Relation Extraction from Text of Various Domains: A Serbian Case Study

Ranka Stanković
University of Belgrade
Faculty of Mining and Geol.
Belgrade, Serbia
ranka@rgf.bg.ac.rs

Cvetana Krstev
Language Resources and
Technologies Society–JeRTeh
Belgrade, Serbia
cvetana@jerteh.rs

Milica Ikonić Nešić, Miloš Utvić
University of Belgrade
Faculty of Philology
Belgrade, Serbia
(milica.ikonic.nesic
|misko)@fil.bg.ac.rs

Abstract

Research on Serbian named entity recognition (NER) has evolved from early rule-based approaches to neural and transformer-based models, culminating in the TeslaNER+ dataset, which contains over 150,000 sentences annotated with eight entity types and linked to Wikidata. This paper explores the use of TeslaNER+ for relation extraction (RE), focusing on identifying and classifying semantic relations between entities in unstructured text. A predefined set of relations tailored to various domains, including literary texts is employed. After analyzing existing RE schemas, 19 relevant relations are selected. The methodology combines rule-based patterns with LLM-based prompting, leveraging lexically grounded patterns and prompts derived from key terms. Experiments were conducted on a subset of 23,000 sentences, and the paper reports initial results for selected relations. The FST approach achieves a precision of 0.77 and recall of 0.39, whereas the LLM model attains a higher recall (0.97) but lower precision (0.65). These results are approximate, as they are computed on a sampled subset rather than the full dataset. In terms of efficiency, the FST approach incurs no computational cost and enables fast extraction, while the LLM approach is significantly slower and involves additional expenses.

Keywords: Serbian NLP; Named Entity Recognition; Relation Extraction; Rule-based methods; Large Language Models; Semantic Relations.

1 Introduction

Work on Serbian named entity recognition (NER) has a long history, beginning with early rule-based systems (Krstev et al., 2014) relying on electronic dictionaries and the Unitex[1] corpus processing tools, continuing with convolutional neural network approaches, and more recently transformer-based models. These tools were essential in developing TeslaNER+, a dataset of over 150,000 sentences. It combines automated annotations (Nešić et al., 2024) with manual curation and features seven NER classes linked to Wikidata: PERS, LOC, ORG, DEMO, ROLE, WORK, EVENT. Recently, an additional class, PRODUCT, has been introduced, although it is not yet annotated throughout the entire corpus. TeslaNER+ integrates manually curated datasets from prior research and expands them with newly prepared texts and an additional linked-data layer.

The goal of this research is to use this dataset for the relation extraction (RE) task to detect and classify semantic links between annotated entities in unstructured text. Relation identification detects pairs of entities that are related to each other in the text, while relation classification assigns a predefined label to each identified entity pair, indicating the nature of their relationship. To the best of our knowledge, no prior research on relation extraction from Serbian texts has been published.

The main contributions of this study are: 1) adapting an existing relation schema to the characteristics of Serbian texts from various domains, including literary, 2) constructing a lexically grounded rule-based framework for relation detection using finite-state transducer (FST) patterns, and 3) exploring the complementarity of rule-based and large language model (LLM) approaches for relation extraction. The results contribute to the development of structured semantic resources derived from corpora and support further research in knowledge graph construction and information extraction for Serbian.

In this paper, we present the first preliminary results for the selected set of relations. Section 2 presents an overview of related work in Serbian

[1] https://unitexgramlab.org/

named entity recognition, types of relation extraction with relevant approaches in multilingual settings. Section 3 describes the dataset, the selection of relations, and the proposed methodology, including both rule-based and prompt-based approaches. Section 4 provides descriptive statistics of the dataset and the distribution of entity types and relations. Section 5 outlines the evaluation setup and reports the obtained results. Section 6 discusses the findings, challenges, and limitations of the approaches. Finally, Section 7 concludes the paper and outlines directions for future work.

2 Related Research

2.1 Overview of Serbian NER

Research on named entity recognition (NER) for Serbian spans more than two decades and reflects a gradual evolution from rule-based approaches to modern neural and LLM-based methods. Early work focused on the linguistic properties of proper names and their representation in lexical resources, as explored in the Prolex project (Vitas et al., 2009). This line of research was further extended through the use of electronic dictionaries and corpus processing tools, enabling the development of rule-based NER systems grounded in rich morphological descriptions (Vitas and Krstev, 2012).

Subsequent efforts addressed the recognition and normalization of named entities, emphasizing the integration of linguistic knowledge with corpus-based methods (Krstev et al., 2012). Rule-based approaches based on local grammars and finite-state technologies, particularly within the Unitex framework, proved effective for Serbian due to its morphological complexity (Krstev et al., 2014). At the same time, multilingual perspectives were introduced through resources such as NERosetta, which aimed to position Serbian within a broader cross-lingual NER space (Krstev et al., 2016).

More recent research has shifted toward machine learning and neural approaches, including the development of models for named entity linking, such as SrpCNNeL, which integrates entity recognition with linking to external knowledge bases (Nešić et al., 2024). Parallel corpora and linked data have also been explored as a means to enhance semantic interoperability and support cross-lingual NER and linking tasks (Stanković et al., 2024).

