Challenges and Opportunities for Computational Construction of Narratives

Pablo Gervás

Introduction

As Artificial Intelligence enters a new era where large language models show surprising capabilities for handling language, it becomes important to review the challenges and insights accumulated by the community of researches that have worked on building computational modelling of literary creativity over the years. The efforts of this community have at times achieved success at small tasks and at other times faced failure when over-ambitious goals were pursued. It would be important for efforts on related tasks undertaken from this point on to keep in mind the insights accumulated on the nature of the task and the challenges it presents. The present paper attempts to summarize some of these insights into a set of challenges that face the automated generation of narratives, but also as a set of opportunities open for the future.

The paper is structured as a review of a number of possible ways to formulate the problem of having a program generate automatically a story on demand, followed by some reflections on how storytelling might be subdivided into a set of related subtasks as suggested by that set of formulations of the problem.

Programs that Deliver Stories

The possible ways of building programs that deliver stories explored in this paper are presented from the simpler formulation towards increasingly complex approaches. At each point the challenges (or lack thereof) and opportunities presented are analyzed, trying at each point to lead on to the approach discussed after it.

How to Request Stories

The simplest possible formulation for requesting a story from a story generation program is one where there is no input and the output should be a story. Such a program might be built by considering some existing repository of stories and to select from it one story at random. The user would get a new story every time, and she would have difficulty to tell whether the story has been generated by the system or selected from a set of prewritten stories. A slightly more elaborate approach would be to consider an input in terms of a series of keywords or phrases to narrow down the set of possible stories returned. The development of a system under this approach requires the existence of a vocabulary for presenting queries formulated in terms of various characteristics such as: details about the concepts in the story world, details about characters and relations between them, plot structures that capture causal and temporal relations between events in the story, or details about the potential effect of the story on its audience. The most elaborate approach would be to provide a text prompt that describes the story in some way, and have the story generator provide a story matching the prompt.

Existing story generators have progressively shifted from the early attempts that worked on an empty input to the most recent neural solutions that accept complex text prompts, with intervening systems that worked on keyword-based queries of different complexity.

Original Stories

If one further establishes the constraint that the outcomes must be original – that is, not already available in some form – the task becomes more difficult. Two different challenges need to be met: how to build a story and how to judge if a newly built story is sufficiently different from those already existing to be considered original.

Building Stories

The task of building a story from constituent elements, in the hope that the result may be original, requires solving two basic problems: what to consider as constituent elements and what type of procedure can be employed to put together those constituents so that the result be a valid story. Several approaches have been considered in the past, but only some are reviewed here as illustrative examples. For more detailed reviews, readers are referred to existing surveys in the field (Kybartas / Bidarra 2016; Alhussain / Azmi 2021). Some approaches consider breaking existing stories into pieces and then recombine the resulting set of pieces from a large set of different stories into new instances of stories.

The simplest approach of this type relies on individual words as the pieces that are obtained from stories and then recombined. Attempts to achieve this based on statistical models of have proven successful in the past at producing legible text, original but usually lacking sense (Brown et al. 2015). This procedure is also the basis for the recent work on transformers (Wolf et al. 2020) that has become the rage in the world of AI. It is only when systems built based on these models have started to be prompted with full sentences describing the desired output that successful results have made their appearance. We will address these models in the section "Acceptable Stories", when I discuss more elaborate ways of requesting a story.

Another approach to generating stories from pieces of prior stories involves using instead of words some abstraction of the meaning of stories and finding ways of recombining these abstractions into conceptual representations of a story that can then be transcribed as text. That is essentially what story generation based on planning does (Young et al. 2013). In this approach, a set of planning operators is constructed as an abstraction of the events that appear in a story. Each *planning operator* represents an event in a story – usually in some form of predicate logic that allows a predicate to represent the action and a set of argument variables to represent the characters or objects that participate in it – but it also encodes additional information of which other predicates are preconditions or post-conditions of that event. This allows a planner to create chains of events that are causally linked to one another, from an initial situation to a goal. For planning-based story generators, both the initial situation and the final goal are usually provided as input.

This type of approach is faced with the important challenge of having to generate a body of planning operators of sufficient coverage to generate a broad range of stories, which constitutes a significant bottleneck for the approach. Attempts have been made to solve the problem by means of advanced methods of knowledge engineering (O'Neill / Riedl 2014) and crowdsourcing (Guzdial et al. 2015).

