Reviewing as storytelling?

Examining narrative themes in video game reviews

Alexandra Barancová[1,*], Giovanni Sileno[1]

[1]Informatics Institute, Faculty of Science, University of Amsterdam, the Netherlands

Abstract

This paper outlines an approach to extracting and analysing narratives in reviews, aiming to represent domain-specific themes and various components of narrativity therein, such as events, entities, and sentiment. We report on experiments with the first two stages of this process. First, we use and evaluate an existing narrative detection model on a small set of textual video game reviews, confirming that the reviews contain narratives, but also finding ambiguities in the coding process. Second, through qualitative analysis, we identify several domain-specific themes that these narratives contain.

Keywords

computational narrative understanding, narrative detection, natural language processing, reviews, video games

1. Introduction

'I bought this and Transformers on the same day. I loaded this first and played for about 1 hour and it was much like an old game called "Horde" (10 yrs ago). I went to install Transformers, I played it for about 1/2 hour and was itching to get back to commanding my army or minions...'

The internet abounds with reviews. Although this genre itself, might we call it one, has a long tradition — especially if taken to include critiques of performances, exhibitions, books, or art pieces — "leaving a review" now happens at scale online. As a result, review datasets can offer longitudinal and spatio-temporal data, in some cases spanning several decades by now, as well as relatively rich content of free-form text, sometimes accompanied by photographs or video. Here, we approach reviews as first-person accounts of lived experiences [1] and are interested in accessing the stories told within. How might we extract and study them in order to better understand the objects, systems, or worlds they are about? In this inquiry, we focus specifically on video game reviews, as they are often based on in-game experiences and can play an important role in forming discourses about and collective understanding of the games themselves — as observed in [2], video game reviews tend to be more than just 'shopping guides'. Our work builds on and aims to contribute to the domain of computational narrative understanding.

To date, much of the computational work on narrative analysis and understanding has focused on storytelling in literary forms. As recent studies such as [3] or [4] point out, narratives — or features of 'narrativity', if we understand narratives as being pieced together from a multidimensional range of properties [5] — differ across domains, making it relevant to study the ways in which stories are told across a varied suite of places and purposes. Although an extensive body of work on review data dealing with tasks such as sentiment analysis or summarisation exists, to the best of our knowledge, little has been done in terms of identifying and exploring the narratives therein.

In this paper we outline an approach to extracting and analysing narratives contained in video game reviews, and report on two experiments, the initial steps in our process. First, we experiment with narrative detection on a set of video game reviews from the Amazon Reviews'23 [6] dataset using a fine-tuned RoBERTa model from the StorySeeker toolkit [4]. Second, through qualitative analysis, we identify several themes contained in narratives in video game reviews; the themes themselves provide

8th Workshop on Computational Models of Narrative (CMN 2025), May 28–30, 2025, Geneva, CH

*Corresponding author.

a.barancova@uva.nl (A. Barancová); g.sileno@uva.nl (G. Sileno)

a first confirmation of the value of this line of research, but more importantly, are intended to inform future computational work on narrative analysis of our data.

This study is therefore a stepping stone towards: (1) the implementation of a simple pipeline for narrative analysis of video game reviews, and (2) a wider exploration of ways in which we can leverage the narratives contained in reviews, focusing on the longitudinal and (relatively) polyvocal nature of such datasets. In the long run, our work aims to enrich our understanding of the kinds of stories told online and contribute to developing methods that help integrate narratives from various, often 'non-expert' sources into the way we understand sociotechnical systems. For ephemeral objects like video games, such methods could also contribute to preservation and archiving efforts.

2. Background

In this section, we briefly sketch out three areas of interest in order to situate our work. Firstly, the increasing availability and relevance of what we might call 'non-expert' knowledge as a resource for understanding sociotechnical systems. Secondly, work on computational narrative understanding, with a focus on narratives found online and methods informed by narrative theory. And thirdly, existing work dealing with reviews, touching on their contextual and historical value.