The latest developments highlight the complementary strengths of different paradigms: rule-based, neural, and LLM-based approaches, support an integrated, holistic approach to NER and NEL. Rather than replacing earlier methods, these paradigms contribute distinct advantages, combining linguistic precision with contextual flexibility and scalability. This progression is summarized in recent work tracing the development of Serbian NER from early rule-based systems to contemporary LLM-driven methods (Nešić et al., 2025). Overall, the evolution of Serbian NER and NEL demonstrates a convergence of methodologies, leading to more robust, adaptable, and semantically enriched systems.

2.2 Relation Types

The most common cases of RE focus on the extraction of binary relations within a sentence (sentence-level RE). Also, Diaz-Garcia and Lopez (2025) describe document-level RE (relationships between entities extend beyond individual sentences, spanning entire documents), multilingual RE (extracting relations from multilingual text), open RE or OpenRE (unveiling previously unknown relation types between entities within open-domain corpora), multimodal RE (integration of text and visual data to enhance understanding of relationships), Few-Shot and Low-Resource RE (developing models that generalize well even with minimal labeled examples), and distant RE (extracting relations between entities based on distant supervised data).

For example, in the sentence "Marko Ilić dismissed Ivan Rogić from his staff," an NEL system should recognize "Marko Ilić" and "Ivan Rogić" as entities and identify "dismissed" as the relationship linking them. Some RE systems rely on predefined types (e.g., `DismissedBy`, `TermEmployee` as short for "employment termination"), while some Open Information Extraction (OpenIE) systems extract structured, domain-independent relational triples: (subject; relation; object), directly from unstructured text without requiring a predefined schema, e.g., (Marko Ilić, dismissed, Ivan Rogić).

Watson NLP RE [2] encapsulates algorithms for extracting 32 relation types between two entity mentions, some of which are:

- `parentOf`: exists between a Person and their children or stepchildren.

[2] https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/watson-nlp-block-relation-extraction.html?context=dph

- `educatedAt`: exists between a Person and the Organization at which s/he is or was educated.
- `colleague`: exists between two Persons who are part of the same Organization.

RE supports the identification of relationships in text and plays a key role in building and populating knowledge graphs. It is essential for downstream applications such as question answering (Chen et al., 2019). Additionally, RE contributes to knowledge acquisition in the humanities, particularly in graph-based datasets, enabling more advanced information retrieval and analysis (Zhu et al., 2017).

2.3 Relation Extraction Techniques

Rule-Based Approaches characterized early relationship extraction systems, which used predefined patterns or syntactic rules to identify relationships. For example, patterns like "X dismissed Y" could help identify a relationship between entities "X" and "Y" (Diaz-Garcia and Lopez, 2025).

Supervised Machine Learning trains classifiers, like Support Vector Machines (SVMs) or decision trees, using labeled datasets to learn relationship patterns. The features of such labeled datasets can be part-of-speech tags, dependency parse trees, and word embeddings. The limitations of supervised learning models are the need for extensive labeled data; however, there is still no guarantee that they will generalize well to unseen relationships.

Deep Learning and Neural Network Approaches use neural network models which process text sequentially or in parallel. These models, e.g., recurrent neural networks (RNNs), convolutional neural networks (CNNs), and transformers, improve relationship extraction accuracy and adaptability by capturing complex patterns and long-range dependencies. Transformer-based models try to understand the surrounding context and use contextual embeddings to enhance relationship extraction.

Distant Supervision is used to align text with existing structured data (such as knowledge graphs), based on the assumption that co-occurring entities in a text, which are also present in the knowledge base with a known relationship, are likely related. This method is suitable for producing large quantities of training data, but it can also generate noise due to incorrect assumptions. Dependency parsing identifies the syntactic structure of sentences and thereby helps detect relationships between entities, while end-to-end neural models combine entity recognition and relation extraction in a single process to reduce error propagation and improve efficiency.

Our approach focuses on extracting binary relations within a single sentence (sentence-level RE), similar to most other RE systems. Since a large part of our corpus consists of literary texts (more than one third), our relation inventory was adapted to the characteristics of this genre. In fiction, relations such as `spouseOf`, `siblingOf`, `parentOf`, `residesIn`, `relativeOf`, and `partOfOrg`, occur frequently, while business-oriented relations (e.g., `shareholdersOf`, `competitor`, `clientOf`) are comparatively rare.

3 Methodology

From the 32 relations defined in Watson NLP RE, we selected and adapted 19 relations relevant for various domains represented in our dataset. The difference is that we distinguish `managerOf` between persons and `managerOfOrg` between person and organization, and we introduced new relations: `eventPers` and `eventLoc`. The proposed methodology combines two complementary approaches: a rule-based approach and an LLM-based approach. For each relation, a list of lexical triggers and phrases is defined and used both for constructing FST patterns and for generating prompts in the LLM-based method.

