text2story: A Python Toolkit to Extract and Visualize Story Components of Narrative Text

Evelin Amorim[1] & Ricardo Campos [1,2] & Alípio Jorge[1,3] & Pedro Mota[3] & Rúben Almeida[1]

[1]INESC TEC
[2]University of Beira Interior
[3]University of Porto

April 30, 2024

INESCTEC

Section 1. Introduction

Some Concepts

Narrative definition

A sequence of events that are related to each other.

The narrative structure encompasses events, time expressions, and participants.

Narrative arrangement helps understanding . Thus, it is present in news, finance documents, books, etc.

Automatically extract narrative components can aid the summarization of data, then fastening the grasping of information.

Extraction of Narratives

1. Human annotation;
2. Data analysis;
3. Training models on the labeled data.

Cumbersome pipeline of tasks . Combination of different tools and models 😰

Section 2. Related Work

Existing tools

- Some programming libraries extract specific narrative components, but ignore the narrative structure (Jin et al., 2021; Wen et al., 2021; Zhang et al., 2022).
- Annotation tools ignore the narrative structure (Stenetorp et al., 2012; Montani and Honnibal, 2018) or disregard programming functionalities Horstmann (2020).

None of the existing tools comprises (1) programming functionalities, (2) batch experiments, and (3) visualization for the narrative structure.

text2story package contributions

🏁 To smooth the narrative extraction pipeline

1. it annotates text with off-the-shelf models, with an extensible annotator class;
2. it reads some well-known annotation file format files;
3. it automatizes batch experiments and their evaluations;
4. it produces three types of visual representation of annotation, namely, Message Sequence Chart (MSC), Knowledge Graphs (KG), and Bub- ble Diagrams (BD).

Section 3. The text2story package

text2story architecture

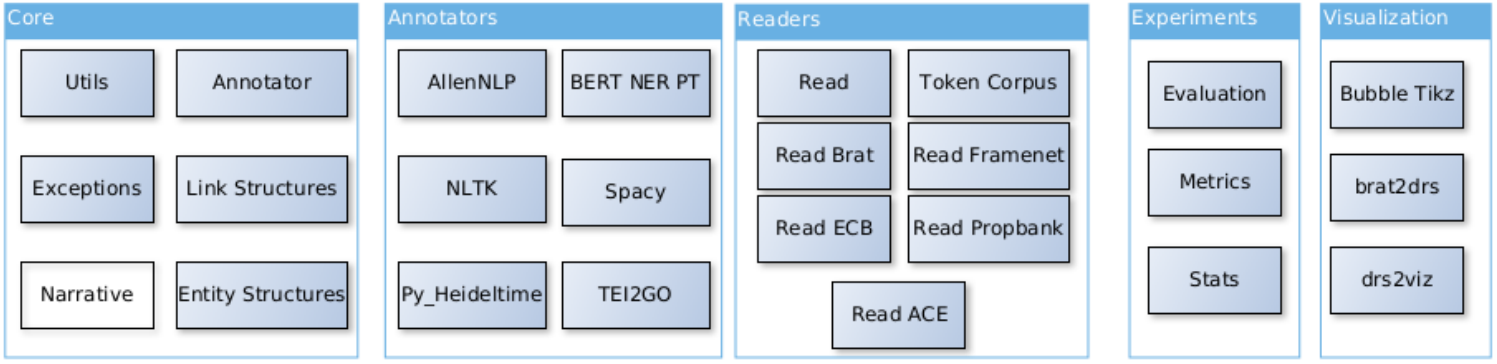


Figure: Main modules of text2story toolkit.

A pipeline for the narrative extraction - Part I

Example

Mrs Potter was Mrs Dursley's sister, but they hadn't met for several years; in fact, Mrs Dursley pretended she didn't have a sister, because her sister and her good-for-nothing husband were as unDursleyish as it was possible to be.