2.1. Stories told online

Over the past few decades, content posted and circulated online has become an important source and subject matter for a number of fields. Domains like science and technology studies, media studies, or anthropology offer a range of tools and perspectives on how to approach the online and hybrid realms our lives have become enmeshed in — see for example [7] or [8].

Research drawing on notions such as folk theories [9], myths [10] or gossip [11] has contributed to underscoring the relevance of lived experiences for making sense of sociotechnical systems [8]. For example, building on the notion of affect, and on phenomenology and ethnography, Bucher [12] explored what she termed the 'algorithmic imaginary' – ways in which 'ordinary users' understand algorithms, in this case those employed by Facebook – by analysing stories told in 2014-2015 Twitter posts and conducting interviews. Challenging narrow conceptions of what counts as 'expert', technical knowledge, such work foregrounds various forms of sense-making and storytelling practices for the study of technical objects and the systems they are part of. Stories about and collected from, for example, blogs, online forums, or social media platforms have become a valuable resource for researchers.

2.2. Computational narrative understanding

The emergence of large datasets and computational methods has made it possible to analyse stories found online at scale. Working with a subset of the Webis-TLDR-17 [13] corpus of Reddit posts and comments, for example, [4] explore the distinctiveness of stories told by different online communities. Such studies have come to form the domain of computational narrative understanding, combining, on the one hand, various methods, tasks, and models from natural language processing (NLP) and artificial intelligence (AI) with theory from literature and philosophy, on the other. For an elaboration on the connection to literary theory, see for example [14], where the authors elaborate the notion of narrativity [5] for computational narrative understanding, informed by theories of narrative developed by scholars like Gérard Genette [15] or Mieke Bal [16]. Recently, numerous helpful reviews on existing methods for tasks such as narrative detection or extraction [17] have been published. To date, most of the work in this domain has focused on textual data.

Much of the work in this domain reflects on the broader requirements of tools for computational narrative analysis and understanding, while grappling with case- and context-specific applications. Antoniak et al. [18] explore narrative, sentiment, and persona-based patterns in birth stories using the parser `spacy.io`, the Valence Aware Dictionary and sEntiment Reasoner (VADER) [19] and a latent Dirichlet allocation (LDA) [20] topic model, respectively. Saxena et al. [21] use Correlation Explanation

(CoRex) [22], a semisupervised topic modeling method through which existing domain knowledge can be integrated into the topic generation process, in a study on child welfare case notes. Such examples demonstrate that narratives have their own particularities across different contexts; what constitutes, or distinguishes, a narrative in one domain may not be the same for another [3]. As [4] show, the features of storytelling in online communities differ not only from those found in for instance, long-form literary works such as novels, but also across different online communities, delineated by subreddits.

2.3. Analysing (video game) reviews

Reviews exist in a wide range of virtual spaces, ranging from small community initiatives to large-scale profit-driven platforms. A correspondingly large body of work has been dedicated to analysing various dimensions of reviews as of the early 2000s, contributing to the development of methods for NLP tasks like text mining and summarization — e.g. Hu and Liu's [23] study on Amazon product reviews – or sentiment analysis — going back to Pang et al.'s [24] seminal study on IMDB movie reviews. Unsurprisingly, a large portion of the work on reviews is geared towards understanding and modelling aspects of consumer behaviour and purchasing patterns.

That said, several examples of story-focused approaches to review data conducting narrative analysis — e.g. [25] — and longitudinal studies — e.g. [26] — can be found in the literature. As relatively well-maintained and standardised longitudinal sets of first-person reports, review datasets can be a fruitful way to study changes in technology over time; framing the 2016 Amazon Product Review dataset as an archival collection, [26] set out to use it to explore technological progress. In the context of video games specifically, prior work has shown that reviews play an important role in creating and shaping our understanding of games and the virtual worlds they are often part of. A study focusing on a set of hand-coded reviews from two dedicated game review sites, IGN and GameSpot, argues that reviews play a central role in discourse on games, outlining a number of functions that they serve (e.g. feedback to the developer community) and highlighting their potential to help preserve video game history, by providing contextual information and documenting players' situated and lived experiences [2].

