

# Report on the 4th International Workshop on Narrative Extraction from Texts (Text2Story 2021) at ECIR 2021

Ricardo Campos  
INESC TEC, Portugal  
Ci2 - Smart Cities Research Center,  
Polytechnic Inst. of Tomar, Portugal

Alípio M. Jorge  
INESC TEC,  
Portugal FCUP  
Univ. of Porto, Portugal

Adam Jatowt  
Univ. of Innsbruck  
Austria

Sumit Bhatia  
Media and Data Science Research Lab,  
Adobe, India  
*Sumit.Bhatia@adobe.com*

Mark Finlayson  
Florida International Univ.,  
USA

João Paulo Cordeiro  
INESC TEC, Portugal  
Univ. of Beira Interior, Portugal

Conceição Rocha  
INESC TEC, Portugal

Alexandre Ribeiro  
INESC TEC, Portugal

Behrooz Mansouri  
Rochester Inst. of Technology,  
Rochester, NY, USA

Jeffery Ansah  
BHP, Australia

Arian Pasquali  
CitizenLab, Belgium

## Abstract

The Fourth International Workshop on Narrative Extraction from Texts (Text2Story'21) was held on the April 1<sup>st</sup>, 2021, in conjunction with the 43<sup>rd</sup> European Conference on Information Retrieval (ECIR 2021). Due to the Covid-19 outbreak, the workshop was held online on Zoom platform. During the course of the day, an average of more than 80 attendees had the opportunity to follow-up and discuss the recent advances in extraction and formal representation of narratives. The workshop consisted of two invited keynotes and ten paper presentations. The proceedings of the workshop are available online at <http://ceur-ws.org/Vol-2860/>

**Date:** 1 April, 2021.

**Website:** <https://text2story21.inesctec.pt/>.



Figure 1: Overview of the Text2Story2021 Online Workshop

## 1 Introduction

The Text2Story series of Workshops started with the aims of bringing together researchers from diverse, related fields such as IR, NLP, AI, design and visualization to come together and share the recent advances in their respective fields towards narrative understanding. Building upon the success of the past editions [Jorge et al., 2018, 2019a; Campos et al., 2020] and on the Text2Story Special Issue at IPM Journal [Jorge et al., 2019b], this year, we organized the fourth edition of the Text2Story workshop, held online (see Figure 1) under the umbrella of ECIR 2021. This report summarizes the key activities at the workshop. We owe the workshop success to the diverse set of participants, from academia and industry, who established a new record of more than 80 attendees (on average) during all day and approximately 120 participants. The enthusiasm and active participation of the attendees, ensured that the virtual edition of the workshop was as intellectually stimulating as the past physical editions. The workshop program included a total of 10 research papers (out of 20 submissions) from 48 authors and two invited keynote talks. The videos<sup>1</sup> and presentation slides<sup>2</sup> of the talks at the workshop are available online.

<sup>1</sup>[https://www.youtube.com/playlist?list=PLUa5fBAPctfRuJh\\_ZyeZkQ5CI1Z\\_IVX6h](https://www.youtube.com/playlist?list=PLUa5fBAPctfRuJh_ZyeZkQ5CI1Z_IVX6h)

<sup>2</sup>[https://text2story21.inesctec.pt/slides/text2story21\\_workshop\\_slides.zip](https://text2story21.inesctec.pt/slides/text2story21_workshop_slides.zip)

---

## 2 The Program

The workshop program consisted of two invited keynote talks, six regular research papers and one demonstration. In addition, we also had work-in-progress, project description and nectar papers that allowed the respective authors to present their ongoing work and get useful feedback from the workshop attendees. The papers presented at the workshop covered diverse aspects of the narrative extraction problem ranging from tools and frameworks to enable annotate and visualize narratives in text to applications covering identifying disinformation about COVID-19 pandemic.

### 2.1 Keynotes

**José Alves** and **Célia Rodrigues** from *Jornal Público*<sup>3</sup>, a leading Portuguese newspaper, presented, with the help of various examples, how they use the art of visual storytelling to present information to the readers in a compelling and engaging way [Alves and Rodrigues, 2021]. The second keynote by **Leila Zia**, Head of Research at the Wikimedia Foundation<sup>4</sup>, was organized in the form of an interactive session where she described various streams of research at the Wikimedia foundation and how narrative science can help improve the quality of content in various Wikimedia projects. The questions asked by the attendees also covered a multitude of topics covering bias, dis-information, and availability of resources and datasets from the Wikimedia foundation [Zia, 2021].

