Identification of Relations between Text Segments for Semantic Storytelling

Georg Rehm[1], Malte Ostendorff[1], Rémi Calizzano[1], Karolina Zaczynska[1] and Julián Moreno Schneider[1]

[1]DFKI GmbH, Berlin, Germany

Abstract

Semantic Storytelling is the (semi)-automatic generation of new content (storylines) based on information extracted from document collections presented with helpful visualisation techniques. This paper summarises our previous work and describes the Semantic Storytelling vision and technical approach. We describe experiments that focus on the identification of relations between text segments extracted from documents written by different authors; discourse relations have, so far, primarily been researched within single documents only. The results confirm our intuition that discourse parsing is difficult to apply to the inter-text level, but they are encouraging as a first step. Similarly, the identification of inter-text relations using pairwise document classification yields promising results. Lastly, we show the effectiveness of paragraph ordering for coherent story generation.

Keywords

Semantic Storytelling, Text Segment Classification, Textual Entailment, Paragraph Ordering

1. Introduction

With the ever-increasing amount of digital content, users face the challenge of coping with enormous quantities of information. This challenge is especially true for digital content curators, i. e., knowledge workers such as, among others, journalists [1], television producers [2], designers [3], librarians [4], or academics [5]. These and other professional profiles have in common that they monitor and process *existing* or *incoming content* to produce *new content.* In several projects [6, 5], we have been developing technical approaches to support knowledge workers in their day-to-day jobs, curating large amounts of mostly textual content more efficiently and effectively. One of our focus areas is the generation of storylines. This includes the (semi)automatic creation of new content as well as helpful presentation and visualisation techniques. We call the approach Semantic Storytelling [7, 8, 1, 9, 2, 10, 11].

Recently, storytelling has mostly been interpreted as a language generation task [see, e. g., 12, 13, 14, 15], where the goal is to generate texts. We interpret the concept differently by concentrating on the *extraction* and *presentation* of stories and their parts, contained in content streams, e. g., document collections or social media feeds. We see storylines as sets of building blocks, which, depending on their combination (temporal, geographical, causal etc.), can be

Qurator 2022: 3rd Conference on Digital Curation Technologies, September 19-23, 2022, Berlin, Germany
georg.rehm@dfki.de (G. Rehm); malte.ostendorff@dfki.de (M. Ostendorff); remi.calizzano@dfki.de (R. Calizzano); karolina.zaczynska@dfki.de (K. Zaczynska); julian.moreno_schneider@dfki.de (J. M. Schneider)

assembled into a story in various ways. Our goal is the recognition of atomic pieces of information (e. g., facts, propositions, events) in document collections and the identification of semantic relations between these pieces. Applications can be conceptualised, e. g., as information systems (for the retrieval of existing content) or recommender systems (for creating new content).

2. Operationalising Storytelling

We define Semantic Storytelling as the (semi-)automatic generation of storylines based on information extracted from documents or social media streams which are processed, classified, annotated and visualised, typically in an interactive way. This section describes the existing components and utilized previous work, followed by an architecture description conceptualized to meet the requirements of our Semantic Storytelling approach.

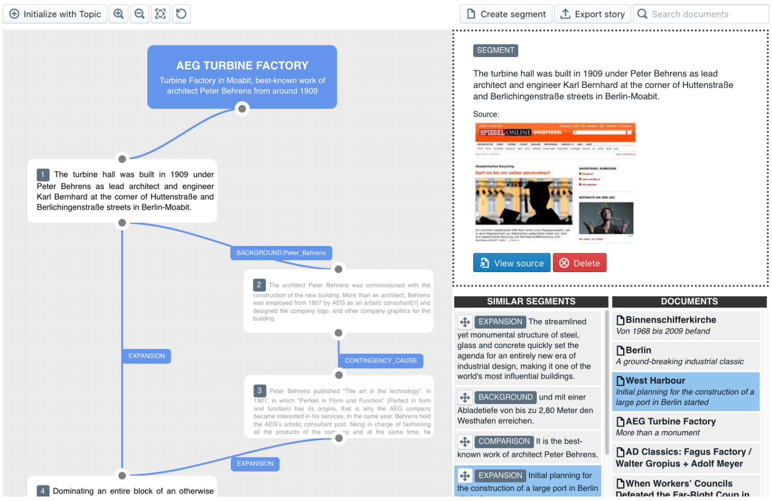

Figure 1: Prototypical UI of a story editor.