Other efforts rely on similar abstractions of the meaning of events in a story as predicates associated with pre-conditions and post-conditions but forego the emphasis on goal-driven causality of the planning paradigm. Such systems design their abstractions for story actions based on different conceptual approaches. The Mexica system (Pérez y Pérez 1999) relies on a set of story actions that associate character emotions and tensions between characters as preconditions or post-conditions. The PropperWryter system (Gervás 2015) relies on existing abstractions on plot relevant actions (Propp 1928) and dependencies established between them by virtue of being actions associated with narrative roles such as villains or heroes.

This approach also suffers from the lack of a broad vocabulary of story actions to employ. This is known in artificial intelligence as the *knowledge acquisition bottleneck* – a long-standing problem (Cullen / Bryman 1988) that remains current to this day (Pasini 2021). Supported by a knowledge engineering effort to annotate plots of musicals (Gervás et al. 2016), the PropperWryter system achieved success when it was engaged in the production of the first computer-generated musical, which was staged at the London West End for two weeks in 2016 (Colton et al. 2016).

Judging Story Originality

However, building complete stories up from elementary blocks is no guarantee that the stories will be original. If the blocks used to build them correspond to elements that appear in already existing stories – and this is usually the motivating requirement when building knowledge resources to use as such building blocks – and the construction procedure is guided by reasonable heuristics, there is a non-zero probability that the building process result in stories very similar to the ones that inspired the construction of the knowledge resource. Furthermore, when story construction efforts of this kind have to be maintained over time, it becomes important to devise means to avoid the repetition of stories built previously by the system.

To address this challenge, research has been carried out on developing means for deciding when a story is similar to another (Fisseni / Löwe 2014; Hervás et al. 2015) or when a story can be considered sufficiently novel (Peinado et al. 2010). The existence of these efforts is very significant, because it opens a new avenue of research on computational storytelling, one where the systems do not only generate stories but also include the ability to produce judgments of some kind over them. This follows a general tendency in the field of computational creativity to progressively evolve from systems that merely generate artifacts to systems that develop their own aesthetic and which are capable of defending why the artifacts they produce are valuable (Colton / Charnley / Pease 2011).

Acceptable Stories

The need for systems that construct artifacts to include procedures for assessing the quality of their outputs arises from the fact that such systems essentially explore a search space of possible artifacts, and not all the artifacts in that search space are equally valid. Once a specific type of artifact is defined in a way that allows instances to be constructed computationally, it becomes very easy to produce a very large number of different instances of it. Ensuring that the produced instances are acceptable is usually a little more difficult. Making them all be valuable instances is a significant challenge.

To address this challenge in the field of narrative generation, recent research efforts have focused on defining metrics for story quality. Following traditional engineering practice, these efforts usually identify one specific aspect that is known to impact the perception of quality and attempt to model it, without necessarily considering other aspects that are also relevant. Examples of aspects that have been considered in the development of metrics for story quality are: character believability (Gomes et al. 2013), semantic coherence across the story on relevant events such as birth / death or romantic entanglements (Gervás / Concepción / Méndez 2021), reproduction of features observed in human-written stories (Leon et al. 2020), consistent use of entities across the narrative (Papalampidi / Cao / Kocisky 2022) or probability of each sentence in the story with and without its preceding story context (Sap et al. 2022).

The set of relevant aspects that should be considered is very large, and efforts should be made to integrate these various models so they can be applied together onto the same stories. The quality of story generators would improve significantly once these metrics can be applied as filters to their output.

The integration of a metric on story quality – essentially a story critic module – into a story generator presents another important challenge. Story generation systems are generally constructed on the basic assumption that they take an input and generate an output that is a story. The idea that one should read the output, reflect upon it and then, based on that reflection, rework it in some way is fundamental for the human approach to writing (Flower / Hayes 1981; Sharples 1999). However, generation systems constructed on a similar basis are few and far between. The Mexica system is a remarkable exception, being based on Sharples' cognitive model of the writing task (Sharples 1999).

There is another view on the quality of stories that differs slightly from whether it is simply a good story. This view considers whether the story is acceptable as an instance of a story described in some way by the user that requests it. Many of the features described above as potential input could be used to narrow down the set of stories that are acceptable in response to a specific input: categories of stories, details on the world or the characters, plot structures or potential effects of the story on its audience. If an interface of this type is made available, a story returned by a story generator in response to a specific request would only be acceptable to the extent that it matches the given request.

