

# Knowledge Extraction from Fiction Using Generative Large Language Models and Multilingual Semantic Parsing

Federico Pianzola

University of Groningen, The Netherlands

## Abstract

Extracting event knowledge from unstructured text is a well-known challenge in natural language processing (NLP) and is particularly difficult when dealing with fiction. Subtext, instead of explicit information, and figurative style in fictional narratives complicate event extraction. Recent advances in Large Language Models (LLMs) have improved performance in various NLP tasks. However, their effectiveness in extracting events from fiction remains underexplored. In this talk, I present an evaluation of different techniques to extract narrative information from fiction, including the prompting of open-weights LLMs and the semantic parsing via Abstract Meaning Representation (AMR). This work has been done in the context of the GOLEM project (“*Graphs and Ontologies for Literary Evolution Models*”).

## Short Bio

Federico Pianzola is Assistant Professor of Computational Humanities at the University of Groningen (Netherlands), where he coordinates the Master’s in Digital Humanities. He is the Principal Investigator of the ERC Starting Grant project GOLEM and previously completed a Marie Skłodowska-Curie Global Fellowship at the University of Milan-Bicocca (Italy) and Sogang University (South Korea). He is member of the scientific advisory board of OPERAS (the Research Infrastructure supporting open scholarly communication in the social sciences and humanities in the European Research Area) and president of the International Society of the Empirical Study of Literature (IGEL). He is also managing editor of the journal *Scientific Study of Literature* and he is on the editorial board of three other journals: the *Journal of Computational Literary Studies*, the *Journal of Cultural Analytics*, and the *Korean Journal of Digital Humanities*. His most recent book *Digital Social Reading: Sharing Fiction in the 21st Century* has been published by MIT Press in January 2025.

In: R Campos, A. Jorge, A. Jatowt, S. Bhatia, M. Litvak (eds.): *Proceedings of the Text2Story 26 Workshop, Delft (The Netherlands), 29-March-2026*

✉ [f.pianzola@rug.nl](mailto:f.pianzola@rug.nl) (F. Pianzola)

🌐 <https://www.rug.nl/staff/f.pianzola/> (F. Pianzola)

🆔 0000-0001-6634-121X (F. Pianzola)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).