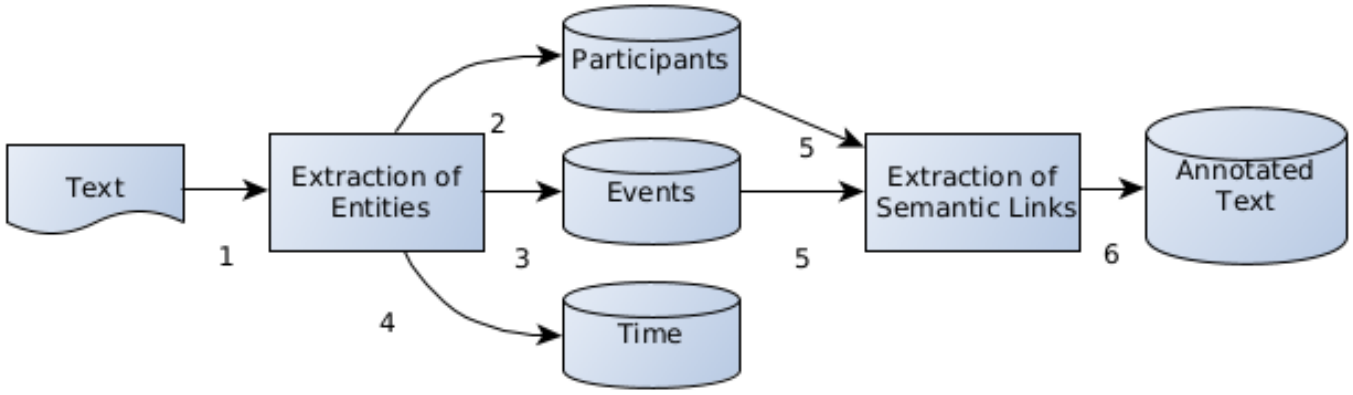


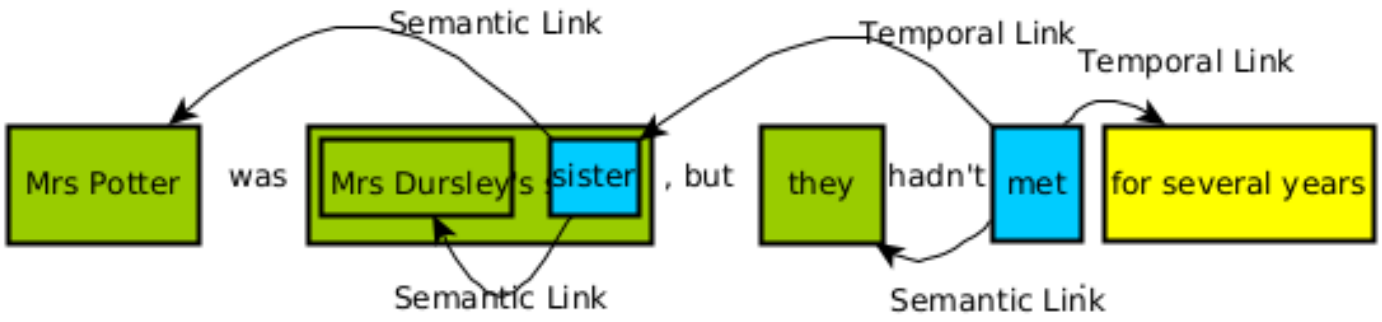
Figure: Pipeline flow

A pipeline for the narrative extraction - Part II

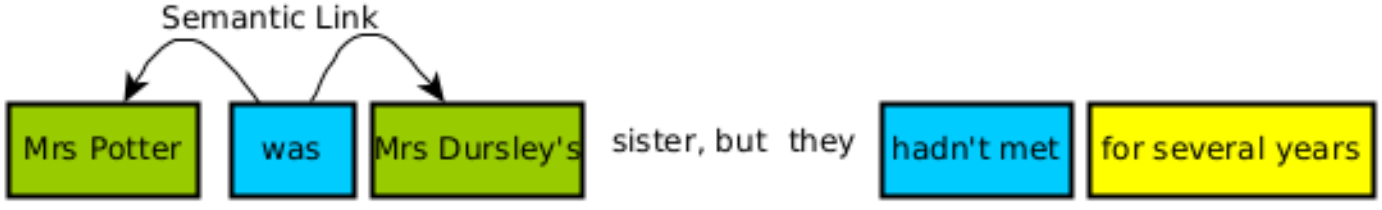

```
1 import text2story as t2s
2
3 narrative_doc = t2s.Narrative("en",doc, "
    2023") # this is the narrative object
4 participants = narrative_doc.
    extract_participants("spacy")
5 times = narrative_doc.extract_times("
    py_heideltime")
6 events = narrative_doc.extract_events("
    allennlp")
7 semanticrole_links = narrative_doc.
    extract_semantic_role_links()
```