### 2.2 Tools and Frameworks for Narrative Detection and Extraction

Amorim et al. [2021] presented Brat2Viz – a tool that displays the visualization of narrative information of an input text annotated in by the Brat annotation framework. The presentation of the narrative annotations in visual format helps the end-user easily grasp various narrative elements (such as actors and events) and semantic links between these elements. Alrashid and Gaizauskas [2021] presented learnings from their efforts in annotating *narrative scenes* in text using the SceneML framework. They found the task to be feasible and recommend several areas of improvement in designing task guidelines that can help improve inter-annotator agreement. Finally, Kroll et al. [2021] presented a demonstration of their system to automatically bind the narrative elements in text to external knowledge repositories such as Wikidata and PubMed.

### 2.3 Interpreting and Understanding Narratives

Gervás [2021] discussed the notion of embedding stories – stories told by a character in a story and argued for the need of computational models to represent such embedded stories within the main storyline. He then proposed a simplified model to account for the recursive structure of embedded stories in text and presented various directions of future work. Oza and Dietz [2021] presented an approach for identifying key entities in a story and found that entity co-occurrences in text to be a strong indicator of connection strength between the identified entities.

---

<sup>3</sup><https://www.publico.pt/>

<sup>4</sup><https://wikimediafoundation.org/>

---

## 2.4 Applications

Catipon and Savas [2021] presented a very timely study comparing the differences between conservative narratives on Twitter and Parler. Their analysis found that the media bias levels are higher for users on Parler than Twitter. They also noted that topics which are more moderated on Twitter gather very divergent views on Parler. Further, they also found that influential news sources are more politically varied on Parler. Hussain et al. [2021] presented a narrative visualization tool that can help an analyst discover the various themes and their associated narratives being discussed in various blogs. Rosa et al. [2021] presented **THEaiTRE 1.0**, a system based on GPT-2 to generate theatre play scripts.

## 2.5 Ongoing Research Efforts

The workshop also included special tracks for work-in-progress and position papers to offer a platform for authors to share initial results of their research efforts and gather useful feedback from the community. Marcoux and Agarwal [2021] presented their ongoing efforts on tracking spread of disinformation regarding COVID-19 across various online information sources. They also made findings of their study as well as misinformation stories identified are available to the public<sup>5</sup>. Vijayaraghavan and Vijayaraghavan [2021] presented their efforts towards modeling social media narratives about caste-related news stories. Mori et al. [2021] presented an approach for analysing incomplete stories and identifying positions of missing information. This is an important task as automated story generation methods can then generate missing parts of a story to produce a complete story.

## 3 Key Contributors

The Program Chairs were Ricardo Campos, Alípio Jorge, Adam Jatowt, Sumit Bhatia, and Mark Finlayson. The proceedings were setup by our Proceedings Chairs João Paulo Cordeiro and Conceição Rocha. Arian Pasquali, Behrooz Mansouri, Alexandre Ribeiro, and Jeffery Ansah were the web and dissemination chairs. In addition to this, we had six session chairs who brilliantly moderated the presentation of the papers. Our thanks go to Jannik Strötgen (Bosch Center for Artificial Intelligence, Germany), João Magalhães (Universidade NOVA de Lisboa), Bruno Martins (IST and INESC-ID - Instituto Superior Técnico, University of Lisbon) and Sérgio Nunes (INESC TEC & University of Porto). Finally, we would like to acknowledge the effort and valuable contribution of the researchers and industry experts that have served on the Program Committee of the Text2Story'21 workshop. Our thanks go to:

- Álvaro Figueira (INESC TEC & University of Porto)
- Andreas Spitz (Ecole polytechnique fédérale de Lausanne)
- António Horta Branco (University of Lisbon)
- Arian Pasquali (CitizenLab)
- Brenda Santana (Universidade Federal do Rio Grande do Sul)
- Bruno Martins (IST and INESC-ID – Instituto Superior Técnico, University of Lisbon)