The recent development of neural-based solutions for text generation such as GPT-2 and GPT-3 (Zhang / Li 2021) makes this type of solution capable of responding to user prompts with large paragraphs of valid and fluid text. These solutions are based on a large language model trained over a neural network representation, and these models essentially capture the relationships between words as featured in examples of text written by humans. The interactive nature of these solutions has made it possible to present requests for stories as simple sentences that describe the desired story. The tests carried out so far on how acceptable the responses of these systems are in terms of their correlation with the given prompts show outstanding results for instances of general conversation, even though testers remain unconvinced of their general validity (Elkins / Chun 2020). More exhaustive tests need to be carried out on the applicability of these solutions in the realm of narrative generation.

Stories for a Specific Purpose

This observation leads into the final but not less important challenge that story generation faces: generating stories for specific purposes. The concept of purpose is very difficult to represent formally in terms that a computational system would understand. Yet it is crucial for guiding the composition of any message in any media that is aimed at communicating to an audience (Smedley 1952).

Story generator systems very rarely consider any representation of purpose among their inputs. In most cases, when a story generator is designed there is some idea in the mind of the designer of what the generated stories should achieve. As a result, story generation modules are sometimes included in systems designed for very clear purposes, such as teaching children about bullying (Aylet et al 2007), supporting emergency rescue training (Hullett / Mateas 2009), or for military training (Zook et al. 2012). In such cases, there is a clear purpose that the generated story has to fulfill, and the generator is designed with that purpose in mind. But the identification of the purpose and the tuning of the generation mechanism to ensure the generated stories achieve it are wholly in the mind of the designer and not explicitly modeled in the system.

The main body of research on storytelling has focused on generating stories whose content is constructed at the same time as the story. This actually sidesteps one of the main purposes of storytelling as used by people, which is to convey a set of events that has actually happened. In this case, the overall description of the task changes significantly. The set of inputs to consider to a potential system that builds such stories must necessarily include a description of the events that the story needs to convey. In most cases, the construction procedure is expected to respect that set of events and not introduce any additional ones that did not occur. This slightly different formulation of the task has been addressed much more rarely but instances exist of systems that generate stories about events that the story generator receives as input: stories about past interactions between a user and an intelligent agent (Behrooz / Swanson / Jhala 2015), stories abstracted from the moves of the pieces in a given chess game (Gervás 2014), stories about a user constructed from data on their routines acquired via sensors (Reddington / Tintarev 2011), stories to explain cybersecurity logs (Afzaliseresht et al. 2020), narratives from personal digital data (Farrow / Dickinson / Aylett 2015) or narrative biographies from knowledge extracted from the web (Kim et al. 2002).

Living as we do in world obsessed with fake news, we need also to consider a revised version of the task where a story is generated that is based on a given set of facts, but purposefully departs from it at some point when building a story about them. This is another storytelling task that is intuitively familiar to most people: biased historians or politicians do it often to their audiences, parents do it to soften the world for their children, transgressors do it to hide their offences. It is also the key task underlying historical fiction or fictional accounts of recent events. From a computational point of view this task can be understood as a combination of the task of generating a new story with the task of telling a known set of facts. This particular view of the task has been addressed computationally by attempting to match the known set of events to a particular plot structure, allowing mismatching real events to be omitted, and using the plot structure to provide additional fictional events to complete the story (Gervás 2018a, 2018b).

Another possible approach is to tell stories tailored for a specific audience. This is, yet again, an aspect of the storytelling job that is considered inseparable from the task itself by human writers, and yet very rarely addressed in story generation systems. Fortunately, we are beginning to see research aimed in this direction, such as a module to enhance a video game by being able to tell stories tailored to a specific player, based on a model for each individual player (Ramirez / Bulitko 2012).

Whether the goal is to achieve a particular purpose or satisfy a specific user, it is clear that generator systems would do a much better job the better they understand the potential reactions of the audience of their output. Computational models of how this might be achieved in terms of having specific models of the reader reaction included in the generation system have been proposed (Gervás / León 2016). Although this type of integration is yet in the future, efforts already exist to build computational models of the reading task. As in the case of metrics on story quality, specific approaches tend to focus on specific aspects of the task, such as understanding the set of events narrated in a given discourse (Niehaus / Young 2014), modeling the reaction of the reader in terms of an evolving curve of suspense (Doust 2015), interpretation of embedded stories told within a frame story (Gervás 2021), or reconstruction of the actual chronology of a story told in a discourse that allows flashbacks (Gervás 2022).