3. Outline of the planned approach

Building on previous work on narrative detection and analysis of online content that has dealt with stories told through, for example social media [4] or blogs [27], we turn our attention to video game reviews. As [14] argue, computational methods can contribute to understanding narratives in different narrative 'situations' through theory validation, theory reduction, and theory development. Adopting the aforementioned notion of narrativity, we hypothesise that video game reviews differ in their degree of story-likeness and that various components, or narrative features, contribute to the stories therein.

Our devised approach combines a limited, yet composite, and potentially extendable set of narrative features. As shown in Figure 1, we first check for the presence of narrativity in video game reviews; we begin with binary narrative detection, but may consider extracting story spans in the future. The next step is to understand the context-specific themes contained in the narratives through qualitative analysis. We then plan to analyse narrativity across these themes computationally through events, entities, and sentiment. We see these as roughly mirroring the basic elements of narrative synthesized by David Herman [5]: *situatedness* (in our case given by the domain of video game reviews), *event sequencing* (the reported events), *worldmaking* (entities that fill the story world), and *feltness* (sentiment(s) expressed by the reviewer). Our choice to model these particular components by proxy of rather established tasks is further motivated by their relative importance for computational narrative understanding — events [28] and animated entities [14], for example, have been shown to be key predictors of narrativity — and by their relative popularity in NLP at large, and therefore availability of existing tools to build on.

We see this as an intermediate step to a holistic approach for extracting more narrative components, as specified by a chosen framework. For example, building on prior work in narratology, [29] use a more fine-grained analysis of the features we focus on here (e.g. 'event sequences' and 'eventfulness') and include additional features not represented in our framework (e.g. 'concreteness' or 'abstraction').

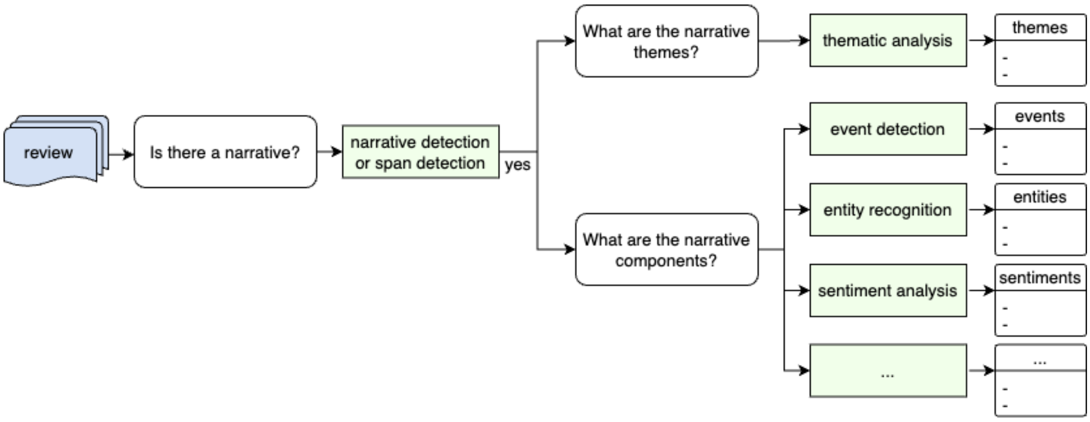

Figure 1: A diagram outlining our key questions and associated tasks for the analysis of narratives in reviews. In this paper we elaborate on experiments with narrative detection and thematic analysis.

Task	Model(s) or Tool(s)
Narrative detection	StorySeeker [4], "Detecting Narrativity across Long Timescales" [3], ...
Event detection	"Event Causality Extraction" [28], BookNLP [31], ...
Entity recognition	FLERT (through Flair) [32], BookNLP [33] ...
Sentiment analysis	Flair [34], VADER [19], ...

Table 1
A table summarizing existing computational methods that can be used for the narrative analysis tasks outlined in Figure 1. These include open models, tools in NLP pipelines and LLM prompts. References point to published work these are based on; in the case of BookNLP and Flair, these are separate papers for different tasks.