The set of relations is predefined, and possible subject–object combinations are specified:

- PERS–PERS: `colleagueOf`, `managerOf`, `parentOf`, `relativeOf`, `siblingOf`, `marriedTo`, `spouseOf`
- PERS–ORG: `educatedAt`, `managerOfOrg`, `affiliatedWith`
- PERS–LOC: `residesIn`, `bornAt`, `diedAt`
- PERS–WORK: `authorOf`
- ORG–ORG: `partOfOrg`
- ORG–LOC: `basedIn`
- LOC–LOC: `partOfLoc`
- EVENT–PERS: `eventPers`
- EVENT–LOC: `eventLoc`

3.1 Dataset

The approach was applied and evaluated on a subset of 23,000 sentences extracted from TeslaNER+. The subset contains the following distribution of annotated entities: LOC (27K), PERS (25K), ROLE (17K), ORG (10K), DEMO (6K), WORK (2K), EVENT (1K), and PRODUCT (1K).

The input to the system consists of text seg-

mented into sentences and enriched with named entity annotations, as illustrated with the sentence 'After all, Goethe married Christiane Vulpius, not Madame von Stein.':

```
<s file_id="burmag-zorgoj.tsv" num="1756">Uostalom,
i <ner class="PERS" qid="Q5879">Gete</ner> se oženio
<ner class="PERS" qid="Q66746">Kristijanom Vulpijus
</ner>, a ne <ner class="ROLE" qid="Q1378024">
gospođom</ner> Fon <ner class="PERS" qid="None">
Štajn</ner>."</s>
```

Each sentence is annotated with entity spans, their types (e.g., PERS, ROLE), and, where available, links to external knowledge base Wikidata (QIDs) (Vrandečić and Krötzsch, 2014). The relation extraction module processes such annotated sentences by identifying candidate pairs of entities and applying predefined patterns (rules) or prompts to determine whether a semantic relation holds between them.

In this example, the system identifies a relation of type `marriedTo` between two entities of type PERS: *Gete* and *Kristijanom Vulpijus*. The relation is triggered by the lexical pattern *se oženio* ("has married"), which serves as a strong indicator of a marital relationship in Serbian. Although additional entities are present in the sentence (e.g., *gospođom* and *Štajn*), the system correctly filters out irrelevant pairs based on entity types and contextual constraints.

The resulting output is represented in a structured format, where the detected relation is explicitly encoded as follows:

```
<nel type="marriedTo">
<ner class="PERS" qid="Q5879">Gete</ner>
<ner class="PERS" qid="Q66746">Kristijanom Vulpijus
</ner> </nel>
```

This example demonstrates how the combination of entity annotation, lexical triggers, and pattern-based or prompt-based processing enables the extraction of semantically meaningful relations from complex sentence structures.

3.2 Rule-based Approach

The FST-based approach to relation extraction relies on the Unitex corpus processing system and is grounded in morphological dictionaries and local grammars (Krstev, 2008; Vitas and Krstev, 2012). Texts are first processed using rich lexical resources that provide morphological and syntactic information, enabling precise recognition of lexical units and their inflected forms. RE is then performed through cascades of FSTs, implemented as local grammars, which encode patterns corresponding to specific semantic relations. These grammars capture lexical triggers, syntactic configurations, and contextual constraints to identify related entity pairs within sentences. The cascade architecture allows the sequential application of increasingly complex patterns and recognition of various relations in the same sentence, allowing an NE to participate in several relations. This rule-based approach is particularly effective in handling morphologically rich languages such as Serbian, although it requires substantial human knowledge and manual effort for grammar design and maintenance.

When designing FSTs for relation extraction, we experimented with two approaches applied to several relations: a relaxed one that relies only on annotated NEs and a set of triggers and a strict one that takes into account numerous syntactic variations (e.g. word order), and possible modifications of NEs and triggers. The obtained results showed that the first approach leads to low precision, and the second one to low recall. Therefore, we decided to produce FSTs that are relaxed with some additional constraints. The Figure 1 presents one such FST designed for `spouseOf` relation extraction. It basically relies on a set of triggers, e.g., *žena*, *supruga* ("wife", "spouse") with an additional constraint that two recognized PERS entities are of the opposite sex (in this case the first is female and the second is male), which corresponds to the traditional understanding of persons united in marriage. Intermediate nodes (`notPers` in presented FST) allow for optional tokens such as punctuation, intervening words, or NEs of other types, capturing variation in natural language expressions.

The Figure 2 presents a FST for the extraction of the `bornAt` relation that links a PERS entity and a LOC entity. This FST also relies on triggers which can be verbs (*poticati* 'come from', *doći na svet* 'lit. come to the world'), nouns (*rodno mesto* 'birthplace'), or noun phrases (*poreklom iz* 'originating from'), but it also takes into account different word order (PERS coming before LOC and vice versa). Intermediate nodes (`notPersLoc` in presented FST) allow for optional tokens such as punctuation, intervening words, or NEs of other types, while nodes `Date` allow the insertion of complex dates.

