

PhD Candidate full name:

Hugo Miguel Oliveira de Sousa

Dissertation Title: Unfolding the Temporal Structure of Narratives**Date:** 2026-02-25 14:30**Location:** UP | FCUP | FC5 278**Higher Education Institution:** University of Porto**Doctoral Programme:** Doctoral Program in Computer Science (FCUP)

Abstract or Public Summary: When reading a story or a news article, humans can understand the chronological order of mentioned events even when such information is vaguely defined. This is a fundamental skill for the comprehension of a narrative. For instance, from the sentence "Bob sent a message to Alice while she was leaving her birthday party." we comprehend that the occurrence of the event "sent" is included in the time span of the event "leaving", despite not being explicitly stated on the text. This PhD thesis addresses the task of temporal information extraction, tackling both core challenges and practical applications across multiple domains and languages. We structure the problem into two main components: temporal entity identification and temporal relation classification. For temporal entity identification, we explore methods across diverse settings. We develop a biomedical entity identification pipeline for Portuguese oncology health records, combining neural models with entity linking. We also study the use of large language models to extract narrative entities from Portuguese news articles via prompt engineering, showing that their effectiveness can be comparable with methods that were fine-tuned for the task. Additionally, we introduce TEI2GO, a suite of multilingual models for temporal expression extraction that achieves state-of-the-art results in four of the six languages evaluated. For temporal relation classification, we propose decomposing interval relations into point relations between entity endpoints. This method reaches a temporal awareness score of 70.1% on the TempEval-3 dataset, establishing a new state of the art on this benchmark. Building on this insight, we introduce a novel formulation of the task which recasts relation classification as a sequential decision-making problem. This perspective enables the application of reinforcement learning algorithms to learn temporal reasoning from experience. All research was conducted on top of tieval, a Python library we developed and open-sourced to support the research community. This framework standardizes temporal information extraction evaluation across multiple corpora and provides domain-specific tools such as temporal closure and temporal awareness metrics. The contributions of this thesis range from practical advances in healthcare and multilingual systems to methodological innovations in temporal entity identification

and temporal relation classification. Together, they advance the state of the art and broaden the foundations of temporal information extraction.

Principal Supervisor at INESC TEC: Alípio Jorge

Additional Supervisor: Ricardo Campos (UBI)

Scientific Domain: [Artificial Intelligence]

Keywords: Temporal information extraction, temporal entity identification, temporal relation classification