---

<sup>5</sup><https://cosmos.ualr.edu/covid-19>

- 
- Daniel Gomes (FCT/Arquivo.pt)
  - Daniel Loureiro (University of Porto)
  - Denilson Barbosa (University of Alberta)
  - Deya Banisakher (Florida International University)
  - Dhruv Gupta (Norwegian University of Science and Technology (NTNU))
  - Dwaipayan Roy (ISI Kolkata)
  - Evelin Amorim (INESC TEC)
  - Florian Boudin (Université de Nantes)
  - Gaël Dias (Université de Caen Normandie)
  - Grigorios Tsoumakas (Aristotle University of Thessaloniki)
  - Henrique Lopes Cardoso (University of Porto)
  - Ismail Sengor Altingovde (Middle East Technical University)
  - Jeffery Ansah (BHP)
  - Jeremy Pickens (OpenText)
  - João Paulo Cordeiro (INESC TEC; Universidade da Beira do Interior)
  - Kiran Kumar Bandeli (Walmart Inc.)
  - Ludovic Moncla (INSA Lyon)
  - Marc Spaniol (Université de Caen Normandie)
  - Nina Tahmasebi (University of Gothenburg)
  - Nuno Moniz (INESC TEC & University of Porto)
  - Pablo Gamallo (University of Santiago de Compostela)
  - Paulo Quaresma (Universidade de Evora)
  - Preslav Nakov (Qatar Computing Research Institute (QCRI))
  - Ross Purves (University of Zurich)
  - Satya Almasian (Heidelberg University)
  - Sérgio Nunes (INESC TEC & University of Porto)
  - Udo Kruschwitz (University of Regensburg)
  - Yihong Zhang (Osaka University)
  - Yating Zhang (Alibaba)
  - Sebastião Miranda (Priberam)

## 4 Recognition Awards

The workshop ended with the announcement of the recognition awards. The Text2Story 2021 Best Paper Award went to *Regina Catipon* and *Onur Savas* for their paper entitled *A Different Story: How Conservative Narratives Diverge Between Twitter and Parler*. In addition to this, the workshop chairs have recognized the following researchers as recipients of the 2021 reviewer award for their insightful and valuable reviews - *Brenda Santana* (Universidade Federal do Rio Grande do Sul), *Bruno Martins* (IST and INESC-ID – Instituto Superior Técnico, University of Lisbon), *Deya Banisakher* (Florida International University).

---

## 5 Concluding Thoughts

This was the fourth edition of the Text2Story workshop series. We started with an objective to bring together the interested participants from different geographies and research expertise to collectively set the agenda for the emerging multi-disciplinary area of *narrative extraction from texts*. We are overwhelmed from the increasingly positive response from the research community and the growing participation in the workshops underlines the relevance of the topic. The fifth edition of the Text2Story workshop <sup>6</sup> will be held in conjunction with ECIR 2022 <sup>7</sup> and we invite the interested reader to be a part of the workshop.

## Acknowledgements

We would like to thank the ECIR 2021 organizers for helping us to organize the workshop, our wonderful keynote speakers, members of the program committee, the paper authors, and the participants who made the workshop a grand success.

## References

- Tarfah Alrashid and Robert J. Gaizauskas. A pilot study on annotating scenes in narrative text using sceneml. In Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Mark A. Finlayson, editors, *Proceedings of Text2Story - Fourth Workshop on Narrative Extraction From Texts held in conjunction with the 43rd European Conference on Information Retrieval (ECIR 2021), Lucca, Italy, April 1, 2021 (online event due to Covid-19 outbreak)*, volume 2860 of *CEUR Workshop Proceedings*, pages 7–14. CEUR-WS.org, 2021. URL <http://ceur-ws.org/Vol-2860/paper1.pdf>.
- José Alves and Célia Rodrigues. Storytelling through infographics. In Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Mark A. Finlayson, editors, *Proceedings of Text2Story - Fourth Workshop on Narrative Extraction From Texts held in conjunction with the 43rd European Conference on Information Retrieval (ECIR 2021), Lucca, Italy, April 1, 2021 (online event due to Covid-19 outbreak)*, volume 2860 of *CEUR Workshop Proceedings*, pages 1–2. CEUR-WS.org, 2021. URL <http://ceur-ws.org/Vol-2860/keynote1.pdf>.
- Evelin Amorim, Alexandre Ribeiro, Brenda Salenave Santana, Inês Cantante, Alípio Jorge, Sérgio Nunes, Purificação Silvano, Antonio Leal, and Ricardo Campos. Brat2viz: a tool and pipeline for visualizing narratives from annotated texts. In Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Mark A. Finlayson, editors, *Proceedings of Text2Story - Fourth Workshop on Narrative Extraction From Texts held in conjunction with the 43rd European Conference on Information Retrieval (ECIR 2021), Lucca, Italy, April 1, 2021 (online event due to Covid-19 outbreak)*, volume 2860 of *CEUR Workshop Proceedings*, pages 49–56. CEUR-WS.org, 2021. URL <http://ceur-ws.org/Vol-2860/paper6.pdf>.