Rethinking Basic Assumptions

In view of the analysis presented to this point, it is important to accept that the concept of computational story-telling should be considered not as a single task, but as a set of interconnected tasks, all related to narrative, but corresponding to formulations that are very different in computational terms, as described above. Most of the existing systems for computational storytelling operate at the level of discourse as understood by Ricoeur (1976): a sequence of sentences, where each sentence involves a predicate applied to some entities that need to be identified by the subject (and objects) of the sentence. This view abstracts away from what Ricoeur describes as “the particular structure of the particular linguistic system”, which would be closer to the text.

From the point of view of a detailed analysis of computational storytelling, this distinction between text and discourse provides a key tool to understand the broad range of solutions that have been developed over the years. Suppose we consider that a story describes a *story world* – often fictional but not necessarily so. Suppose also that its narrative structure can be represented by some form of *discourse,* that may be rendered as *text* but which may allow rendering in other formats. Under these assumptions, computational storytelling may be described in terms of a number of component sub-tasks that embody transitions between story world, discourse and text. The sub-task of *narrative generation* involves construction *ex novo* of either text, discourse or world. The sub-task of *narrative composition* involves building a discourse that tells some story from a world, or building a text that tells a given discourse. The sub-task of *narrative interpretation* involves reconstructing the conceptual discourse that underlies a given text, or reconstructing the world that underlies a give discourse. These ideas are expanded in the following sections.

Constructing Narratives: Narrative Generation and Narrative Composition

Let us consider narrative generation as the task of constructing a story that did not exist before. Faced with the distinction between story world, discourse and text, an engineer wishing to design a story generator needs to decide how many of these levels she is willing to represent within her design. All these levels are interconnected, in the sense that a reader faced with a text will interpret a discourse as its meaning, and imagine a story world as its content. The question is at how many of these levels will the program operate.

The system could generate text directly, with no additional representations in terms of discourse or story world. This is indeed the choice favoured by neural approaches to the story generation task.

The system could focus on generating discourse, abstracting away from the complexities of generating language but still building a sequence of discourse – represented conceptually – that has a valid narrative structure. This presents the advantage that the story world does not have to be represented in the system. The approaches described above that rely on abstract representations of the meaning of stories as story actions operate under this choice.

Because people are better readers of text than of predicate logic representations, systems that generate discourse usually include a module that transcribes the discourse onto text. This process is not really generating the story, but rather finding a way of telling as text a story that already exists as discourse. We refer to this task as narrative composition.

The system could include an explicit representation of the story world. Stories can now be built directly by generating events or characters, or locations in the story world directly and then find a way of telling the story of what happens in her story world. This is the choice favoured by story generators based on simulation of an underlying story world (Ryan 2018). Under this approach, two further tasks are needed: one that transcribes possibly a selection of the events in the story world onto a discourse, and one that renders the resulting discourse as text. Both are instances of narrative composition.

With respect to the classic pipeline for natural language generation systems (Reiter / Dale 2000) the transition from story world to discourse corresponds to the task of content planning, and the transcription from discourse to text to the stages of sentence planning and surface realization. The content planning task will optionally include a task of selecting which events in the story world are included in the story, which aligns well with the recently formulated task of story sifting (Ryan 2018). Research efforts related to generation of stories based on known facts covered in the section "Stories for a Specific Purpose" correspond to tasks of narrative composition.

Processing Narratives: Narrative Interpretation

When the situation is presented in this manner, it becomes clear that there is a related set of possible computational tasks corresponding to the transitions from text to discourse and from discourse onto a conceptual representation of the story world. These correspond to subtasks of narrative interpretation. This task of narrative interpretation has been a long-sought goal in the field of artificial intelligence, originally known as natural language understanding (Allen 1995). Historically, a succession of un-met expectations over the years lead the natural language processing community to operate under a progressively shrinking scope, focusing more and more on specific sub-tasks instead of attempting anything like processing text onto an exhaustive representation of its meaning. Recent efforts to model identification of narrative structure (Gervás 2022) would correspond to the transition between discourse and story world, including the possibility of a given discourse referring not to one single story world but to a number of interrelated story worlds.

The attempts at developing computational models of the reading task would also constitute instances of the narrative interpretation task. Very little is known about this task, because it has only recently started to be treated computationally. This makes it stand out as an open field for challenges and opportunities for future research.