Different frameworks could also be considered, for example William Labov's theory of narrative, consisting of the 'abstract', 'orientation', 'complicating action', 'evaluation', 'resolution' and 'coda' [30].

Initially, we plan to test several existing open methods for the tasks of narrative detection [4][3], event detection [28][31], entity recognition [32][33], and sentiment analysis [34][19], in order to build on and connect established resources (Table 1 provides an overview). Although implementing the full framework is beyond the scope of this paper, in the next section we report on preliminary experiments with the first two components of Figure 1; narrative detection and qualitative analysis of our data. In the remainder of this section, we present the data we plan to work with.

3.1. Data

We plan to use two datasets containing reviews of video games: Amazon Reviews'23 [6] and Steam reviews.[1] Amazon is an e-commerce platform that sells a range of consumer products, including video games. Steam is a digital distribution platform dedicated to video games. Although Steam is the more popular gaming platform and resource where reviews can be linked to relevant metadata like the number of hours a reviewer had played a game at the time of reviewing it, for our initial experiment reported in this paper we focus only on the former, as we found that the data required fewer preprocessing steps.

The full 2023 version of the Amazon review dataset [6] contains 571.5M reviews from the period May 1996 to September 2023 that span 33 different product categories. The entire 'Video_Games' category contains 4.6M reviews by 2.8M different users on 137.2K unique items. Alongside video games, items in this category contain many associated peripheral products, such as hardware or accessories. We filter out all items with tags that signal non-game products such as 'PlayStation VR Hardware', 'Cases' or 'Protectors'.[2] Filtering reduces the set of unique items to 73.1K with a mean=31.5, std=185.0, and median=4.0 for the corresponding number of reviews per item, leaving a total of 2.3M reviews.

[1]A collection of Steam reviews for bestselling games from the period 2010-1019, which we may consider using, is available at: https://www.kaggle.com/datasets/luthfim/steam-reviews-dataset

[2]We did a quick manual proof-check of the filtering, comparing the first 50 entries in the non-filtered set, with the first 50 in the filtered version, even though not *all* hardware/accessories etc. had been removed all video games had been retained.

4. Preliminary experiments

As an initial exploration, we set out to determine (1) *Do we find stories in video game reviews?* and, if so, (2) *What characterises their narrativity?* To address these questions, we run a narrative detection model on a subset of our dataset and qualitatively examine a portion of the reviews.

4.1. Narrative detection

For this task, we use a pre-trained RoBERTa model [35] from the StorySeeker toolkit [4], which the authors fine-tuned on a corpus of Reddit posts and comments, annotated specifically for the task of story detection in online communities. Although these texts do not cover reviews, to the best of our knowledge, they come closest to our data compared to other solutions. Additionally, the authors found the model to generalise reasonably well for narrative detection on non-Reddit data [4].

We run the narrative detection model on a small subset of our filtered dataset. In order to assemble a substantial, yet still feasibly human-readable number of reviews per game, we draw a random sample by selecting all items that have exactly 100 reviews available in the dataset. Although seemingly arbitrary, this choice is meant to focus our qualitative analysis on several games with relatively good follow-up. This yields 37 items, which, following a manual check for residual non-game items, leaves 28 video games with 2,800 corresponding reviews. Following [4] we do not preprocess the individual texts, passing them directly to the model for prediction. This yields a binary (narrative/non-narrative) label. We find that 27.0% of the reviews are labelled 'narrative' by the model.

As our dataset is not annotated for narrativity, determining a precise accuracy score for the model's performance is beyond our current scope. Nevertheless, we approximate an evaluation using our own (cross-validated) annotations. One of the authors annotated all the reviews for 10 games, following the codebook in [4]. To account for one dimension that may cause stories told in reviews to differ, namely the sentiment about the given game, we select the 4 games with the lowest mean rating in our set (3.0-3.7/5.0), 2 middle (4.3/5.0), and 4 highest (4.7-4.8/5.0). For cross-validation, the second author annotated a random sample of 100 reviews. The inter-annotator agreement falls in the 'substantial' range (Cohen's k=0.619). The agreement between our labels and the model predictions is 83.3%.