3.3 LLM-based Approach

Wadhwa et al. (2023) revisit the evaluation of generative relation extraction by using human assessment instead of exact matching. Their results show that few-shot prompting with GPT-3 achieves near

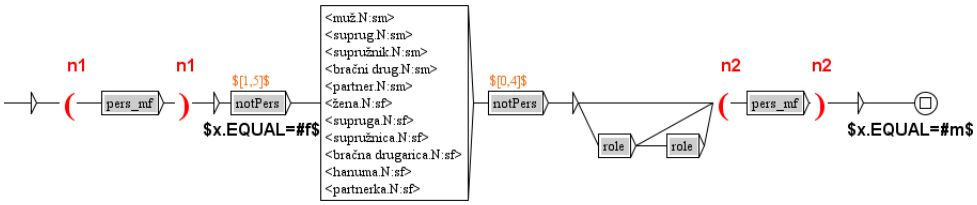

Figure 1: A simplified FST for the extraction of `spouseOf` relation (output omitted)

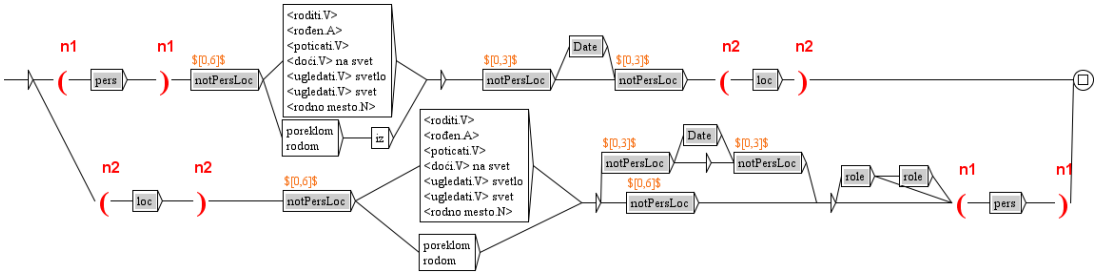

Figure 2: A simplified FST for the extraction of `bornAt` relation (output omitted)

state-of-the-art performance, comparable to fully supervised models, while Flan-T5 performs weaker in few-shot settings but reaches state-of-the-art results when fine-tuned with Chain-of-Thought explanations generated by GPT-3.

The LLM-based approach formulates relation extraction as an instruction-following task, where an LLM is prompted to identify and classify semantic relations between pairs of pre-annotated entities in text. Instead of relying on explicitly trained classifiers, the model leverages its pretrained knowledge and contextual understanding to infer relations directly from the input sentence or passage.

For each sentence, candidate pairs of entities are generated, and a structured prompt is constructed. The prompt includes: (i) the input passage, (ii) a list of recognized entities with their types, (iii) a predefined set of allowed relation labels, and (iv) task-specific instructions. The model has to decide whether a relation exists between a given entity pair and, if so, assign exactly one label from the predefined set.

To ensure consistency and facilitate automatic evaluation, the model is constrained to produce output in a fixed JSON format containing the head and tail entity identifiers, the predicted relation, a short textual evidence span extracted from the passage, and a confidence score. Lexical triggers associated with each relation are optionally included in the prompt to guide the model, although the final decision must rely on contextual evidence rather than surface patterns alone.

The prompt for the classification of the candidate pair is (it is translated in English for the sake of readability):

```
PAIR_CLASSIFICATION_PROMPT = """
You are a relation classification system for Serbian
literary text.

Task:
Decide whether there is a semantic relation
between the two target entities in the given
passage.
If yes, assign exactly one relation label
from the allowed set.
If no relation is clearly supported,
return "NO_RELATION".

Allowed relation labels:
{allowed_relations}

Typical Serbian lexical triggers
(hints only, not proof):
{trigger_block}

Rules:
- Use only the evidence from the sentence.
- Prefer the most specific relation.
- Use only the given entities.
- Return only valid JSON.
- If the relation is not clearly supported,
  return NO_RELATION.
- Evidence must be a short exact quote
  from the passage.
- Confidence must be a number between 0 and 1.

Return JSON in exactly this format:
{{
  "head": "{head_id}",
  "tail": "{tail_id}",
  "relation": "LABEL",
  "evidence": "short exact quote from the text
  or empty string",
  "confidence": 0.0
}}

Passage:
{passage}

Entities:
{entities_json}

Target pair:
- Head: {head_id} ({head_text}, {head_type})
- Tail: {tail_id} ({tail_text}, {tail_type})
"""
```

This approach allows flexible adaptation to different domains and relation schemas without additional model training. However, its performance

depends on prompt design, the clarity of the relation set, and the model's ability to handle ambiguity and complex sentence structures, which are particularly common in literary texts.

4 Statistics

Our rule-based system extracted a total of 1,676 relations from the dataset, covering a diverse set of semantic relation types, with an uneven distribution of relations (see Figure 3). Overall, the distribution reveals a dominance of location- and organization-related relations, while more fine-grained interpersonal and event-based relations occur less frequently. More precisely, the number of identified relations involving the entity ORG is 1,106 for five different relations, that is, 66.0% of relations retrieved. Similarly, the number of identified relations involving the entity LOC is 973 for six relations (58.1%). On the contrary, not many PERS–PERS relations were retrieved – 261 occurrences for seven relations (15.6%). The relatively low frequency of certain relations may indicate limitations in either lexical coverage or pattern detection.