Conclusions

The panorama of research efforts described in this paper suggest that story telling needs to be considered not as a single monolithic task but as a set of interconnected sub-tasks. A simple distinction between text, discourse and story world provides means for describing some of these sub-tasks in terms of processes of narrative generation, narrative composition and narrative interpretation

In terms of how the existing research on computational narrative relates to human performance on equivalent basic goals related to the construction of stories, it is clear that computational solutions have a long way to go. The analysis presented in the sections "Programs that Deliver Stories" and "Rethinking Basic Assumptions" constitutes a fair review of the particular tasks considered to this point in past research on computational narrative. However, it becomes clear that, regardless of which of the combinations of transitions between the levels have already been explored in specific computational systems, the approach that humans apply to the storytelling task takes advantage of a very broad combination of this set of subtasks. Some writers plan out the plot of their story to the very end and only then sit down to write it. This would align with the idea of building a discourse and then telling it. Other authors construct a whole word and then find stories to tell in that world. Others sit down each day to produce a number of pages of text with no idea of how the plot is going to proceed as a result. Even the same writer may start from a sentence that sounds promising to him (text), work out in his head how the characters might react to the resulting situation (simulation on the story world) then rework it all to ensure that the revealing moment takes place at the optimal point in the scene to maximize its impact on the reader (iteration of discourse revision guided by a metric for impact), possibly guided by a dynamic model of how the reader progressively react to the sentences as they appear in the text.

As a result, to the various specific challenges described in the paper one must add the grander challenge of exploring how all of these specific challenges may interact in more complex processes of story writing.

Acknowledgments

This paper has been partially funded by the project CANTOR: Automated Composition of Personal Narratives as an aid for Occupational Therapy based on Reminiscence, Grant. No. PID2019-108927RB-I00 (Spanish Ministry of Science and Innovation) and the ADARVE (Análisis de Datos de Realidad Virtual para Emergencias Radiológicas) Project funded by the Spanish Consejo de Seguridad Nuclear (CSN), Grant Ref. SUBV-20/2021.

Bibliography

Afzaliseresht, Neda / Yuan Miao / Sandra Michalska / Qing Liu / Hua Wang (2020): "From Logs to Stories: Human-centred Data Mining for Cyber Threat Intelligence". *IEEE Access* 8: 19089–19099. https://doi.org/10.1109/ACCESS.2020.2966760.

Alhussain, Arwa I. / Aqil M. Azmi (2021): "Automatic Story Generation: A Survey of Approaches", *ACM Computing Surveys (CSUR)* 54(5): 1–38.

Allen, James (1995): *Natural Language Understanding*. San Francisco: Benjamin-Cummings Publishing Co., Inc.

Aylett, Ruth / Marco Vala / Pedro Sequeira / Ana Paiva (2007): "Fearnot! An Emergent Narrative Approach to Virtual Dramas for Anti-bullying Education". *Proceedings of the 4th International Conference on Virtual Storytelling: Using Virtual Reality Technologies for Storytelling*, Berlin: Springer Verlag: 202–205.

Behrooz, Morteza / Reid Swanson / Arnav Jhala (2015): "Remember that Time? Telling Interesting Stories from Past Interactions", H. Schoenau-Fog / L. E. Bruni / S. Louchart / S. Baceviciute (eds.) *Interactive Storytelling*. Cham: Springer International Publishing: 93–104.

Brown, Peter F. / Vincent J. Della Pietra / Peter V. deSouza / Jennifer C. Lai / Robert L. Mercer (1992): "Class-based N-gram Models of Natural Language". *Computational Linguistics* 18(4): 467–480. https://aclanthology.org/J92-4003 (21 January 2024).

Colton, Simon / John W. Charnley / Alison Pease (2011): "Computational Creativity Theory: The FACE and IDEA Descriptive Models". *ICCC:* 90–95.

Colton, Simon et al. (Jun. 2016): "The Beyond the Fence Musical and Computer Says Show Documentary". *ICCC*, Paris: Sony CSL.

Cullen, Jet / Alan Bryman (1988): "The Knowledge Acquisition Bottleneck: Time for Reassessment?", *Expert Systems* 5(3): 216–225.

Doust, Richard (2015): *A Domain-independent Model of Suspense in Narrative*. Ph.D. thesis. Open University (United Kingdom).