2.1. Semantic Storytelling: Existing Components

Three groups of text analytics components are the foundation of our architecture [see, e. g., 7, 16, 8, 2]: (1) services that analyse documents or document collections to provide document-level metadata, (2) services that extract, annotate and enrich specific parts of the incoming content, and (3) services that transform the content, e. g., via summarization or translation.

The tools are in different stages of technical maturity. They are orchestrated using a workflow manager [17]. Named entity recognition and linking as well as time expression analysis are performed to identify named entities of various types and classes. The integration of time expression analysis allows reasoning over temporal expressions and anchoring entities and events to a timeline. We use topic detection to assign abstract topics to individual sentences, paragraphs, chapters, and documents. For the annotations, we use NLP Interchange Format [NIF, 18], which allows the exploitation of the Linked Data paradigm and Linked Open Data resources. We distinguish between different classes or genres of documents, i. e., we experiment with different approaches for identifying document structures [10] and, on top of this, document genres [19]. An ontology to represent a heterogeneous set of document characteristics, tying together the different parts of annotations mentioned above, is currently under development.

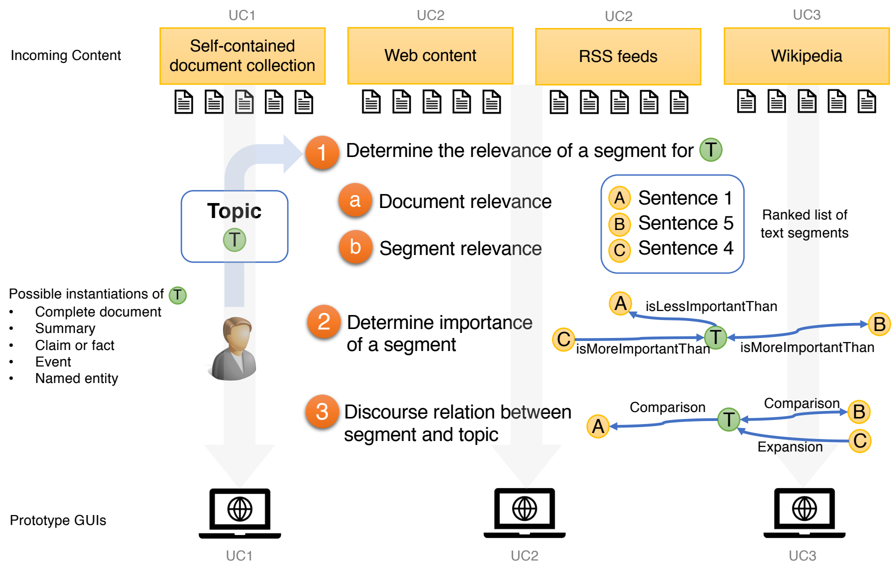

Figure 2: Semantic Storytelling architecture.

We implemented a number of experimental prototypes and user interfaces [7, 9, 8, 2]. On top of the semantic analysis of documents, we map the extracted information, whenever possible, to Linked Open Data and visualise the result [20, 3]. By providing feedback to the output of certain semantic services, content curators have control over the workflow. The storytelling UIs involve the dynamic and interactive recomposition and visualisation of extracted information. This involves arranging content elements (documents, paragraphs, sentences, events) on a dynamic timeline or as a graph. Figure 1 shows an example that visualises a story as a graph in which individual story units (content pieces) are represented as nodes. Different content types such as topics and text segments are displayed in different colors. Edges represent specific relations.

2.2. Architecture

At the core of the Semantic Storytelling approach are three processing steps. Figure 2 and the following paragraphs illustrate these steps in more detail.

2.2.1. Step 1: Determine the Relevance

We start with a topic *T*, instantiated through a text segment, e. g., a named entity, headline or document. To identify content pieces relevant for *T*, we process an incoming content stream and decide for each piece whether it is relevant for *T*. Relevance can be computed in various ways, e. g., text similarity measures, or we can compute the overlap in terms of named entities.

The accuracy depends on the length of the segments. For example, we can perform pairwise comparisons of document similarity starting with the seed document d_s of which we know that it represents topic T and measure its similarity to other candidate documents. Document pairs with a high similarity score are assumed to cover the same topic. If an incoming segment, e. g., a news article, is relevant to T or its seed document d_s, the next steps involve identifying the important atomic segments (i. e., sentences or paragraphs) and determining the relations that hold between these atomic segments and T.

