

Visual Story-Writing: Writing by Manipulating Visual Representations of Stories

Damien Masson
Université de Montréal
Montreal, Quebec, Canada
damien.masson@umontreal.ca

Zixin Zhao
University of Toronto
Toronto, Ontario, Canada
zzhao1@cs.toronto.edu

Fanny Chevalier
University of Toronto
Toronto, Ontario, Canada
fanny@dgp.toronto.edu

Abstract

We define “visual story-writing” as using visual representations of story elements to support writing and revising narrative texts. To demonstrate this approach, we developed a text editor that automatically visualizes a graph of entity interactions, movement between locations, and a timeline of story events. Interacting with these visualizations results in suggested text edits: for example, connecting two characters in the graph creates an interaction between them, moving an entity updates their described location, and rearranging events on the timeline reorganizes the narrative sequence. Through two user studies on narrative text editing and writing, we found that visuals supported participants in planning high-level revisions, tracking story elements, and exploring story variations in ways that encourage creativity. Broadly, our work lays the foundation for writing support, not just through words, but also visuals.

CCS Concepts

• **Human-centered computing** → **Interactive systems and tools**; **Visualization theory, concepts and paradigms**.

Keywords

creative writing, visualization, creativity support, LLM, AI

ACM Reference Format:

Damien Masson, Zixin Zhao, and Fanny Chevalier. 2025. Visual Story-Writing: Writing by Manipulating Visual Representations of Stories. In *The 38th Annual ACM Symposium on User Interface Software and Technology (UIST '25)*, September 28–October 01, 2025, Busan, Republic of Korea. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3746059.3747758>

1 Introduction

Creative writers manage many moving parts, from character arcs and causal chains to spatial coherence and narrative timing [4, 57]. Keeping track of everything is challenging, and it gets even more complicated when experimenting with story ideas [83]. Modifications require careful planning and multiple rounds of edits.

Consider a task as simple as changing the location of a character in a story: the cat no longer goes to the barn but instead wanders about the lake. Simply replacing all instances of “barn” would cause inconsistencies unless descriptions and actions pertaining to the barn are also updated to reflect the new context of the lake. Moreover, the fact that the cat wanders about the lake does not mean

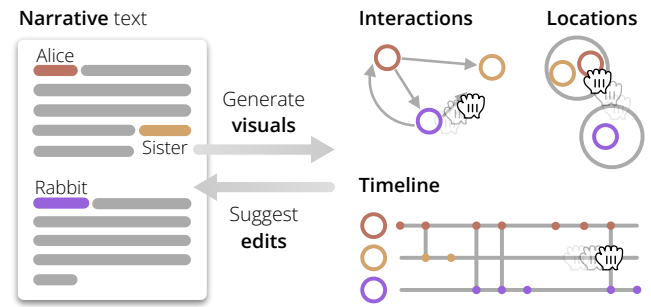


Figure 1: Visual story-writing supports writing by generating visualizations that help review the story and that can be manipulated to suggest edits to the narrative text.

that the barn needs to be removed entirely from the story. Instead, before revising the text, writers must identify references to the barn and tediously review the other characters’ locations to maintain spatial and semantic consistency. This busywork is prone to errors, and writers often resort to creating external documents, such as maps, spreadsheets, and timelines, that they painstakingly update to keep track of characters, events, and locations [1, 3, 54, 66, 83]. Large language models (LLMs) could help, but expressing specific intents with a prompt is difficult [49, 81], and trying to explain which barn and which cat to edit will leave much room for misinterpretation. Further, the suitability of the changes made by the generative model could be difficult to verify.

This challenge stems from the mismatch between the reasoning and representation spaces: in our *barn* example, the writer does spatial reasoning with a textual representation. A better approach would be to use a spatial representation, such as a map of the story world, where changing the location of a character becomes as simple as moving it on the map. This same argument applies to all other aspects of a story, such as time, events, and characters, for which we posit superior representations can be designed.

Therefore, we propose using representations that match the reasoning space. We define **visual story-writing** as the use of visual representations of story elements as both a reviewing tool and an input medium for expressing writing intents (Figure 1). This complements the writing workflow of writers with various experience levels. For example, an experienced writer may prefer editing the story’s text, in which case the visual representations updates, offering a visualization to help the writer track story elements. Alternatively, a beginner writer could manipulate the visual representations through direct manipulation to explore story variations.

We articulate a framework grounded in narratology to design new visual representations of stories (§3). Applying our framework, we explore the potential of visual story-writing through a prototype



This work is licensed under a Creative Commons Attribution 4.0 International License.
UIST '25, Busan, Republic of Korea
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2037-6/25/09
<https://doi.org/10.1145/3746059.3747758>

software (§4) with three illustrative visual representations: (1) a diagrammatic view of the entities in the story and their relationships, allowing to add, remove, and edit characters and add new actions between them; (2) a spatial view of the entities in the story, allowing to add new locations and move entities; and (3) a timeline of the events in the story, allowing to reorder the narrative, quickly find specific scenes, and precisely modify the entities and locations of the selected scenes by leveraging the other views.