---

<sup>6</sup><https://text2story22.inesctec.pt/>

<sup>7</sup><https://ecir2022.org/>

- 
- Ricardo Campos, Alípio Jorge, Adam Jatowt, and Sumit Bhatia. The 3<sup>rd</sup> international workshop on narrative extraction from texts: Text2story 2020. In Joemon M. Jose, Emine Yilmaz, João Magalhães, Pablo Castells, Nicola Ferro, Mário J. Silva, and Flávio Martins, editors, *Advances in Information Retrieval - 42nd European Conference on IR Research, ECIR 2020, Lisbon, Portugal, April 14-17, 2020, Proceedings, Part II*, volume 12036 of *Lecture Notes in Computer Science*, pages 648–653. Springer, 2020. doi: 10.1007/978-3-030-45442-5\\_86. URL [https://doi.org/10.1007/978-3-030-45442-5\\_86](https://doi.org/10.1007/978-3-030-45442-5_86).
- Regina Catipon and Onur Savas. A different story: How conservative narratives diverge between twitter and parler. In Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Mark A. Finlayson, editors, *Proceedings of Text2Story - Fourth Workshop on Narrative Extraction From Texts held in conjunction with the 43rd European Conference on Information Retrieval (ECIR 2021), Lucca, Italy, April 1, 2021 (online event due to Covid-19 outbreak)*, volume 2860 of *CEUR Workshop Proceedings*, pages 15–22. CEUR-WS.org, 2021. URL <http://ceur-ws.org/Vol-2860/paper2.pdf>.
- Pablo Gervás. A model of interpretation of embedded stories. In Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Mark A. Finlayson, editors, *Proceedings of Text2Story - Fourth Workshop on Narrative Extraction From Texts held in conjunction with the 43rd European Conference on Information Retrieval (ECIR 2021), Lucca, Italy, April 1, 2021 (online event due to Covid-19 outbreak)*, volume 2860 of *CEUR Workshop Proceedings*, pages 23–31. CEUR-WS.org, 2021. URL <http://ceur-ws.org/Vol-2860/paper3.pdf>.
- Muhammad Nihal Hussain, Hayder Al Rubaye, Kiran Kumar Bandeli, and Nitin Agarwal. Stories from blogs: Computational extraction and visualization of narratives. In Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Mark A. Finlayson, editors, *Proceedings of Text2Story - Fourth Workshop on Narrative Extraction From Texts held in conjunction with the 43rd European Conference on Information Retrieval (ECIR 2021), Lucca, Italy, April 1, 2021 (online event due to Covid-19 outbreak)*, volume 2860 of *CEUR Workshop Proceedings*, pages 33–40. CEUR-WS.org, 2021. URL <http://ceur-ws.org/Vol-2860/paper4.pdf>.
- Alípio M. Jorge, Ricardo Campos, Adam Jatowt, and Sérgio Nunes. First international workshop on narrative extraction from text (text2story’18). In Pasi G., Piwowarski B., Azzopardi L., Hanbury A. (eds). *Advances in Information Retrieval. ECIR 2018 (Grenoble, France. March 26 – 29). Lecture Notes in Computer Science, vol 10772*, pages 833–834, 2018.
- Alípio M. Jorge, Ricardo Campos, Adam Jatowt, and Sumit Bhatia. Second international workshop on narrative extraction from text (text2story’19). In Azzopardi L., Stein B., Fuhr N., Mayr P., Hauff C., Hiemstra D. (eds), *Advances in Information Retrieval. ECIR’19 (Cologne, Germany. April 14 – 18). Lecture Notes in Computer Science, vol 11438*, pages 389–393, 2019a.
- Alípio M. Jorge, Ricardo Campos, Adam Jatowt, and Sérgio Nunes. Special issue on narrative extraction from texts (text2story): Preface. In *Information Processing Management an International Journal. Elsevier, Vol 56(5)*, pages 1771–1774, 2019b. URL <https://www.sciencedirect.com/science/article/abs/pii/S0306457319304455>.