2.2.2. Step 2: Determine the Importance

Given a document d related to T, we need to determine the importance of d (and its segments) with regard to T. Various cues and indicators can be exploited, e. g., an incoming news piece on T that was published only seconds ago and that includes the cue "BREAKING" in its title. One way of determining the topical importance of an individual segment is to treat it as a segment-level question answering task. Given a document d that consists of a sequence of segments $(t_1, t_2, \dots t_n)$, the aim is to find the segment t_i that contains the answer to the input question. In our context, the input would be the topic T instead of a question. A weighting schema could be applied such that, e. g., novel news pieces are preferred over old ones.

2.2.3. Step 3: Determine the Relation

Relations can exist between two segments on different levels. While we are primarily looking at discourse or coherence relations, relations can also relate to more simple levels such as segment order. The modeling of semantic, discourse, or coherence relations in textual content, where propositions, statements or events are the individual units, is at the core of discourse parsing frameworks. These analyse a text for, typically, *intra*-textual but not *inter*-sentential, relations. We borrow from discourse parsing and experiment with PDTB2 annotations [21] (see Section 3.1), but there are several added challenges. Crucially, our system needs to be able to robustly find (discourse) relations of short segments extracted from *different* texts while we have ample evidence from step 1 that the two texts are relevant to each other.

3. Identification of Relations between Text Segments

While steps 1 and 2 can be considered common NLP tasks, step 3 is less standardized with only little related work in the literature, which is why we elaborate on the identification of relations between text segments in more detail. We summarise three separately published experiments that represent different approaches to the task: PDTB2 discourse classification [22], Wikipedia article relations [23], and text segment ordering [24].

3.1. Experiment 1: PDTB2-based Discourse Relation Classification

The goal of this experiment [22] is to get preliminary results regarding the (ambitious) task of identifying discourse relations between two arbitrary text segments that are relevant to each other and that both relate to the same topic. The experiments are based on PDTB2 [21], which

Table 1
Results of DistilBERT for multi-class predictions of PDTB2 relations.

PDTB Relation	Precision	Recall	F1
Comparison	0.50	0.47	0.48
Contingency	0.38	0.65	0.48
Expansion	0.50	**0.79**	**0.61**
Temporal	**0.51**	0.55	0.53
None	0.49	0.73	0.59
Micro avg.	0.47	0.67	0.55

contains approx. 40,000 annotated relations. We establish a baseline training a classifier for the four top level PDTB2 senses (*Temporal*, *Contingency*, *Comparison*, *Expansion*) or *None* if none of the classes apply. We use DistilBERT to obtain contextual vector representations of the text segments [25]. For the relation classification, we use a Siamese architecture.

Table 1 shows the results of the classifier which performs best for *Expansion*, by far the most frequent class. The performance is lower than state-of-the-art approaches, but a comparison is not straightforward. First, our classification performs considerably lower than other approaches because we have not implemented features relating to the connective. Second, we only use the four PDTB2 top level classes. Most other approaches use a more fine-grained set, resulting in many more classes and lower performance. See for example [26], who report an F1-score of 40.70 for implicit relations. For our experiments, however, we argue that a coarse classification, with more training examples and higher accuracy, is better suited.

3.2. Experiment 2: Semantic Relations of Wikipedia Articles

Table 2
Results for Wikipedia relation classification as micro avg. F1-scores.

Methods	F1	Std.
Avg. GloVe [27]	0.875	± 0.0036
Paragraph Vectors [28]	0.845	± 0.0019
Siamese BERT [29, 30]	0.870	± 0.0067
Siamese XLNet [31, 30]	0.864	± 0.0096
BERT [29]	**0.933**	± 0.0039
XLNet [31]	0.926	± 0.0016

In this experiment [23], we transfer the relation classification from a sentence-level to a document-level and from PDTB relations to more generic ones. Given a seed document d_s, we are interested in finding a target document d_t that shares the semantic relation r_i with d_s. We model the task of finding the relation r of a document pair (d_s, d_t) as a pairwise multi-class document classification problem. Wikipedia articles are utilized as documents and Wikidata properties [32] define their semantic relations. For example, the Wikipedia article on *Albert Einstein* and its Wikidata item are connected to *German Empire* through the property *country of citizenship*. The Wikidata property acts as the relation between the article pair and the class label in the training data for this pair of documents. We selected nine relations ranging from *country of citizenship* to *facet of* to *opposite of*. Besides the number of available Wikipedia article pairs, diversity was also a criterion in our selection with respect to the different semantic meanings of properties. We evaluate six different methods: Document vectors from average GloVe word vectors [27], Paragraph Vectors [28], BERT [29], XLNet [31], and Siamese variations of BERT and XLNet [33, 30].