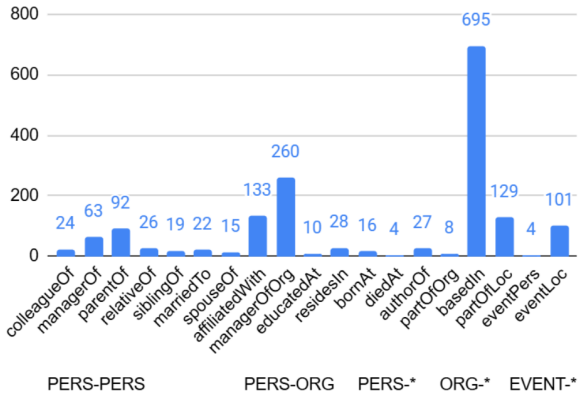

Figure 3: Results of RE by the rule-based system

We used the GPT-4.1-mini API for the prompt-based approach on a subset of 6,600 sentences, rather than the full dataset, due to slow performance. The total cost for this processing was $97. It extracted 24,607 relation candidates, of which 13,907 have a confidence score (CF, given by LLM) ≥ 0.9, and 14,139 have ≥ 0.85.

The productivity of relations was very different. Figure 4 show statistics, where for very productive (`colleagueOf`, `affiliatedWith`, `partOfOrg`, `basedIn`, `partOfLoc`), the threshold for CF $\geq$ 0.85. For evaluation we selected 2,629 relation candidates. The largest portion is dominated by LOC–LOC relations (`partOfLoc`, 8,642 instances CF ≥ 0.85, 12,616 total), followed by ORG–LOC (`basedIn`, 2,034 CF ≥ 0.85, 3,783 total) and ORG–ORG (`partOfOrg`, 1,022 CF ≥ 0.85, 1,541 total), indicating a strong emphasis on spatial and organizational hierarchies.

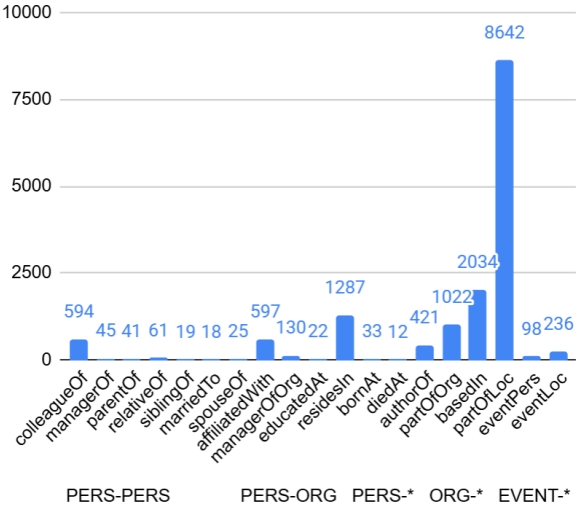

Figure 4: Results of RE by the LLM-based system

In contrast, other PERS–PERS relations are less frequent: `colleagueOf` with 594 candidates has CF ≥ 0.85(2,411 total), while more specific relations such as `marriedTo`, `siblingOf`, and `relativeOf` are sparsely represented. Similarly, PERS–ORG relations (722) are mainly driven by `affiliatedWith`, 597 candidates have CF ≥ 0.85 (1,532 total), whereas `educatedAt` is underrepresented.

PERS–LOC relations (1,332) are also skewed toward `residesIn` having 1,287 candidates, with `bornAt` and `diedAt` appearing only marginally. Among PERS–WORK relations, `authorOf` is also well represented (421), while EVENT-related relations remain relatively limited (98 + 236).

5 Evaluation

Sentences from the dataset presented in Section 3.1 and annotated using the rule-based system were transformed into so-called WebAnno TSV 3.2 File format.[3] Figure 5 presents examples of relation annotation in the INCEpTION environment (Klie et al., 2018) of several sentences from literary texts.

[3] https://webanno.github.io/webanno/releases/3.4.5/docs/user-guide.html#sect_webannotsv

