# Identification of Participants of Narratives Using Knowledge Bases

**Juliana Machado**[1,2], **Evelin Amorim**[1]

[1] Laboratory of Artificial Intelligence and Decision Support (LIAAD)
Institute for Systems and Computer Engineering, Technology and Science (INESC TEC)
Porto, Portugal

[2]Department of Automation and Systems Engineering (DAS)
Federal University of Santa Catarina (UFSC)
Florianópolis, Brazil

julianamachado99@gmail.com, evelin.f.amorim@inesctec.pt

***Abstract.*** *Identifying participants in narratives is important to understand and extract meaning from unstructured texts. This paper investigates the use of DBpedia and Wikifier for this task. We tested these two knowledge base platforms to evaluate their performance in recognizing and extracting entities in Portuguese-language journalistic narrative texts. The results show that both DBpedia and Wikifier present similar results in identifying participants, around 0.40 in the f1-score. The objective of this paper is to study the potential of knowledge bases to improve the understanding of narratives, in addition to suggesting directions for future research in this domain.*

## 1. Introduction

Experts usually define the narrative structure as a sequence of events related to each other and arranged chronologically [Santana et al. 2023]. Commonly, the narrative structure also includes participants that engage in the existing events. This arrangement of information facilitates the grasp of knowledge by people. Thus, the narrative structure is employed in several text types, from medical records [Moharasan and Ho 2019] to financial reports [Zmandar et al. 2021]. The automatic extraction of such a structure can aid in summarizing information and facilitating the selection of relevant information. Narrative extraction from text often involves a pipeline of models, with each model identifying one of the main components of a narrative (events, participants, and time expressions) and their relationships. In this work, we focus on the participant identification task.

The identification of participants in a narrative is similar to the task of named entities, nonetheless, a participant can be not only a named entity but also an expression, an object, or other contextual elements [Santos et al. 2023]. For instance, in the sentence "The dog ran until he was tired", "dog" is not a named entity, but it is a participant in the event "ran". To evaluate the difficulty of this task, we propose to employ pre-trained entity linking models to identify participants in narrative text. The entity linking task is the process of matching a mention of a name with an entry in a knowledge base [Sevgili et al. 2022]. Two well-known knowledge bases (KB) are DBPedia[1] and Wikipedia[2]. Several approaches for entity linking employ Wikipedia and DB-Pedia, one of them is the Wikifier [Brank et al. 2017] and another is the DBpedia Spotlight [Daiber et al. 2013]. None of these approaches, as far as we know, employed these tools for entity linking in the identification of participants of narrative text.

In this work, we evaluate entity-linking tools for their effectiveness in identifying participants within narrative texts. In particular, we employ Wikifier and DBpedia Spotlight to identify the participants from the Text2Story Lusa dataset [Nunes et al. 2024]. This dataset comprises Portuguese news stories, with all narrative components manually annotated by linguists.

The main contribution of this work is a quantitative evaluation of multi-lingual entity linking tools in the identification of participants in narrative text. We hope that these contributions aid researchers in selecting more appropriate tools and advance the research of the identification of participants in narratives.

## 2. Related Work

Regarding the identification of participants, the study of [Santos et al. 2023] focuses on identifying them in narrative stories, such as novels and soap operas in the Portuguese language. These stories are from the DIP (Character Identification Challenge) dataset, which was created based on data from the Digital Library of Literature in Portuguese Language (BDLP). This dataset includes stories added by the authors themselves, the vast majority of which are in Brazilian Portuguese. Additionally, the authors aimed to include heterogeneity in the dataset, featuring both Brazilian and Portuguese stories. In our study, however, we wish to identify participants in journalistic narratives, and our dataset is news from Lusa: Notícias do dia, Portugal's news agency, which makes our results more focused on European Portuguese.

Several studies on entity linking using knowledge bases, a task employed to identify the participants of a text, include the works of [Wu et al. 2018], [Jia et al. 2021] and [Xia et al. 2020]. For example, [Wu et al. 2018] investigates entity linking in various datasets of short texts, including those extracted from Twitter (a social media platform) and Wikipedia, among others. Similar to our study, they use knowledge bases to link entities, some of these most common knowledge bases are DBpedia, Wikipedia, and YAGO. DBpedia and Wikipedia will be used in our study.