Elkins, Katherine / Jon Chun (2020): "Can GPT-3 Pass a Writer's Turing Test?", *Journal of Cultural Analytics* 5(2): 1–16.

Farrow, Elaine / Thomas Dickinson / Matthew P. Aylett (2015): "Generating Narratives from Personal Digital Data: Using Sentiment, Themes, and Named Entities to Construct Stories", J. Abascal et al. (eds.) *Human- Computer Interaction*. Cham: Springer International Publishing: 473–477.

Fisseni, Bernhard / Benedikt Löwe (2014): "What Makes Stories Similar? Report on a Research Project, 2011–2014", Mark A. Finlayson / Jan C. Meister / Emile Bruneau (eds.) *2014 Workshop on Computational Models of Narrative*, Dagstuhl: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik: 9–12. https://doi.org/10.4230/OASIcs.CMN.2014.9.

Flower, Linda / John Hayes (1981): "A Cognitive Process Theory of Writing". *College Composition and Communication* 32(4), 365–387.

Gervás, Pablo (2014): "Composing Narrative Discourse for Stories of Many Characters: A Case Study Over a Chess Game". *Literary and Linguistic Computing* 29(4).

Gervás, Pablo (2015): "Computational Drafting of Plot Structures for Russian Folk Tales", *Cognitive Computation* 8(2), 187–203.

Gervás, Pablo (2018): "Storifying Observed Events: Could I Dress this up as a Story?". *5th AISB Symposium on Computational Creativity*, Liverpool: University of Liverpool.

Gervás, Pablo (2018): "Targeted Storyfying: Creating Stories About Particular Events". *Ninth International Conference on Computational Creativity*, Salamanca: Association of Computational Creativity.

Gervás, Pablo (2021): "A Model of Interpretation of Embedded Stories". *Text2Story: 4th International Workshop on Narrative Extraction from Texts*, Lucca: CEUR Workshop Proceedings.

Gervás, Pablo (2022): "A Discourse Interpretation Engine Sensitive to Truth Revisions in a Story". *Tenth Annual Conference on Advances in Cognitive Systems*, Arlington.

Gervás, Pablo / Eugenio Concepción / Gonzalo Méndez (2021): "Assessing Multiplot Stories: From Formative Analysis to Computational Metrics". *International Conference on Computational Creativity*. Ciudad de México.

Gervás, Pablo / Raquel Hervás / Carlos León / C. V. Gale (2016): "Annotating Musical Theatre Plots on Narrative Structure and Emotional Content". *Seventh International Workshop on Computational Models of Narrative*, Kravov.

Gervás, Pablo / Carlos León (2016): "Integrating Purpose and Revision into a Computational Model of Literary Generation", M. Degli Esposti / E. Altmann / F. Pachet (eds.) *Creativity and Universality in Language*, Cham: Springer: 105–121.

Gomes, Paulo / Ana Paiva / Carlos Martinho / Arnav Jhala (2013): "Metrics for Character Believability in Interactive Narrative", H. Koenitz et al. (eds.) *Interactive Storytelling*. Cham: Springer International Publishing: 223–228.

Guzdial, Matthew / Brent Harrison / Boyang Li / Mark O. Riedl (2015): "Crowdsourcing Open Interactive Narrative", *Proceedings of the 10th International Conference on the Foundations of Digital Games*. Paci c Grove: FDG 2015.

Hervás, Raquel / Antonio Sánchez-Ruiz / Pablo Gervás / Carlos León (2015): "Calibrating a Metric for Similarity of Stories Against Human Judgment". *Creativity and Experience Workshop, International Conference on Case-Based Reasoning*, Bad Homburg, Frankfurt.

Hullett, Kenneth / Michael Mateas (2009): "Scenario Generation for Emergency Rescue Training Games". J. Whitehead / R. M. Young (eds.) *Proceedings of the 4th International Conference on Foundations of Digital Games*, Orlando: ACM: 99–106. https://doi.org/10.1145/1536513.1536538.

Kim, Sanghee et al. (2002): "Artequakt: Generating Tailored Biographies from Automatically Annotated Fragments from the Web", *Workshop on Semantic Authoring, Annotation & Knowledge Markup, 15th European Conference on Artificial Intelligence*, Lyon: 1–6. https://www.researchgate.net/publication/37535731_Artequakt_Generating_Tailored_Biographies_from_Automatically_Annotated_Fragments_from_the_Web (21. January 2024).