Code Python for the Extraction of Narrative of Harry Potter's Example

Result of the Automatic Labeling



(a) Human labeling snippet text



(b) Automatic labeling snippet text

The Visualization of Annotations - MSC

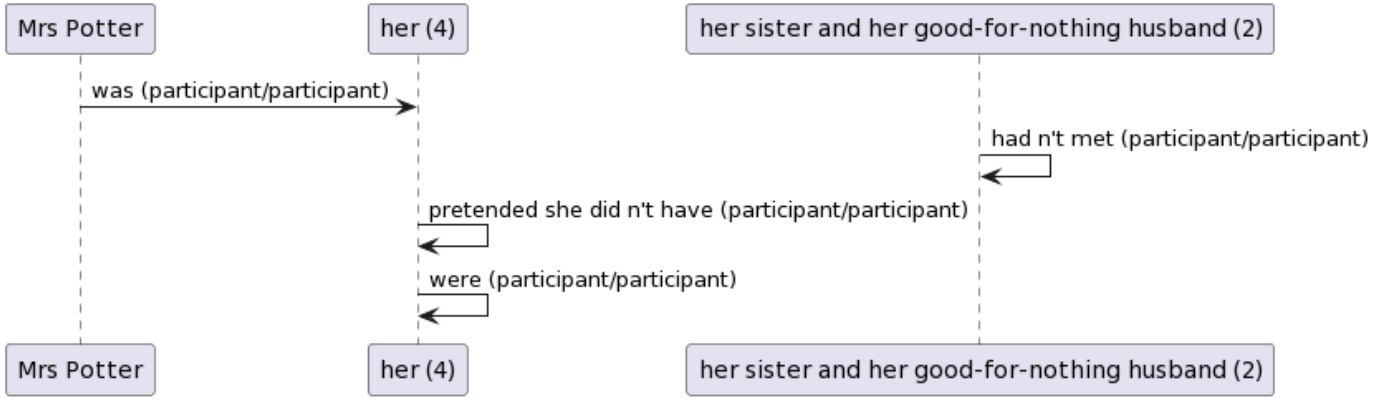


Figure: MSC representation built from the automatic labeling of a sentence of Harry Potter's book.

The Visualization of Annotations - KG

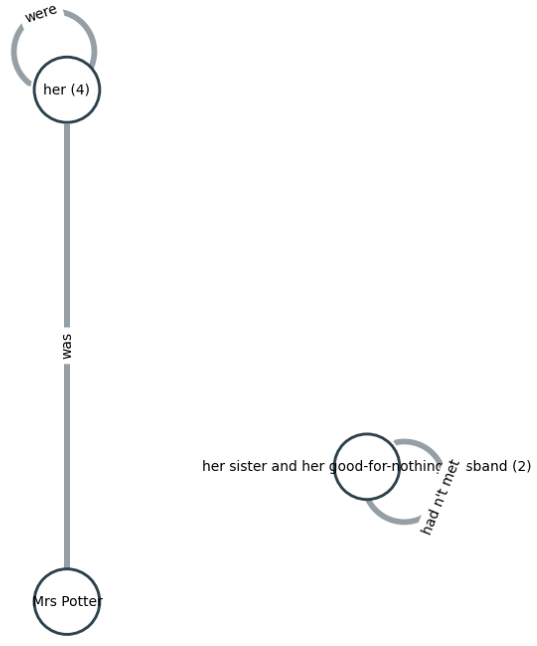


Figure: The Knowledge Graph Representation built from the automatic labeling of a sentence of Harry Potter's book.