The works by [Jia et al. 2021] and [Xia et al. 2020], as well as [Wu et al. 2018], also explore the use of entity linking, but this time in datasets containing data from various news sources, similar to our study. These datasets include data from news agencies such as the New York Times, Reuters, and the Associated Press. [Xia et al. 2020], in addition to using this news story data in English, also uses a dataset with news and micro-blogs in Chinese. Furthermore, [Xia et al. 2020] highlights an important issue in current research: most studies are conducted in English, leaving other languages, such as Chinese and Portuguese, underexplored. The authors of that study focused on Chinese, while our paper focuses on Portuguese.

In all these studies, the authors highlight several important challenges of entity

[1] `https://www.dbpedia.org/`
[2] `https://www.wikipedia.org/`

linking, such as the ambiguity and variability of entity mentions, and the large volume of data they must use to perform this task accurately. These challenges still need to be overcome and were encountered in this study as well.

## 3. Methodology

In this section, we describe the methodology employed in our study, which is part of the Text2Story project. We will first provide an overview of the Text2Story project, followed by a detailed explanation of the tools and technologies used, the source of the test data, and the metrics for evaluating the results.

### 3.1. Text2Story Project

The Text2Story project, as described by [Amorim et al. 2024], aims to facilitate the automatic extraction of narratives by identifying events, participants, temporal aspects, and their relationships. This enhances the semantic understanding and usability of textual data, transforming it into structured narratives. The project provides a toolkit integrating several natural language processing (NLP) techniques to process and visualize narrative texts. This makes it a valuable resource for researchers and developers working with narrative data.

Our study, conducted within the scope of the Text2Story project, focuses on evaluating entity-linking tools for extracting participants from narratives. Specifically, we employed Wikifier [Brank et al. 2017] and DBpedia Spotlight [Daiber et al. 2013], tools that link entities to their corresponding entries in knowledge bases such as Wikipedia and DBpedia. In this way, it was possible to evaluate the performance of these tools in real scenarios, using data from Text2Story Lusa dataset [Nunes et al. 2024], with Portuguese news stories from Lusa[3] (an important news agency in Portugal), thus being able to identify possible improvements for future implementations.

### 3.2. Tools Used

The two entity-linking tools employed in this study were Wikifier and DBpedia Spotlight. As mentioned previously, Wikifier [Brank et al. 2017] is a tool that automatically identifies and links entities mentioned in a text to their corresponding concepts on Wikipedia. The authors created it with advanced algorithms that analyze the context of mentions and associate them with Wikipedia articles. It is useful for disambiguation entities with multiple meanings, providing a good semantic understanding of the narrative text. This tool was chosen for its strong performance and the extensive database provided by Wikipedia.

Similarly, DBpedia Spotlight [Daiber et al. 2013]is an open-source tool that links mentions in a text to resources in the DBpedia knowledge base, which is a structured multilingual resource extracted from Wikipedia. According to the authors, DBpedia Spotlight brings an efficient and accurate approach to entity linking, taking advantage of DBpedia's structured data to provide a good semantic analysis of texts. Just like Wikifier, this tool was chosen to be studied and implemented in Text2Story because it demonstrated good results and also because it has a large structured knowledge base and therefore can extract information more precisely. In addition to that, both tools support the Portuguese language, which is the target language of our dataset.

[3]https://www.lusa.pt/

**Table 1. Dataset statistics**

| Narrative Component | Lusa News | |
|---|---|---|
| | Train | Test |
| Participants | 622 | 2,644 |
| #token | 3,707 | 16,805 |
| #documents | 20 | 90 |

The two tools were integrated into the Text2Story project to extract participants from journalistic narratives. We conducted quantitative tests using the Text2Story Lusa dataset to compare their performance with other tools implemented in the project. More details about the dataset will be presented in the next subsection.

### 3.3. Test Data

The dataset used in this study comes from the Portuguese news Agency Lusa, one of the main news agencies in Portugal. The "Text2Story Lusa" [Nunes et al. 2024] dataset was developed for analyzing narratives in journalistic articles written in European Portuguese. It includes a diverse range of news stories on topics such as politics, economics, and culture, providing a rich resource for evaluating NLP tools.

The dataset contains 117 annotated texts, with annotations that cover participants, events, and temporal aspects of the narratives, separating each word or phrase, based on the context of the narrative into a specific entity. These annotations were made by linguistic experts and are structured to facilitate the evaluation of tasks such as entity linking and participant extraction, which is the focus of our study.

For this study, we used only the test data, which includes 90 articles, to allow for a comprehensive comparison with [Amorim et al. 2024] study, where several evaluations of different tools were carried out, which will be presented in the results in the next section. Table 1 details some characteristics of the split we used.