Kybartas, Ben / Rafael Bidarra (2016): "A Survey on Story Generation Techniques for Authoring Computational Narratives", *Transactions on Computational Intelligence and AI in Games* 9(3): 239–253.

León, Carlos / Pablo Gervás / Pablo Delatorre / Alan Tapscott (2020): "Quantitative Characteristics of Human-written Short Stories as a Metric for Automated Storytelling". *New Generation Computing* 38: 635–671.

Niehaus, James / R. Michael Young (2014): "Cognitive Models of Discourse Comprehension for Narrative Generation", *Literary and Linguistic Computing* 29(4): 561–582.

O'Neill, Brian / Mark O. Riedl (2014): "Applying Qualitative Research Methods to Narrative Knowledge Engineering", *Workshop on Computational Models of Narrative*, Quebec City: 139–153. https://doi.org/10.4230/OASIcs.CMN.2014.139.

Papalampidi, Pinelopi / Kris Cao / Tomas. Kocisky (2022): "Towards Coherent and Consistent Use of Entities in Narrative Generation". *CoRR* abs/2202.01709. https://doi.org/10.48550/ArXiv.2202.01709.

Pasini, Tommaso (2021): "The Knowledge Acquisition Bottleneck Problem in Multilingual Word Sense Disambiguation", *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*: 4936–4942.

Peinado, Federico / Virginia Francisco / Raquel Hervás / Pablo Gervás (2010): "Assessing the Novelty of Computer-generated Narratives Using Empirical Metrics". *Minds and Machines* 20(4): 565–588. https://doi.org/10.1007/s11023-010-9209-8.

Pérez y Pérez, Rafael (1999): *MEXICA: A Computer Model of Creativity in Writing*. Ph.D. thesis, Falmer: The University of Sussex.

Propp, Vladimir (1928): *Morphology of the Folk Tale*. Leningrad: Akademija.

Ramirez, Alejandro J. / Vadim Bulitko (2012): "Telling Interactive Player-specific Stories and Planning for it: Asd+ passage= past," *Eighth Artificial Intelligence and Interactive Digital Entertainment Conference*.

Reddington, Joseph / Nava Tintarev (2011): "Automatically Generating Stories from Sensor Data," *Proceedings of the 16th International Conference on Intelligent User Interfaces*. New York: Association for Computing Machinery: 407- 410. https://doi.org/10.1145/1943403.1943477.

Reiter, Ehud / Robert Dale (2000): *Building Natural Language Generation Systems*. Cambridge: Cambridge University Press.

Ricoeur, Paul (1976): *Interpretation Theory: Discourse and the Surplus of Meaning*. Fort Worth: TCU Press.

Ryan, James (2018): *Curating Simulated Storyworlds*. Ph.D. thesis. https://escholarship.org/uc/item/1340j5h2 (21 January, 2024).

Sap, Maarten / Anna Jafarpour / Yejin Choi / Noah A. Smith / James W. Pennebaker / Eric Horvitz (2022): "Quantifying the Narrative Flow of Imagined Versus Autobiographical Stories". *Proceedings of the National Academy of Sciences* 119(45). https://doi.org/10.1073/pnas.2211715119

Sharples, Mike (1999): *How We Write: Writing as Creative Design*. Milton Park: Routledge.

Smedley, Ralph C. (1952): *Speech Engineering: 25 Ways to Build a Speech*. Englewood: Toastmasters International.

Wolf, Thomas et al. (2020): "Transformers: State-of-the-art Natural Language Processing". *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations,* Online: Association for Computational Linguistics: 38–45.

Young, R. Michael / Stephen G. Ware / Brad A. Cassell / Justus Robertson (2013): "Planning in Narrative Generation: A Review of Plan-based Approaches to the Generation of Story, Discourse and Interactivity in Narratives". *Sprache und Datenverarbeitung, Special Issue on Formal and Computational Models of Narrative* 37(1-2): 41–64.

Zhang, Min / Juntao Li (2021): "A Commentary of GPT-3 in MIT Technology Review 2021". *Fundamental Research* 1(6): 831–833.

Zook, Alexander / Stephen Lee-Urban / Mark O. Riedl / Heather K. Holden / Robert A. Sottilare / Keith W. Brawner (2012): "Automated Scenario Generation: Toward Tailored and Optimized Military Training in Virtual Environments". *Proceedings of the 7th international conference on the Foundations of Digital Games*: 164–171.