Section 2 presents the empirical results. BERT yields the best micro average F1-score with 0.933, followed by XLNet with 0.926 F1. The vanilla Transformers, BERT and XLNet, generally outperform their Siamese counterparts. The shared contextual information during the encoding of document pairs most likely yields the better performance for vanilla Transformers. Even abstract relations, like *facet of*, yield a considerable high F1-score (0.91 for BERT). Siamese BERT (0.870 F1) and Siamese XLNet (0.870 F1) are even outperformed by AvgGloVe (0.875 F1) despite AvgGloVe requiring only a fraction of the computing resources compared to the Transformer models. A qualitative evaluation and detailed analysis is presented in [23]. Our results suggest that pairwise classification is suitable for classifying semantic relations between documents. In another study [34], we confirm this finding also for the domain of research papers.

3.3. Experiment 3: Text Segment Ordering

In the last experiment, we study the relation of coherence between text segments [24]. When segments are related to the same topic, we want to determine which of the segments (sentences or paragraphs) should precede the other in order to maximise discourse coherence. In the best case, the predicted order would correspond to a complete and coherent story line although being composed of individual segments. As opposed to pairwise ordering (similar to pairwise classification, Section 3.2), we apply direct ordering of all segments using a Pointer network [35] combined with the pre-trained encoder-decoder model BART [36]. As baseline, we rely on Hierarchical Attention Networks (HAN) inspired by [35] that uses a Pointer network, Multi-Head Attention and LSTMs for sequence representations. We evaluate the BART-Pointer and the HAN baseline on two new paragraph ordering datasets tailored to Semantic Storytelling. As opposed to common segment ordering datasets using only single sentences as segments, for Semantic Storytelling, we are also interested in paragraphs with more than one sentence. Hence, we construct two new datasets for paragraph ordering based on the CNN DailyMail (CNN-DM) dataset and on Wikipedia.

Our model outperforms the HAN baseline on both datasets. With a Perfect Match Ratio (PMR) of 0.3699 for Wikipedia and 0.0171 for CNN-DM, the BART-Pointer combination is significantly better than the baseline, which yields 0.2100 PMR for Wikipedia and 0.0049 PMR for CNN-DM. An evaluation with Kendall's Tau metric (τ) shows that for 36.99% of the test samples our model is able to perfectly order the shuffled paragraphs from the introduction of Wikipedia articles while only 1.71% of the CNN-DM articles can be ordered perfectly. CNN-DM seems to be a greater challenge to the model by the number of paragraphs with an average of 14.5 sequences to order, whereas the Wikipedia dataset has an average of 6.29 paragraphs. We assume that the introductions in Wikipedia articles is often more consistent than those in CNN-DM, thus, the order is easier to learn. To sum up, we evaluate the BART-Pointer combination as suitable for ordering paragraphs to create a coherent text from an unordered collection of segments.

3.4. Discussion

The (semi)automatic identification and generation of storylines from text segments is still in its infancy. In this paper, we focus on one crucial step of our Semantic Storytelling approach, i. e., the identification of the relation between text segments. We approach the task from different angles, i.e., the notion of relation is defined as PDTB2 discourse relation, Wikidata properties, or order relation. The results of the PDTB2 experiments (Section 3.1) reveal a below state-of-art performance of the Siamese BERT approach. We attribute the low performance to the inability of the Siamese network to encode relational information of the two text segments.

This assumption is confirmed by the method evaluation of the second experiment (Section 3.2) carried out as a pairwise multi-class document classification task with more training data and on a variety of non-discourse relations. The methods from the experiment yield substantially higher accuracy scores compared to experiment 1. On a methodological level, vanilla Transformer models turn out to be more suitable for the relation classification task as their Siamese counterparts. We find that also presumable difficult relations like *facet of* achieve promising results that would be suitable for our use cases. However, the pairwise document classification approach has one drawback: the approach only classifies a single pair of segments while stories consists of multiple segments.

To address this, experiment 3 explores relation identification as a paragraph ordering task (Section 3.3). The order relation ensures a coherent story generation, i. e., more than two segments are arranged in a meaningful manner. In our experiment, we demonstrate that a Pointer network in combination with an encoder-decoder model like BART, is capable of not only ordering sentences but also text segments of paragraph length. Given the difficulty of this task, our results are very encouraging.

4. Related Work

This brief overview of related work refers to several areas including narratology, discourse theory, as well as applied work in computational linguistics and language technology.