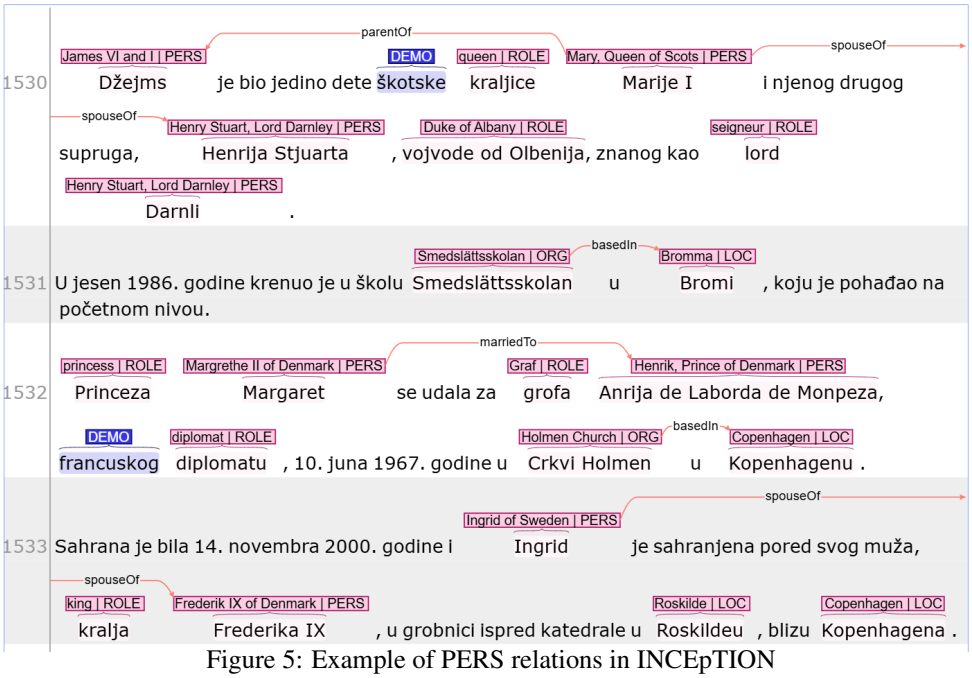

Figure 5: Example of PERS relations in INCEpTION

Named entities are annotated as PERS and LOC, and semantic relations are marked between relevant pairs. In the sentence 1530, the entities *Džejms* and *Marije I* are linked by the relation `parentOf`, triggered by the lexical cue *je bio dete* ("was child"), and the same PERS *Marija I* is linked to *Henrija Stjuarta* by the `spouseOf`. In the example 1531, ORG *Smedslättsskolan* is `basedIn` LOC *Bromma*. In the 1352 sentence, the entities *Margaret* and *Anri de Labord de Monpez* are linked by the relation `marriedTo`, indicated by the phrase *se udala* ("has married"). These examples illustrate how different types of interpersonal relations are identified through lexical triggers and contextual cues within narrative text.

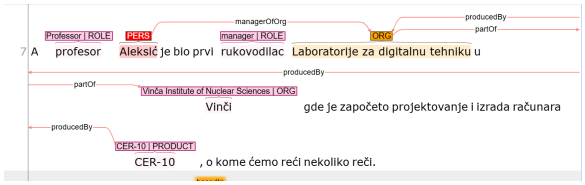

Figure 6: Example of ORG relations in INCEpTION

Illustrative examples of relations between entities of different types are presented in Figure 6. *Aleksić* is annotated as PERS, *profesor* "professor" and *rukovodilac* "manager" as ROLE, and *Laboratorije za digitalnu tehniku* "Lab for digital technology" as ORG. The relation `managerOfOrg` links the role *rukovodilac* to the organization, while additional relations such as `partOf` capture structural associations between organizations. The relation `producedBy` links an ORG entity with a PRODUCT entity.

Figure 7 presents the evaluation of extracted relations using FST-based approach, showing the high precision for most relations (blue). The evaluation results for relations extracted using LLM-based approach are presented in Figure 8. The LLM approach shows more deviations than the FST approach when extracting different relations. The results are approximate, as they are based on an evaluated sample of 2,629 candidates, rather than the full dataset.

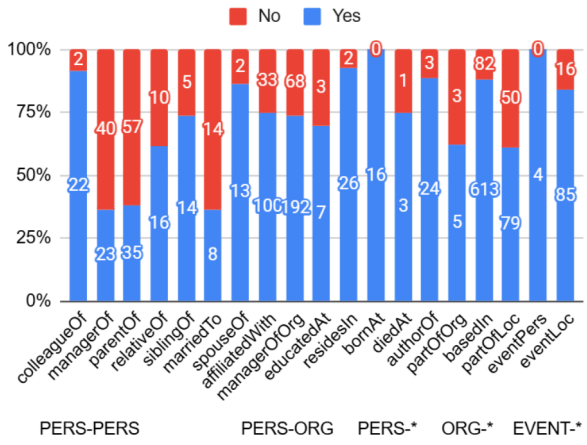

Figure 7: Evaluation of FST-based relation extraction

The precision of extracted relations across different relation types is presented in Figure 9. Over-

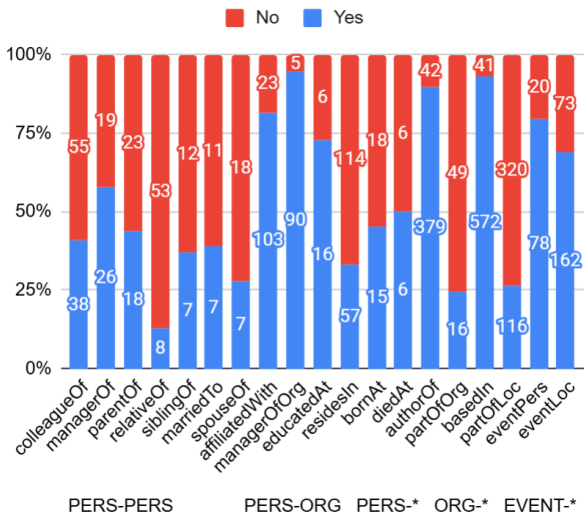

Figure 8: Evaluation of LLM-based relation extraction

all, performance is high, with overall precision $P = 0.77$. The best precision ($P = 1$) is observed for `eventPers` and `bornAt`, but not many occurrences were extracted. The most productive is `basedIn` (695) with $P = 0.88$. The best precision for a relation with a significant number of candidates was achieved for `residesIn` with $P = 0.93$. In contrast, `managerOf`, `parentOf`, and `marriedTo` show lower precision, indicating greater difficulty in capturing these from texts. The worst results were for `marriedTo` $P = 0.36$, where the problem might be the existence of a similar relation `spouseOf`, which was difficult to distinguish both for models and for evaluators.