### 3.4. Evaluation Metrics

To evaluate the effectiveness of innovative entity-linking tools in this project for participant extraction, three standard specifications were used: precision, recall, and F1-score. This information provides good guidance on how the tools are performing. Precision measures the proportion of correctly linked entities among those identified by the tool, recall measures the proportion of correctly linked entities among all relevant entities in the dataset, and F1-score provides a harmonic mean of precision and recall, which provides a more accurate assessment of the tool in question.

Following [UzZaman et al. 2013], we applied these metrics in two ways: strict and relaxed. In the strict evaluation, all narrative elements must be labeled by the automatic annotation tool exactly as the human annotation was done. In the relaxed evaluation, a partial overlap between human and automatic annotation is considered a true positive. For example, if the automatic annotator detected "automobilista (motorist)" and the human annotator annotated "um automobilista (a motorist)", it would be a false negative in the strict evaluation but a true positive in the relaxed evaluation. Thus, the relaxed evaluation is useful because it highlights partial correspondences, facilitating the identification of relevant narrative elements, therefore this was the evaluation used in this study.

## 4. Results

The Table 2 presents the results of the entity linking tools used in this project, compared with the results of other tools also implemented in the Text2Story project and presented in the [Amorim et al. 2024] study. The metrics presented are precision ($P_r$), recall ($R_r$), and F1-score ($F_{1_r}$), all applied to the task of extracting participants from journalistic narratives using the Text2Story Lusa dataset.

**Table 2. Results for the Annotators implemented in this study and the results of Annotators implemented in the [Amorim et al. 2024] study**

| | $P_r$ | $R_r$ | $F_{1_r}$ |
|---|---|---|---|
| Wikifier | 0.61 | 0.30 | 0.40 |
| DBpedia | 0.48 | 0.32 | 0.38 |
| SRL | 0.93 | 0.15 | 0.26 |
| SPACY | 0.77 | 0.33 | 0.45 |
| GPT-3 | 0.70 | 0.77 | 0.72 |

The results indicate that GPT-3, as described in the study by [Amorim et al. 2024], obtained the best overall performance in participant extraction, achieving excellent test results, with a precision of 0.70, a recall of 0.77 and an F1-score of 0.72. This superior performance is likely due to GPT-3's large language model, which is trained on vast amounts of diverse data, enabling it to understand and extract participants from complex narratives more effectively than traditional entity linking tools.

In contrast, Wikifier and DBpedia Spotlight, the tools implemented and evaluated in this study, showed good precision (0.61 and 0.48, respectively) but low recall (0.30 and 0.32, respectively). This suggests that while these tools can accurately identify participants when they are recognized, they struggle to detect all relevant entities. This limitation might stem from the coverage of the underlying knowledge bases (Wikipedia for Wikifier and DBpedia for Spotlight), which may not encompass the full range of entities or nuances present in the dataset.

Comparatively, the Semantic Role Labeling (SRL) and SPACY tools, also evaluated by [Amorim et al. 2024] exhibited strong precision but significantly lower recall, as can be seen in Table 2. This pattern suggests a broader challenge across various tools in fully detecting all relevant participants in journalistic narratives, possibly due to the inherent complexity and variability of the language used in news stories.

These results highlight the importance of a balance between precision and recall, and mainly the difficulty of tools detecting all participants. Although some tools are highly accurate, their usefulness is limited if they cannot identify a sufficient number of relevant entities (which can be seen from low recall). Tools like GPT-3 that can maintain a good balance between these metrics are preferable in participant extraction scenarios. Still, the first results from Wikifier and DBpedia are important, but they demonstrate that more studies are still needed on these tools so that they can improve.

These findings highlight a key issue: while high precision indicates accurate identification of known entities, low recall reveals a significant gap in detecting all relevant participants. This could be due to the limitations in the knowledge bases or the specific algorithms used by Wikifier and DBpedia Spotlight, which might not be as robust in handling the diverse and often ambiguous nature of participant references in journalistic texts. Still, the first results from Wikifier and DBpedia are important, and demonstrate that more studies are needed on these tools so that they can be improved.

## 5. Conclusion and discussion

In this article, we evaluate the effectiveness of the entity linking tools Wikifier and DBpedia Spotlight for participant extraction from journalistic narratives using the Text2Story Lusa dataset. The results demonstrated that, although both tools offer good precision, the recall results were low, which demonstrates that they still face challenges in detecting all relevant participants.