We validate this approach through two user studies looking at the different components of the writing process [25]. Results from our first study revealed that visuals helped plan high-level revisions, search, and promote critical reflection over having only text. Results from our second study show the potential of visual story-writing to express editing intents and help writers explore story variations in a way that supports creative expression. This points to the potential of visual story-writing and encourages further exploration.

2 Related Work

Visual story-writing draws inspiration from work on visualizing stories and efforts to give alternative representations to edit content.

2.1 Visualizing Stories

Visual representations of stories and story patterns have a long-standing presence in writing, with staple examples including Freytag's pyramid story [26] and abstract diagrams capturing the "shape of the story" [22, 60, 64]. Such visual representations are not only useful for teaching the craft of storytelling but are also considered useful aids to support the writing process. Referring to the visual representations, screenplay writer E. Williams says: "*I find story shape to be a great tool [...] to make sure my story holds together*" [78].

There exist many techniques to visualize stories, spanning from classical charts, like word clouds, to more artistic visualizations like calligrams [10]. Often, these visualizations attempt to highlight a specific facet of the story (e.g., emotions [45], motion [20], fortune [19, 22], or topic recurrence [84]). For example, Storyline visualizations [52, 72] depict the evolution of the relationships between characters through timelines. StoryCake [62] uses a hierarchical plot to visualize the structure of nonlinear stories. Story Curves [40] visualizes non-linear narratives by showing both the order in which the events are told and the story's chronological order. StoryPrint [76] shows the scenes and character emotions through circular timelines. Each of these visual representations presents the ability to support readers, writers, and analysts alike in appreciating or even analyzing the stories from a specific lens.

Other story visualizations have been used in tools to help with script and story writing. For example, CARDINAL [46] helps screenwriters visualize their script through different views, including a visualization of the interactions between characters and animated 2D and 3D views to play out the scene and replay character movements. Similarly, Portrayal [31] attempts to help writers develop their characters by extracting and presenting information related to the characters in interactive visualizations. Our work seeks to inform the design of story visualizations by proposing a new framework. This framework helps generate representations that might serve as a medium to edit the story, a topic we discuss next.

2.2 Editing Using Alternative Representations

Visualizations are increasingly used to not only encode data but also collect and modify new data [12]. In our work, visual representations can be edited to modify the narrative text in addition to exploring and analyzing it, making it an input visualization.

One domain which has embraced the use of input visualizations is computer programming. For example, visual programming [13, 53] allows generating programs by manipulating elements in more than one dimension, code projections [28, 41] augments programming languages with tools and widgets that are alternative representations of fragments of code, and bi-directional programming [29, 51] allows manipulating the output of a program to edit the program itself. Similar to our work, the motivation behind these approaches is that some tasks are easier done using different representations: it is easier to find a colour using a colour picker than by typing its hexadecimal code, and it is easier to move an object through direct manipulation [67] than by typing coordinates.

In the creative writing domain, the idea of manipulating projections or visual representations to edit stories is largely untapped despite the reliance on visual representations to capture story concepts [22, 26, 83]. Instead, most writing support tools ask users to manipulate narratives using text, typically through prompts or suggestions [15, 16, 24, 42, 65, 80]. Some exceptions include the proposal of Kempen et al. [39] to manipulate the syntax tree of a sentence to restructure it or the proposal to directly drag the words to reorder them [2]. Beyond syntactic level modifications, Dang et al. [23] proposed a system that allows reorganizing, merging, and removing paragraphs by manipulating textual summaries. Moving even further from text as an input medium, the machine learning task of "visual storytelling" consists in generating a story from a sequence of images [6, 32–34, 74]. Similarly, TaleBrush [19] leverages a canvas where sketching allows modulating the fortune of characters in a story, whereas Textoshop [47] offers features inspired by drawing software such as tone pickers and text layers.

Closest to our approach are systems that provide an editable alternative view of the text. For example, VISAR [82] generates a visual outline of an argumentative text, which can be edited to change the structure of the argument. Toyteller [20] generates story text based on character motions, as if the characters were toys. Similarly, XCreation [79] supports creativity and cross-modality edits through the use of clip-arts and a relationship diagram as an intermediate representation between text and images.

Our work contributes to this line of research by formalizing *visual story-writing*. The concept is rooted in the same principles as visual programming and code projections, positing that elements of a story become easier to understand and manipulate when presented in alternative forms. Specifically, we propose a framework grounded in structural narratology theories to inform the design of visualizations and extend the concept to all story elements, including events, like XCreation [79], but also space and time.

3 A Framework for Story Constructs

To inform the design of visual representations that support the visual story-writing workflow, we articulate a generative framework. Our framework builds upon narrative theories by structuralists like Gérard Genette [27] and Mieke Bal [4, 5]. Elements of stories

outlined by structuralists [4, 5] match what we observed in existing story visualizations, such as organizing story elements based on time [52, 76], space [46], or a combination thereof [35]. Story elements do not work in isolation. Rather, many important aspects of stories stem from the intricate weaving of these elements, such as the movement of characters across locations over time.