Several approaches grounded in narratology address storytelling as a way of automatising the detection of instances of story grammars [37], especially events, in texts. Caselli and Vossen [38] present a data set for the detection of temporal and causal relations and use a plot structure [39] to order events found in narratives or text documents, chronologically and logically. According to Bal [39], narratives follow a plot structure that consists of ordered events, told by an agent or author and caused or experienced by actors. Yarlott and Finlayson [40] use Propp's (1968) morphology of Russian hero tales for story detection and generation systems. His book, first published in 1928, Propp analyzes the structural elements of Russian folk tales, which always occur in a fixed, consecutive order. Yan et al. [42] describe a system that learns "functional story schemas" as sets of functional structures (e. g., character introduction, conflict setup, etc.) in social media narratives. They extract patterns of functional structures. Afterwards, their formation in a story is analyzed across all stories to find schematic structures. Vice versa, Gordon et al. [43] use stories from blog articles to perform automated causal reasoning. Bois et al. [44] recommend articles based on simple lexical similarity. They link news articles in the form of a graph and label links to inform users on the nature of the relation between two news pieces. Ribeiro et al. [45] cluster news articles based on identified event instances and word alignment. They attempt to form clusters of online articles that deal with a certain event type. Nie et al. [46] use dependency parsing and discourse relations to determine sentence relations by learning vector representations. Yarlott et al. [47] apply the discourse theory by Dijk [48] to examine how paragraphs behave when used as discourse structure units in news articles.

5. Conclusions and Future Work

After laying the groundwork for Semantic Storytelling through various experiments, we are now approaching the final phase, in which we attempt to combine the components described in Section 2.1, including text analytics and enrichment services that operate on documents, document collections or text segments [10], with the emerging set of technologies described in Section 2. To this end, we identified the key building blocks for Semantic Storytelling technologies. While step 1 can be implemented using one of several known approaches, steps 2 and 3 are much more challenging (Section 2). Our approach is grounded in the assumption that *different texts* that deal with the *same* topic but that are from *different authors* and *different sources* can be interconnected in a meaningful way through relations, which we attempt to extract automatically. We want to support content curators and make use of these relations holding between two segments by exposing them explicitly and exploiting them in the construction of storylines in a semiautomatic or fully automatic way. While our experiments are promising [49, 24, 23] they also show that additional research is needed before we can integrate the technologies into prototypes. Data sets annotated for rhetorical or discourse structure are still rather limited both in availability and in size. Our future work will focus on expanding our setup, especially with regard to the analysis and classification of discourse relations and more sophisticated processing of connectives. We will integrate a more flexible approach with regard to the processing of single documents by concentrating on larger parts of a document including longer summaries and paraphrased variants to increase coverage. Taking into account explicit ontological knowledge to identify semantic relations between texts will also be an important next step towards the completion of the envisaged Semantic Storytelling prototype [10].

Acknowledgments

The work presented in this paper has received funding from the German Federal Ministry of Education and Research (BMBF) through the projects QURATOR (Wachstumskern no. 03WKDA1A) and PANQURA (no. 03COV03E).

References

[1] J. Moreno Schneider, A. Srivastava, P. Bourgonje, D. Wabnitz, G. Rehm, Semantic storytelling, cross-lingual event detection and other semantic services for a newsroom content curation dashboard, in: Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism, ACL, 2017, pp. 68–73.

[2] G. Rehm, J. Moreno Schneider, P. Bourgonje, A. Srivastava, R. Fricke, J. Thomsen, J. He, J. Quantz, A. Berger, L. König, S. Räuchle, J. Gerth, D. Wabnitz, Different Types of Automated and Semi-Automated Semantic Storytelling: Curation Technologies for Different Sectors, in: G. Rehm, T. Declerck (Eds.), Language Technologies for the Challenges of the Digital Age: 27th Int. Conf., GSCL 2017, Berlin, Germany, September 13-14, 2017, Proceedings, number 10713 in Lecture Notes in Artificial Intelligence (LNAI), Springer, 2018, pp. 232–247. 13/14 September 2017.

[3] G. Rehm, J. He, J. Moreno-Schneider, J. Nehring, J. Quantz, Designing User Interfaces for Curation Technologies, in: S. Yamamoto (Ed.), Human Interface and the Management of Information: Information, Knowledge and Interaction Design, 19th Int. Conf., HCI Int. 2017 (Vancouver, Canada), number 10273 in Lecture Notes in Computer Science (LNCS), Springer, 2017, pp. 388–406.

[4] C. Neudecker, G. Rehm, Digitale Kuratierungstechnologien für Bibliotheken, Zeitschrift für Bibliothekskultur 027.7 4 (2016).