False negatives included phrases retelling in-game stories, often in the present tense, e.g. 'A new dragon is in town and a romance is sparking.' False positives included short statements like 'I been waiting to play. Great game.' often containing past-tense verb phrases. As annotators, we agreed that such cases reported states that were too nonspecific to fit the narrativity criteria outlined in the codebook we followed. However, they did raise ambiguities, as we wondered what role the implicit context of reviewing should be given; seeing as the writer is likely reporting on their own experience, can we not assume that even seemingly "general" states are, in fact, accounts of concrete events? This hints that the codebook may require some re-adjustments.

4.2. Thematic analysis

Next, we qualitatively examined the reviews of the 10 games mentioned above to characterise — as human readers — the types of stories told in our data. Here, we present several emergent themes, accompanied by excerpts from reviews in our dataset. These themes are neither exhaustive nor mutually exclusive, but illustrate the value that narrative analysis can provide in this context. Video game reviews tell stories about:

the "product journey". 'It was sealed and brand new, but came with this horrible crack that goes into the plastic of the game box.' — written from the perspective of a consumer, assessing the logistics and services involved more so than the game itself. Often somewhat mundane, bordering on generic, aimed to either vouch for or warn against a service provider's credibility. Nevertheless, these stories can reveal information about things like material dimensions of video games, ranging from their means of distribution to packaging.

reason(s) for interest. 'Been a fan since the original Dead Space first released. I had really high hopes for this game upon seeing the first trailer. Day one release was a bit rocky with some texture anomalies (Especially on PS5) but if you are reading this now the developers have released a patch that has completely fixed the issue.' — in recounting concrete motivations for purchase, such narratives can reveal information about player bases of games, occasions that lead to people playing them (birthday gifts, price drops, others' positive reviews, etc.) or the development process of the game (announcements, trailers, pre-releases, etc.). In some cases, they can help situate a game within a wider network of media, like other games, films, or books that might share stories, universes, or genres.

technical obstacles. 'It keeps saying, no matter name I write in, "You're not the first to bear this name, please choose another name..." I tried removing the programs and re-installing them again and I even tried using a different laptop." — personal accounts of wrangling video games to get them working. These reveal the technical requirements of games, set-ups players would use, bugs that need(ed) fixing, while offering insights into players' own expectations, technical aptitude, and frustrations.

particular points in (game) time. 'It would not allow me to play chess in any mode without graduating from its pre-requisites.After an hour of frustration my five year old asked if we could just go back to the original "board game". I threw the program in the garbage.' — reports that capture how a player felt at the moment they delved into a game for the very first time, a number of hours in, or on the day they finished it. These are often loosely 'timestamped' in relation to a stage of the game in writing, and alongside feelings, specify the actions or events that ensued as a result.

anecdotes in/with the game. 'I threw a guy off a bridge and there was a ship passing underneath it, there was a special heat cutscene for that.' — highlight in-game events, which are for whatever reason illustrative or noteworthy; sometimes these might be key moments in a game's plot, other times just serendipitous occurrences that stuck with the reviewer. Such stories can help trace the journeys players take through games or contextualise things like glitches. Especially for non-linear or open world games, the former can help illustrate how people spent their time in otherwise undirected worlds.

drifting between worlds. 'But now, even though my son no longer plays, he still dreams about beating King Black.' — various out-of-game anecdotes that show ways in which in-game encounters and events become enmeshed in people's lived realities. These might range from relatively pragmatic reports on lessons learned from gameplay (e.g. whether or not the anticipated skillset or knowledge was acquired from an educational game) to ways in which players relate to concrete characters.

These themes demonstrate several axes through which narrativity can be studied in reviews, providing different views of the sociotechnical system under question, in this case a series of video games.