- 
- Hermann Kroll, Denis Nagel, Morris Kunz, and Wolf-Tilo Balke. Demonstrating narrative bindings: Linking discourses to knowledge repositories. In Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Mark A. Finlayson, editors, *Proceedings of Text2Story - Fourth Workshop on Narrative Extraction From Texts held in conjunction with the 43rd European Conference on Information Retrieval (ECIR 2021), Lucca, Italy, April 1, 2021 (online event due to Covid-19 outbreak)*, volume 2860 of *CEUR Workshop Proceedings*, pages 57–63. CEUR-WS.org, 2021. URL <http://ceur-ws.org/Vol-2860/paper7.pdf>.
- Thomas Marcoux and Nitin Agarwal. Narrative trends of COVID-19 misinformation. In Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Mark A. Finlayson, editors, *Proceedings of Text2Story - Fourth Workshop on Narrative Extraction From Texts held in conjunction with the 43rd European Conference on Information Retrieval (ECIR 2021), Lucca, Italy, April 1, 2021 (online event due to Covid-19 outbreak)*, volume 2860 of *CEUR Workshop Proceedings*, pages 77–80. CEUR-WS.org, 2021. URL <http://ceur-ws.org/Vol-2860/paper10.pdf>.
- Yusuke Mori, Hiroaki Yamane, Yusuke Mukuta, and Tatsuya Harada. The nectar of missing position prediction for story completion. In Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Mark A. Finlayson, editors, *Proceedings of Text2Story - Fourth Workshop on Narrative Extraction From Texts held in conjunction with the 43rd European Conference on Information Retrieval (ECIR 2021), Lucca, Italy, April 1, 2021 (online event due to Covid-19 outbreak)*, volume 2860 of *CEUR Workshop Proceedings*, pages 65–69. CEUR-WS.org, 2021. URL <http://ceur-ws.org/Vol-2860/paper8.pdf>.
- Pooja Oza and Laura Dietz. Which entities are relevant for the story? In Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Mark A. Finlayson, editors, *Proceedings of Text2Story - Fourth Workshop on Narrative Extraction From Texts held in conjunction with the 43rd European Conference on Information Retrieval (ECIR 2021), Lucca, Italy, April 1, 2021 (online event due to Covid-19 outbreak)*, volume 2860 of *CEUR Workshop Proceedings*, pages 41–48. CEUR-WS.org, 2021. URL <http://ceur-ws.org/Vol-2860/paper5.pdf>.
- Rudolf Rosa, Tomáš Musil, Ondrej Dusek, Dominik Jurko, Patrícia Schmidtová, David Marecek, Ondrej Bojar, Tom Kocmi, Daniel Hrbek, David Kosták, Martina Kinská, Marie Nováková, Josef Dolezal, Klára Vosecká, Tomáš Studeník, and Petr Zabka. Theatre 1.0: Interactive generation of theatre play scripts. In Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Mark A. Finlayson, editors, *Proceedings of Text2Story - Fourth Workshop on Narrative Extraction From Texts held in conjunction with the 43rd European Conference on Information Retrieval (ECIR 2021), Lucca, Italy, April 1, 2021 (online event due to Covid-19 outbreak)*, volume 2860 of *CEUR Workshop Proceedings*, pages 71–76. CEUR-WS.org, 2021. URL <http://ceur-ws.org/Vol-2860/paper9.pdf>.
- Prashanth Vijayaraghavan and Lavanya Vijayaraghavan. Modeling social media narratives about caste-related news stories. In Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Mark A. Finlayson, editors, *Proceedings of Text2Story - Fourth Workshop on Narrative Extraction From Texts held in conjunction with the 43rd European Conference on Information Retrieval (ECIR 2021), Lucca, Italy, April 1, 2021 (online event due to Covid-19 outbreak)*,



---

volume 2860 of *CEUR Workshop Proceedings*, pages 81–86. CEUR-WS.org, 2021. URL <http://ceur-ws.org/Vol-2860/paper11.pdf>.

Leila Zia. How wikimedia understands the text2narrative process. In Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Mark A. Finlayson, editors, *Proceedings of Text2Story - Fourth Workshop on Narrative Extraction From Texts held in conjunction with the 43rd European Conference on Information Retrieval (ECIR 2021), Lucca, Italy, April 1, 2021 (online event due to Covid-19 outbreak)*, volume 2860 of *CEUR Workshop Proceedings*, pages 3–6. CEUR-WS.org, 2021. URL <http://ceur-ws.org/Vol-2860/keynote2.pdf>.