[5] G. Rehm, M. Lee, J. M. Schneider, P. Bourgonje, Curation Technologies for a Cultural Heritage Archive: Analysing and transforming a heterogeneous data set into an interactive curation workbench, in: A. Antonacopoulos, M. Büchler (Eds.), DATeCH 2019: Proceedings of the 3rd Int. Conf. on Digital Access to Textual Cultural Heritage, Brussels, Belgium, 2019, pp. 117–122. 8-10 May 2019.

[6] G. Rehm, F. Sasaki, Digitale Kuratierungstechnologien – Verfahren für die effiziente Verarbeitung, Erstellung und Verteilung qualitativ hochwertiger Medieninhalte, in: Proceedings der Frühjahrstagung der Gesellschaft für Sprachtechnologie und Computerlinguistik (GSCL 2015), Duisburg, 2015, pp. 138–139.

[7] P. Bourgonje, J. Moreno Schneider, G. Rehm, F. Sasaki, Processing Document Collections to Automatically Extract Linked Data: Semantic Storytelling Technologies for Smart Curation Workflows, in: A. Gangemi, C. Gardent (Eds.), Proceedings of the 2nd Int. Workshop on Natural Language Generation and the Semantic Web (WebNLG 2016), The Assoc. for Computational Linguistics, Edinburgh, UK, 2016, pp. 13–16.

[8] J. Moreno Schneider, P. Bourgonje, J. Nehring, G. Rehm, F. Sasaki, A. Srivastava, Towards Semantic Story Telling with Digital Curation Technologies, in: L. Birnbaum, O. Popescu, C. Strapparava (Eds.), Proceedings of Natural Language Processing meets Journalism – IJCAI-16 Workshop (NLPMJ 2016), New York, 2016.

[9] G. Rehm, J. Moreno Schneider, P. Bourgonje, A. Srivastava, J. Nehring, A. Berger, L. König, S. Räuchle, J. Gerth, Event Detection and Semantic Storytelling: Generating a Travelogue from a large Collection of Personal Letters, in: T. Caselli, B. Miller, M. van Erp, P. Vossen, M. Palmer, E. Hovy, T. Mitamura (Eds.), Proceedings of the Events and Stories in the News Workshop, ACL, 2017, pp. 42–51.

[10] G. Rehm, K. Zaczynska, J. M. Schneider, Semantic Storytelling: Towards Identifying Storylines in Large Amounts of Text Content, in: A. Jorge, R. Campos, A. Jatowt, S. Bhatia (Eds.), Proceedings of Text2Story – Second Workshop on Narrative Extraction From Texts co-located with 41th European Conf. on Information Retrieval (ECIR 2019), Cologne, Germany, 2019, pp. 63–70. 14 April 2019.

[11] M. Raring, M. Ostendorff, G. Rehm, Semantic Relations between Text Segments for Semantic Storytelling: Annotation Tool – Dataset – Evaluation, in: N. Calzolari, F. Béchet,

P. Blache, C. Cieri, K. Choukri, T. Declerck, H. Isahara, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis (Eds.), Proceedings of the 13th Language Resources and Evaluation Conference (LREC 2022), European Language Resources Association (ELRA), Marseille, France, 2022, pp. 4923–4932. June 20-25, 2022.

[12] A. Fan, M. Lewis, Y. Dauphin, Hierarchical neural story generation, in: Proc. of the 56th Annual Meeting of the Assoc. for Computational Linguistics (Volume 1: Long Papers), 2018, pp. 889–898.

[13] A. Fan, M. Lewis, Y. Dauphin, Strategies for structuring story generation, arXiv preprint arXiv:1902.01109, 2019.

[14] P. Xu, M. Patwary, M. Shoeybi, R. Puri, P. Fung, A. Anandkumar, B. Catanzaro, MEGATRON-CNTRL: Controllable story generation with external knowledge using large-scale language models, in: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), Association for Computational Linguistics, Online, 2020, pp. 2831–2845. doi:10.18653/v1/2020.emnlp-main.226.

[15] X. Hua, A. Sreevatsa, L. Wang, DYPLOC: Dynamic planning of content using mixed language models for text generation, in: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Association for Computational Linguistics, Online, 2021, pp. 6408–6423. doi:10.18653/v1/2021.acl-long.501.

[16] P. Bourgonje, J. Moreno Schneider, J. Nehring, G. Rehm, F. Sasaki, A. Srivastava, Towards a Platform for Curation Technologies: Enriching Text Collections with a Semantic-Web Layer, in: H. Sack, G. Rizzo, N. Steinmetz, D. Mladenic, S. Auer, C. Lange (Eds.), The Semantic Web, number 9989 in Lecture Notes in Computer Science, Springer, 2016, pp. 65–68. ESWC 2016 Satellite Events. Heraklion, Crete, Greece, May 29 – June 2.