Section 4. Experiments

Datasets

Narrative Component	ACE		Lusa News	
	Train	Test	Train	Test
Participants	5,948	37,071	622	2,644
Events	585	3,692	524	2,332
Times	670	3,700	67	338
#token	34,208	213,273	3,707	16,805
#documents	80	455	20	90

Table: Datasets statistics

Experimental Design

- A prompting strategy was developed on top of subset to text GPT-3 Sousa et al. (2023);
- The remaining pre-trained models were employed as zero-shot annotators;
- A relaxed version of f-1 metric was employed UzZaman et al. (2013): if there is intersection of words, then there is a match.

Results of Experiments - ACE

		P_r	R_r	$F1_r$
Time	TEI2GO	**0.75**	**0.60**	**0.64**
	Heideltime	0.68	0.53	0.57
	GPT-3	0.61	0.44	0.46
Participants	SRL	0.29	0.02	0.08
	SPACY	**0.76**	0.25	0.36
	GPT-3	0.68	**0.52**	**0.56**
Events	SRL	0.10	**0.45**	**0.15**
	GPT-3	**0.16**	0.079	0.08

Table: Results for the Annotators of text2story modules and GPT-3 in the ACE 2005 dataset

Results of Experiments - Lusa

		P_r	R_r	$F1_r$
Time	TEI2GO	0.70	**0.81**	**0.73**
	Heideltime	0.70	0.80	**0.73**
	GPT-3	**0.82**	0.52	0.61
Participants	SRL	**0.93**	0.15	0.26
	SPACY	0.77	0.33	0.45
	GPT-3	0.70	**0.77**	**0.72**
Events	SRL	**0.65**	0.37	**0.68**
	GPT-3	0.51	**0.71**	0.57

Table: Results for the Annotators of text2story modules and GPT-3 in the Lusa News dataset

Section 5. Conclusion

Conclusions and Limitations

- Text2story package streamlines the use of off-the-shelf libraries and offers an extensible framework for annotation, reading, and visualization.
- We present visual results that can aid the inspection of annotations and quantitative results for the narrative extraction task of pre-trained models;
- However, one limitation is that still only use BERT, when there are more modern models;
- Another limitation of the package is the lack of an interactive functionality for the visualizations results;
- Prompting is a technique subject to bias and to changes in LLMs.

Bibliography

Horstmann, J. (2020). Undogmatic literary annotation with catma. Annotations in Scholarly Editions and Research, pages 157–176.

Jin, Z., Chen, Y., Sui, D., Wang, C., Xue, Z., and Zhao, J. (2021). CogIE: An information extraction toolkit for bridging texts and CogNet. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: System Demonstrations, pages 92–98, Online. Association for Computational Linguistics.

Montani, I. and Honnibal, M. (2018). Prodigy: A new annotation tool for radically efficient machine teaching. Artificial Intelligence to appear.

Sousa, H., Guimarães, N., Jorge, A., and Campos, R. (2023). Gpt struct me: Probing gpt models on narrative entity extraction. In The 22nd IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology, WI = Artificial Intelligence in the Connected World.

Stenetorp, P., Pyysalo, S., Topić, G., Ohta, T., Ananiadou, S., and Tsujii, J. (2012). BRAT: a web-based tool for NLP-assisted text annotation. In Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics, pages 102–107.

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.

Wen, H., Lin, Y., Lai, T., Pan, X., Li, S., Lin, X., Zhou, B., Li, M., Wang, H., Zhang, H., Yu, X., Dong, A., Wang, Z., Fung, Y., Mishra, P., Lyu, Q., Surís, D., Chen, B., Brown, S. W., Palmer, M., Callison-Burch, C., Vondrick, C., Han, J., Roth, D., Chang, S.-F., and Ji, H. (2021). RESIN: A dockerized schema-guided cross-document cross-lingual cross-media information extraction and event tracking system. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human nologies: Demonstrations, pages 133–143, Online. Association for Linguistics.

Tao, L., Yu, H., Ye, H., Qiao, S., Xie, X., Chen, X., Li, Z., and Li, L. (2022). DeepKE: A deep learning based knowledge extraction toolkit for knowledge base population.