5. Conclusion and discussion

Experimenting with the first step of narrative detection, we conclude that video game reviews contain narratives and anticipate that the pretrained model from [4] tested here can identify these reasonably well. Inevitably, we find some ambiguities in what makes for a "narrative" review. During our annotation process, we found differentiating events from states — a distinction also flagged by e.g. [31] and [4] — particularly tricky in some cases, as many texts are descriptive, written in an objective/generalising tone. Although these often did not fit the criteria for narratives based on the codebook (and we therefore did not annotate them as such here), we speculated whether or not the coding for review data may need to be slightly revised to better account for both in- and out-of-game narratives and take into consideration the specific context of reviewing.

Based on our qualitative analysis, we observe that the stories in video game reviews contain various distinct themes. In order to ensure generalisability beyond the video game domain, it will be relevant to explore how these domain-specific narratives tie to more established narrative archetypes and functions. Within our domain, it will be relevant to determine how our own and existing characterisations of video game review content, such as in [2],[3] will structure the modelling work in the next stages. Integrating further qualitative methods, next to our proposed computational framework, may be a relevant dimension for future work — [9], [12], or [11], for example, all make use of interviewing, a method that might be equally enriching in this context. Our preliminary observations need to be tested further; more systematically and using more data. Game reviews on Steam may, for instance, have slightly different characteristics than those on Amazon, to which our exploration was limited so far.

Alongside testing the models summarised in Table 1, further work will involve determining useful evaluation methods, suited for a domain like ours with little to no pre-existing datasets or annotations focused on narrative features. One approach could be by 'triangulating' the results from several models. Additionally, we may consider implementing the suite of tasks using an open source large language model; [29] set a thorough precedent for this.

References

[1] D. Chandler, R. Munday, A dictionary of media and communication, Oxford University Press, USA, 2011.

[2] J. P. Zagal, A. Ladd, T. Johnson, Characterizing and understanding game reviews, in: Proceedings of the 4th International Conference on Foundations of Digital Games, ACM, Orlando Florida, 2009, pp. 215–222.

[3] A. Piper, S. Bagga, L. Monteiro, A. Yang, M. Labrosse, Y. L. Liu, Detecting Narrativity Across Long Time Scales, in: CEUR Workshop Proceedings, CEUR-WS.org, Amsterdam, 2021.

[4] M. Antoniak, J. Mire, M. Sap, E. Ash, A. Piper, Where Do People Tell Stories Online? Story Detection Across Online Communities, 2024.

[5] D. Herman, Basic elements of narrative, John Wiley & Sons, 2009.

[6] Y. Hou, J. Li, Z. He, A. Yan, X. Chen, J. McAuley, Bridging Language and Items for Retrieval and Recommendation, 2024.

[7] N. Seaver, Algorithms as culture: Some tactics for the ethnography of algorithmic systems, Big Data & Society 4 (2017).

[8] R. Kitchin, Thinking critically about and researching algorithms, in: The Social Power of Algorithms, Routledge, 2018.

[9] M. Eslami, K. Karahalios, C. Sandvig, K. Vaccaro, A. Rickman, K. Hamilton, A. Kirlik, First I "like" it, then I hide it: Folk Theories of Social Feeds, in: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, ACM, San Jose California USA, 2016, pp. 2371–2382.

[10] V. Mosco, The digital sublime: Myth, power, and cyberspace, Mit Press, 2005.

[11] S. Bishop, Managing visibility on YouTube through algorithmic gossip, New Media & Society 21 (2019) 2589–2606.

[12] T. Bucher, The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms, Information, Communication & Society 20 (2017) 30–44.

[13] M. Völske, M. Potthast, S. Syed, B. Stein, TL;DR: Mining Reddit to Learn Automatic Summarization, in: Proceedings of the Workshop on New Frontiers in Summarization, Association for Computational Linguistics, Copenhagen, Denmark, 2017, pp. 59–63.

[14] A. Piper, S. Bagga, Toward a Data-Driven Theory of Narrativity, New Literary History 54 (2022) 879–901.

