRIT-REVA) – CNRS, Université Paul Sabatier - Toulouse III –  
Institut de recherche en informatique de Toulouse - IRIT2 rue Charles Camichel 31071 Toulouse Cedex  
7, France

<sup>3</sup>INU Jean François Champollion – Institut national universitaire Champollion [Albi] – France

<sup>4</sup>Institut de Recherche Intersite Études Culturelles (Montpellier) (IRIEC) – Université Paul Valéry -  
Montpellier III – Université Paul-Valéry - BRED 112A -Route de Mende - 34 199 Montpellier Cedex 5,  
France

<sup>5</sup>Systèmes d’Informations Généralisées (IRIT-SIG) – CNRS – IRIT118 Route de Narbonne31062  
Toulouse Cedex 9, France

## Résumé

Les relations sémantiques entre entités comme les lieux, les personnes et les concepts sont couramment utilisées pour représenter les connaissances. PromptORE (Prompt-based Open Relation Extraction) optimise l’extraction de relations avec des modèles de langue(LLM) pour des documents généraux, mais s’avère moins performant pour des textes historiques non anglophones. Cet article propose une adaptation de PromptORE pour des documents spécialisés, ici des transcriptions de procès de l’Inquisition Espagnole. Notre méthode ajuste les transformers sur des données spécifiques, un processus que nous avons appelé ” biais ”, pour mieux détecter les entités complexes et les questions de genre dans les textes espagnols. Cette approche est validée par des experts et montre que les modèles PromptORE biaisés améliorent la précision jusqu’à 50 % par rapport aux modèles standards.

**Mots-Clés:** Systèmes d’information, Extraction d’information, Large modèle de langue, Ingénierie du prompt, Documents historiques

---

\*Intervenant

†Auteur correspondant: Josiane.Mothe@irit.fr

‡Auteur correspondant: david.panzoli@univ-jfc.fr

§Auteur correspondant: david.kahn@univ-jfc.fr

¶Auteur correspondant: boeglin.um3@gmail.com

||Auteur correspondant: data\_analytics\_HLH@protonmail.com