Wikifier and DBpedia Spotlight achieved good precision (0.61 and 0.48, respectively), but presented low recalls (0.30 and 0.32, respectively). In contrast, the GPT-3 model, as evaluated by [Amorim et al. 2024], achieved much better precision and recall results, reflected in an F1-score of 0.72. This disparity suggests that GPT-3's ability to leverage general knowledge and context is more effective in this task than the more specialized but narrower focus of Wikifier and DBpedia Spotlight.

Also analyzing the results of the other tools evaluated in the study by [Amorim et al. 2024], it is possible to see that the greatest difficulty of all tools is detecting all relevant participants, showing a general trend in the difficulty of identifying all relevant participants in journalistic narratives. This shows a need for improvements mainly in strategies to increase the recall of these tools.

Future research should explore strategies to improve recall, such as integrating more extensive or specialized knowledge bases, refining the algorithms for participant detection, or combining different approaches (e.g., LLMs with entity linking tools) to leverage the strengths of each. Additionally, expanding the evaluation to include other languages and types of narratives could provide a more comprehensive understanding of these tools' effectiveness across different contexts.

The limitations of this study, such as the exclusive focus on European Portuguese and the specific dataset used, also point to the need for broader evaluations. By addressing these challenges, future work can contribute to developing more robust and effective tools for participant extraction in diverse narrative contexts. We hope that this study will contribute to the selection and development of more suitable tools for extracting participants in journalistic narratives, encouraging future research and continuous improvements in natural language processing technologies.

## References

Amorim, E., Campos, R., Jorge, A., Mota, P., and Almeida, R. (2024). text2story: A python toolkit to extract and visualize story components of narrative text. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 15761–15772.

Brank, J., Leban, G., and Grobelnik, M. (2017). Annotating documents with relevant wikipedia concepts. *Proceedings of SiKDD*, 472.

Daiber, J., Jakob, M., Hokamp, C., and Mendes, P. N. (2013). Improving efficiency and accuracy in multilingual entity extraction. In *Proceedings of the 9th International Conference on Semantic Systems (I-Semantics)*.

Jia, N., Cheng, X., Su, S., and Ding, L. (2021). Cogcn: Combining co-attention with graph convolutional network for entity linking with knowledge graphs. *Expert Systems*, 38(1):e12606.

Moharasan, G. and Ho, T.-B. (2019). Extraction of temporal information from clinical narratives. *Journal of Healthcare Informatics Research*, 3:220–244.

Nunes, S., Jorge, A. M., Amorim, E., Sousa, H., Leal, A., Silvano, P. M., Cantante, I., and Campos, R. (2024). Text2story lusa: A dataset for narrative analysis in european portuguese news articles. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 15773–15782.

Santana, B., Campos, R., Amorim, E., Jorge, A., Silvano, P., and Nunes, S. (2023). A survey on narrative extraction from textual data. *Artificial Intelligence Review*, 56(8):8393–8435.

Santos, D., Mota, C., Pires, E., Langfeldt, M. C., Fuao, R. S., and Willrich, R. (2023). Dip-desafio de identificação de personagens: objectivo, organização, recursos e resultados. *Linguamática*, 15(1):3–30.

Sevgili, Ö., Shelmanov, A., Arkhipov, M., Panchenko, A., and Biemann, C. (2022). Neural entity linking: A survey of models based on deep learning. *Semantic Web*, 13(3):527–570.

UzZaman, N., Llorens, H., Derczynski, L., Allen, J., Verhagen, M., and Pustejovsky, J. (2013). Semeval-2013 task 1: Tempeval-3: Evaluating time expressions, events, and temporal relations. In *Second joint conference on lexical and computational semantics (* SEM), volume 2: Proceedings of the seventh international workshop on semantic evaluation (SemEval 2013)*, pages 1–9.

Wu, G., He, Y., and Hu, X. (2018). Entity linking: an issue to extract corresponding entity with knowledge base. *IEEE Access*, 6:6220–6231.

Xia, Y., Wang, X., Gu, L., Gao, Q., Jiao, J., and Wang, C. (2020). A collective entity linking algorithm with parallel computing on large-scale knowledge base. *The Journal of Supercomputing*, 76(2):948–963.

Zmandar, N., El-Haj, M., Rayson, P., Litvak, M., Giannakopoulos, G., Pittaras, N., et al. (2021). The financial narrative summarisation shared task fns 2021. In *Proceedings of the 3rd Financial Narrative Processing Workshop*, pages 120–125.