[15] G. Genette, A. Levonas, Boundaries of Narrative, New Literary History 8 (1976) 1.

[16] M. Bal, C. v. Boheemen, Narratology: introduction to the theory of narrative, 3rd ed., University of Toronto Press, Toronto, 2009.

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

[3] [2] identify the following themes: Description, Personal Experience, Reader Advice, Design Suggestions, Media Context, Game Context, Technology, Design Hypotheses, Industry

[18] M. Antoniak, D. Mimno, K. Levy, Narrative Paths and Negotiation of Power in Birth Stories, Proceedings of the ACM on Human-Computer Interaction 3 (2019) 1–27.
[19] C. Hutto, E. Gilbert, Vader: A parsimonious rule-based model for sentiment analysis of social media text, in: Proceedings of the international AAAI conference on web and social media, volume 8, 2014, pp. 216–225.
[20] D. M. Blei, A. Y. Ng, M. I. Jordan, Latent dirichlet allocation, Journal of machine Learning research 3 (2003) 993–1022.
[21] D. Saxena, E. S.-Y. Moon, A. Chaurasia, Y. Guan, S. Guha, Rethinking "Risk" in Algorithmic Systems Through A Computational Narrative Analysis of Casenotes in Child-Welfare, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, ACM, Hamburg Germany, 2023, pp. 1–19.
[22] R. J. Gallagher, K. Reing, D. Kale, G. Ver Steeg, Anchored Correlation Explanation: Topic Modeling with Minimal Domain Knowledge, Transactions of the Association for Computational Linguistics 5 (2017) 529–542.
[23] M. Hu, B. Liu, Mining and summarizing customer reviews, in: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '04, ACM, New York, NY, USA, 2004, p. 168–177.
[24] B. Pang, L. Lee, S. Vaithyanathan, Thumbs up? sentiment classification using machine learning techniques, in: Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10, EMNLP '02, Association for Computational Linguistics, USA, 2002, p. 79–86.
[25] B. Fisseni, A. Kurji, D. Sarikaya, M. Viehstädt, Story Comparisons: Evidence from Film Reviews, OASIcs, Volume 32, CMN 2013 32 (2013) 94–99.
[26] Y. Zhang, A. Jatowt, K. Tanaka, Causal relationship detection in archival collections of product reviews for understanding technology evolution, ACM Trans. Inf. Syst. 35 (2016).
[27] K. K. Bandeli, M. N. Hussain, N. Agarwal, A framework towards computational narrative analysis on blogs., in: Text2Story@ ECIR, 2020, pp. 63–69.
[28] Y. Sun, Q. Chao, B. Li, Event causality is key to computational story understanding, in: Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), Association for Computational Linguistics, Mexico City, Mexico, 2024, pp. 3493–3511.
[29] A. Piper, S. Bagga, Using Large Language Models for Understanding Narrative Discourse, in: Proceedings of the The 6th Workshop on Narrative Understanding, Association for Computational Linguistics, Miami, Florida, USA, 2024, pp. 37–46.
[30] W. Labov, The language of life and death: The transformation of experience in oral narrative, Cambridge University Press, 2013.
[31] M. Sims, J. H. Park, D. Bamman, Literary event detection, in: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Florence, Italy, 2019, pp. 3623–3634.
[32] S. Schweter, A. Akbik, Flert: Document-level features for named entity recognition, arXiv preprint arXiv:2011.06993 (2020).
[33] D. Bamman, S. Popat, S. Shen, An annotated dataset of literary entities, 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), Association for Computational Linguistics, Minneapolis, Minnesota, 2019, pp. 2138–2144.
[34] A. Akbik, T. Bergmann, D. Blythe, K. Rasul, S. Schweter, R. Vollgraf, FLAIR: An easy-to-use framework for state-of-the-art NLP, in: NAACL 2019, 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics, 2019, pp. 54–59.
[35] Y. Liu, Roberta: A robustly optimized bert pretraining approach, arXiv preprint arXiv:1907.11692 364 (2019).