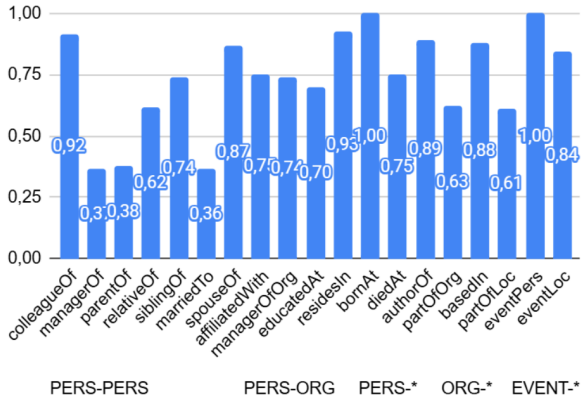

Figure 9: Precision of FST-based extraction

These results suggest that the FST model performs best on structurally explicit and less ambiguous relations, while more context-dependent social relations remain more challenging.

Figure 10 presents the precision of LLM-based relation extraction. Among PERS–PERS relations, moderate performance is observed, with `managerOf` and `parentOf` achieving the highest scores, while `relativeOf` remains notably lower. In the PERS–ORG group, `affiliatedWith` shows the highest score ($P = 0.95$), followed by `managerOfOrg` and `educatedAt`, indicating that institutional relations are more reliably detected.

For PERS–* relations, `authorOf` stands out with a high score, while relations such as `bornAt` and `residesIn` achieve moderate results. Within the ORG–* category, `basedIn` performs strongly, whereas `partOfOrg` shows lower values. Finally, EVENT–* relations, including `eventPers` and `eventLoc`, demonstrate relatively good performance. Overall, the results suggest that more structured and frequent relations are recognized more accurately, while less frequent or more complex relations are more prone to errors.

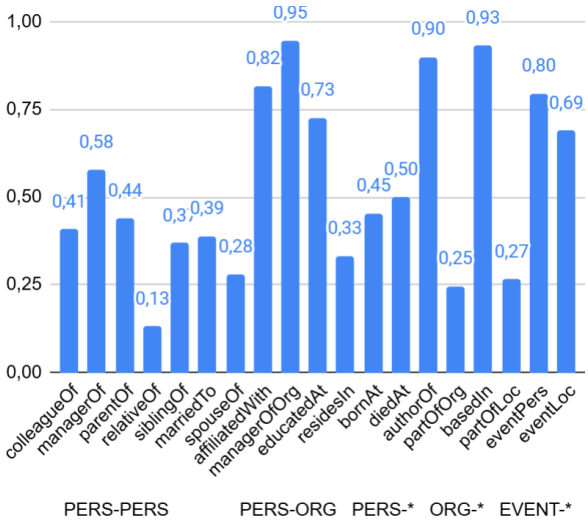

Figure 10: Precision of LLM-based extraction

Out of 2,638 evaluated instances, the LLM approach yielded 1,721 correct (Yes) and 917 incorrect (No) predictions. An additional 56 correct relations were identified by the FST approach which were not extracted by the LLM. In total, the FST system produced 1,285 correct (Yes) and 391 incorrect (No) predictions. For a fair comparison, recall was approximated on the same subset of sentences used for the LLM evaluation (6,600 sentences), where the FST system retrieved 796 candidates, of which 697 were correct and 99 incorrect. The results indicate that the FST approach achieves a precision of 0.77 and approximate recall

of 0.39 whereas LLM model yields much higher recall (0.97) at the cost of reduced precision (0.65).

6 Discussion

RE faces several well-known challenges, including ambiguity and polysemy, complex sentence structures, domain-specific language and limited generalization across domains, as well as the scarcity of labeled data. Additionally, texts often contain multiple overlapping relationships among several entities simultaneously, further increasing the difficulty of the task. We also observed these challenges in our dataset, particularly in literary texts, where nuanced language use and intricate narrative structures make relation identification and classification more demanding.

A key drawback of the LLM prompt-based approach is its relatively slow processing speed, coupled with a dependence on paid services, which may limit scalability and practical deployment.

LLM overgenerates, which can be illustrated with the sentence *Konferenciju je otvorio. . . , Sonja Liht ispred Fonda za otvoreno društvo, Meri Blek ispred Unicefa i Džim Stivens ispred Svetske banke.* ("The conference was opened by. . . , Sonja Licht on behalf of the Open Society Foundation, Mary Black on behalf of UNICEF, and Jim Stevens on behalf of the World Bank.") It generated 5 relations `affiliatedWith` instead of 3 (marked Yes):

- *Sonja Liht - Fonda za otvoreno društvo* (Yes)
- *Sonja Liht - Unicefa* (No)
- *Meri Blek - Unicefa* (Yes)
- *Meri Blek - Svetske banke* (No)
- *Džim Stivens - Svetske banke* (Yes)

The LLM exhibits difficulties with very long sentences, as processing times become excessively long and the model tends to generate nearly all possible combinations of entities. Training a specialized transformer-based model on the prepared dataset is expected to combine the strengths of both the rule-based and LLM-based approaches, leveraging the high precision of the former and the broader coverage and contextual flexibility of the latter. The dataset produced in this work enables such training. We expect that this integrated approach will lead to more robust and scalable relation extraction, particularly in complex and domain-specific texts such as literary corpora.