[17] J. Moreno-Schneider, P. Bourgonje, F. Kintzel, G. Rehm, A Workflow Manager for Complex NLP and Content Curation Pipelines, in: G. Rehm, K. Bontcheva, K. Choukri, J. Hajic, S. Piperidis, A. Vasiljevs (Eds.), Proceedings of the 1st Int. Workshop on Language Technology Platforms (IWLTP 2020, co-located with LREC 2020), Marseille, France, 2020, pp. 73–80. 16 May 2020.

[18] S. Hellmann, J. Lehmann, S. Auer, M. Brümmer, Integrating NLP using Linked Data, in: Proc. of the 12th Int. Semantic Web Conf., 2013. 21-25 October 2013.

[19] G. Rehm, Hypertextsorten: Definition – Struktur – Klassifikation, Books on Demand, Norderstedt, 2007. PhD thesis in Applied and Computational Linguistics, Justus-Liebig-Universität Giessen, 2005.

[20] J. Moreno Schneider, P. Bourgonje, G. Rehm, Towards User Interfaces for Semantic Storytelling, in: S. Yamamoto (Ed.), Human Interface and the Management of Information: Information, Knowledge and Interaction Design, 19th Int. Conf., HCI Int. 2017 (Vancouver, Canada), number 10274 in Lecture Notes in Computer Science (LNCS), Springer, Cham, Switzerland, 2017, pp. 403–421. Part II.

[21] R. Prasad, N. Dinesh, A. Lee, E. Miltsakaki, L. Robaldo, A. K. Joshi, B. L. Webber, The Penn Discourse TreeBank 2.0, in: Proc. of LREC 2008, ELRA, 2008.

[22] G. Rehm, K. Zaczynska, P. Bourgonje, M. Ostendorff, J. Moreno-Schneider, M. Berger, J. Rauenbusch, A. Schmidt, M. Wild, J. Böttger, J. Quantz, J. Thomsen, R. Fricke, Semantic Storytelling: From Experiments and Prototypes to a Technical Solution, in: T. Caselli,

E. Hovy, M. Palmer, P. Vossen (Eds.), Computational Analysis of Storylines: Making Sense of Events, Studies in Natural Language Processing, Cambridge University Press, Cambridge, 2021, pp. 240–259.

[23] M. Ostendorff, T. Ruas, M. Schubotz, G. Rehm, B. Gipp, Pairwise Multi-Class Document Classification for Semantic Relations between Wikipedia Articles, in: Proc. of the 2020 ACM/IEEE Joint Conf. on Digital Libraries (JCDL'20), 2020. arXiv:2003.09881.

[24] R. Calizzano, M. Ostendorff, G. Rehm, Ordering Sentences and Paragraphs with Pre-trained Encoder-Decoder Transformers and Pointer Ensembles, in: Proceedings of the 21st ACM Symposium on Document Engineering (DocEng 2021), Association for Computing Machinery, Limerick, Ireland, 2021, pp. 1–9.

[25] V. Sanh, L. Debut, J. Chaumond, T. Wolf, DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter, CoRR (2019) 1–5.

[26] Z. Dai, H. Taneja, R. Huang, Fine-grained structure-based news genre categorization, in: Proc.of the Workshop Events and Stories in the News 2018, ACL, 2018, pp. 61–67.

[27] J. Pennington, R. Socher, C. Manning, Glove: Global Vectors for Word Representation, in: Proc. of the 2014 Conf. on Empirical Methods in Natural Language Processing (EMNLP), 2014, pp. 1532–1543. doi:10.3115/v1/D14-1162.

[28] Q. V. Le, T. Mikolov, Distributed Representations of Sentences and Documents, Proc. of the 31st Int. Conf. on Machine Learning 32 (2014) 1188–1196.

[29] J. Devlin, M. Chang, K. Lee, K. Toutanova, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, in: Proc. of the 2019 Conf. of the North American Chapter of the Assoc. for Computational Linguistics, 2019, pp. 4171–4186. doi:10.18653/v1/N19-1423.

[30] N. Reimers, I. Gurevych, Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks, in: The 2019 Conf. on Empirical Methods in Natural Language Processing (EMNLP 2019), 2019.

[31] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, Q. V. Le, XLNet: Generalized Autoregressive Pretraining for Language Understanding, in: Advances in Neural Inf. Processing Syst. 32, 2019, pp. 5754–5764.

[32] D. Vrandecic, M. Krötzsch, Wikidata: a free collaborative knowledgebase, Commun. ACM 57 (2014) 78–85.

[33] J. Bromley, J. Bentz, L. Bottou, I. Guyon, Y. Lecun, C. Moore, E. Sackinger, R. Shah, Signature verification using a Siamese time delay neural network, Int. J. of Pattern Recognition and Artificial Intelligence 7 (1993).

[34] M. Ostendorff, T. Ruas, T. Blume, B. Gipp, G. Rehm, Aspect-based document similarity for research papers, in: Proceedings of the 28th International Conference on Computational Linguistics, International Committee on Computational Linguistics, Barcelona, Spain (Online), 2020, pp. 6194–6206. URL: https://aclanthology.org/2020.coling-main.545. doi:10.18653/v1/2020.coling-main.545.

[35] T. Wang, X. Wan, Hierarchical attention networks for sentence ordering, in: Proc. of the AAAI Conf. on Artificial Intelligence, volume 33, 2019, pp. 7184–7191.

[36] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, L. Zettlemoyer, Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension, arXiv (2019) arXiv–1910.

[37] D. E. Rumelhart, Notes on a Schema for Stories, in: D. G. Bobrow, A. Collins (Eds.), Representation and Understanding – Studies in Cognitive Science, Academic Press, New York, San Francisco, London, 1975, pp. 211–236.

[38] T. Caselli, P. Vossen, The event storyline corpus: A new benchmark for causal and temporal relation extraction, in: Proc. of the Events and Stories in the News Workshop, ACL, 2017, pp. 77–86.

[39] M. Bal, Narratology: Introduction to the theory of narrative, University of Toronto Press (1985). Trans. by Christine van Boheemen.

[40] W. V. H. Yarlott, M. A. Finlayson, ProppML: A Complete Annotation Scheme for Proppian Morphologies, in: B. Miller, A. Lieto, R. Ronfard, S. G. Ware, M. A. Finlayson (Eds.), 7th Workshop on Computational Models of Narrative (CMN 2016), volume 53 of *OpenAccess Series in Informatics (OASIcs)*, Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2016, pp. 8:1–8:19.

[41] V. Y. Propp, Morphology of the Folktale, Originally published in 1928; Translation 1968, The American Folklore Society and Indiana University, University of Texas Press, 1968.

[42] X. Yan, A. Naik, Y. Jo, C. Rose, Using functional schemas to understand social media narratives, in: Proc. of the Second Workshop on Storytelling, ACL, 2019, pp. 22–33.

[43] A. S. Gordon, C. A. Bejan, K. Sagae, Commonsense causal reasoning using millions of personal stories, in: Proc. of the Twenty-Fifth AAAI Conf. on Artificial Intelligence, 2011.

[44] R. Bois, G. Gravier, E. Jamet, M. Robert, E. Morin, P. Sébillot, M. Robert, Language-based construction of explorable news graphs for journalists, in: Proc. of the 2017 EMNLP Workshop on Natural Language Processing meets Journalism, ACL, 2017, pp. 31–36.

[45] S. Ribeiro, O. Ferret, X. Tannier, Unsupervised event clustering and aggregation from newswire and web articles, in: Proc. of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism, Assoc. for Computational Linguistics, Copenhagen, Denmark, 2017, pp. 62–67.

[46] A. Nie, E. Bennett, N. Goodman, Dissent: Learning sentence representations from explicit discourse relations, in: Proc. of the 57th Annual Meeting of the Assoc. for Computational Linguistics, ACL, 2019, pp. 4497–4510.

[47] W. V. Yarlott, C. Cornelio, T. Gao, M. Finlayson, Identifying the discourse function of news article paragraphs, in: Proc. of the Workshop Events and Stories in the News 2018, Assoc. for Computational Linguistics, Santa Fe, New Mexico, USA, 2018, pp. 25–33.

[48] T. v. Dijk, News as discourse, Communication Series, L. Erlbaum Associates, 1988.

[49] G. Rehm, K. Zaczynska, J. M. Schneider, M. Ostendorff, P. Bourgonje, M. Berger, J. Rauenbusch, A. Schmidt, M. Wild, Towards Discourse Parsing-inspired Semantic Storytelling, in: A. Paschke, C. Neudecker, G. Rehm, J. A. Qundus, L. Pintscher (Eds.), Proceedings of QURATOR 2020 – The Conf. for intelligent content solutions, Berlin, Germany, 2020. CEUR Workshop Proceedings, Volume 2535. 20/21 January 2020.