7 Conclusion

The presented results are promising and prove that use of FST and LLM-based approaches have their benefits and that the best performance may be achieved when used both. Future research will explore a range of applications of relation extraction, particularly its role in knowledge graph construction and enrichment. Additional directions include its integration into question answering systems, as well as its use in improving information retrieval and text summarization. Further potential lies in applying relation extraction to sentiment and opinion analysis, social network analysis, and domain-specific areas.

In future work, we plan to investigate how strong negative sampling, span filtering, and localized context representations can enable efficient search over all possible spans in an input sentence, as demonstrated by Eberts and Ulges (2020), thus making comprehensive relation extraction more feasible.

Acknowledgments

This research was supported by the Science Fund of the Republic of Serbia, #7276, Text Embeddings–Serbian Language Applications– TESLA and based upon work from COST Action CA23147 GOBLIN-Global Network on Large-Scale, Cross-domain and Multilingual Open Knowledge Graphs, supported by COST (European Cooperation in Science and Technology, https://www.cost.eu)

References

Zi-Yuan Chen, Chih-Hung Chang, Yi-Pei Chen, Jijnasa Nayak, and Lun-Wei Ku. 2019. UHop: An Unrestricted-Hop Relation Extraction Framework for Knowledge-Based Question Answering. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 345–356, Minneapolis, Minnesota. Association for Computational Linguistics.

Jose A Diaz-Garcia and Julio Amador Diaz Lopez. 2025. A Survey on Cutting-edge Relation Extraction Techniques Based on Language Models. *Artificial Intelligence Review*, 58(9):287.

Markus Eberts and Adrian Ulges. 2020. Span-Based Joint Entity and Relation Extraction with Transformer Pre-Training. In *ECAI 2020*, pages 2006–2013. IOS Press.

Jan-Christoph Klie, Michael Bugert, Beto Boullosa, Richard Eckart De Castilho, and Iryna Gurevych.

2018. The INCEpTION Platform: Machine-assisted and Knowledge-oriented Interactive Annotation. In *Proceedings of the 27th international conference on computational linguistics: system demonstrations*, pages 5–9.

Cvetana Krstev. 2008. *Processing of Serbian – Automata, Texts and Electronic Dictionaries*. University of Belgrade, Faculty of Philology, Belgrade.

Cvetana Krstev, Jelena Jaćimović, and Duško Vitas. 2012. Recognition and Normalization of Some Classes of Named Entities in Serbian. In *Proceedings of the Fifth Balkan Conference in Informatics (BCI '12)*, pages 52–57, New York, NY, USA. ACM.

Cvetana Krstev, Ivan Obradović, Miloš Utvić, and Duško Vitas. 2014. A System for Named Entity Recognition Based on Local Grammars. *Journal of Logic and Computation*, 24(2):473–489.

Cvetana Krstev, Anđelka Zečević, Duško Vitas, and Tita Kyriacopoulou. 2016. NERosetta for the Named Entity Multi-lingual Space. In *Human Language Technology. Challenges for Computer Science and Linguistics: 6th Language and Technology Conference, LTC 2013, Poznań, Poland, December 7-9, 2013. Revised Selected Papers*, pages 327–340, Cham. Springer International Publishing.

Milica Ikonić Nešić, Saša Petalinkar, Ranka Stanković, and Ruslan Mitkov. 2025. From Zero to Hero: Building Serbian NER from Rules to LLMs. In *Proceedings of the First Workshop on Comparative Performance Evaluation: From Rules to Language Models*, pages 87–96.

Milica Ikonić Nešić, Saša Petalinkar, Ranka Stanković, Miloš Utvić, and Olivera Kitanović. 2024. SrpCNNeL: Serbian Model for Named Entity Linking. In *2024 19th Conference on Computer Science and Intelligence Systems (FedCSIS)*, pages 465–473. IEEE.

Ranka Stanković, Milica Ikonić Nešić, Olja Perisic, Mihailo Škorić, and Olivera Kitanović. 2024. Towards Semantic Interoperability: Parallel Corpora as Linked Data Incorporating Named Entity Linking. In *Proceedings of the 9th Workshop on Linked Data in Linguistics@ LREC-COLING 2024*, pages 115–125.

Duško Vitas and Cvetana Krstev. 2012. Processing of Corpora of Serbian Using Electronic Dictionaries. *Prace Filologiczne*, LXIII:279–292.

Duško Vitas, Cvetana Krstev, and Denis Maurel. 2009. A Note on the Semantic and Morphological Properties of Proper Names in the Prolex Project. In Satoshi Sekine and Elisabete Ranchhod, editors, *Named Entities - Recognition, Classification and Use*, volume 19 of *Benjamins Current Topics*, pages 117–136. John Benjamins Publishing Company.

Denny Vrandečić and Markus Krötzsch. 2014. Wikidata: a Free Collaborative Knowledgebase. *Communications of the ACM*, 57(10):78–85.

Somin Wadhwa, Silvio Amir, and Byron C. Wallace. 2023. Revisiting Relation Extraction in the era of Large Language Models. *Proceedings of the conference. Association for Computational Linguistics. Meeting*, 2023:15566–15589.

Yongjun Zhu, Erjia Yan, and Il-Yeol Song. 2017. A Natural Language Interface to a Graph-based Bibliographic Information Retrieval System. *Data & Knowledge Engineering*, 111:73–89.