We propose an *operator* mechanism to combine elements into such meaningful composites. Our operators are applied sequentially to *story elements* to form a story construct that can then be represented and manipulated. This framework can describe story constructs in existing visualizations and also generate new ones.

3.1 Story Elements

Story elements are the fundamental building blocks of story constructs found in story visualizations. Within narratology, the early Russian formalists separated the chronological order of the story (*fabula*) from the order of the plot or how it is represented to readers (*syuzhet*) [4, 43]. Structuralists like Genette and Bal borrowed these concepts to analyze narrative structurally. Bal [4, 5] describes a *fabula*, the factual events within a story, to consist of four main elements: actors¹, time, location, and events. The *syuzhet*, or how the story is told, also consists of four elements: characters, temporality, space, and focalization. See table 1 for examples.

Table 1: The eight story elements of our framework, extracted from *fabula* and *syuzhet* elements described by Bal [4, 5]

Element of <i>Fabula</i>	Element of <i>Syuzhet</i>
Actor (e.g., hero, villain)	Character (e.g., Alice)
Location (e.g., Alice’s house)	Space (e.g., home, eerie)
Time (chronological)	Temporality (narrated order)
Event (what happens)	Focalization (point of view)

Although elements of *fabula* and *syuzhet* may seem similar, they refer to distinct concepts. For instance, characters are concrete entities (e.g., Alice, the white rabbit), whereas actors are abstract roles or functions characters can take (e.g., hero, villain). Time corresponds to the chronological timeline of events, whereas temporality corresponds to how those events are narrated (e.g., through flashbacks, ellipsis). And locations are concrete, physical places (e.g., Alice’s house, a forest), whereas spaces refer to the setting and the narrated representation of these locations (e.g., home, eerie). We describe visual story-writing as the use of visuals that represent and manipulate story constructs, i.e. composites of these eight elements.

3.2 Operators

Our operators are functions that combine story elements (operands) into higher-level story constructs. They can be chained to create more elaborate constructs. Based on existing story visualizations, we derived four binary operators acting on operands x and y ²:

- **POSITION** ($x \textcircled{P} y$) places elements of x based on the story element y (Figure 2.(P)). For example, positioning *time* by *locations* results

in a map of the movements in the story. Note that the only valid y operands are *location* and *space*.

- **ASSOCIATE** ($x \textcircled{A} y$) adds the elements of y and associates them to the elements of x (Figure 2.(A)). For example, associating *time* with *focalization* creates a timeline of point of view.
- **CONNECT** ($x \textcircled{C} y$) adds edges between elements of x , according to that of y (Figure 2.(C)). For example, connecting *characters* by *events* creates a graph of character interactions.
- **UNFOLD** ($x \textcircled{U} y$) duplicates and organizes elements of x based on that of y (Figure 2.(U)). This essentially duplicates the x elements according to each element of y . For example, unfolding *locations* by *characters* creates a list of locations visited by each character.

3.3 Describing Existing Story Visualizations

We test our framework by describing existing story visualizations.

- *Storyline* [52] shows character interactions: start with *time*, UNFOLD by *characters*, and CONNECT by *events*.
- *StoryCurve* [40] shows nonlinear narratives: start with *time*, UNFOLD by *temporality* to visualize the non-linear narrative, ASSOCIATE with *locations* (using colours) and ASSOCIATE with *events*.
- *StoryPrint* [76] shows scenes, character presence and emotions: start with *time*, UNFOLD by *characters*, and ASSOCIATE with *events* or other elements (StoryPrint also encodes emotions, which goes beyond the scope of structural narratology).
- *Geo-Storylines* [35] shows locations of characters over time: start with *Storyline*, POSITION by *locations*. The map glyph variation: start with *locations*, POSITION by *location*, UNFOLD by *time*, and ASSOCIATE with *characters*.
- “*All fights from Dragon Ball Z*”³ [21] shows the battles between characters within a saga: start with *events* (only the fights), UNFOLD by *time*, and CONNECT by *characters*.

3.4 Generating New Story Constructs

Figure 2 shows examples of chaining operators to create more complex story constructs. The number of unique combinations of operators and elements is vast. Our goal is not to argue which specific story construct is optimal for writers nor which visual design is best to represent these constructs—we leave it to future work to explore variations within these categories (and beyond). Instead, our work focuses on the potential of visual story-writing, i.e. the use of alternative representations of stories, as an input medium for expressing intent. As such, we designed three simple visual representations of story constructs interconnected through interactivity [56, 75]: a timeline of entities and events (start with *time*, UNFOLD by *characters* and CONNECT by *events*); a graph of actions and entities (start with *characters* and CONNECT by *events*). And a view of the locations (start with *characters* and POSITION by *locations*).

There are other important components to a story, such as character traits and emotions, but in Bal’s narratology model [4, 5], these are properties of the story elements. One way to represent them within the framework is as an attribute of a story element.

¹Actors refer to “agents that perform actions. They are not necessarily human” [4]. To avoid confusion, we use the term *entity* instead in our prototype system.

²Note that our operators are noncommutative. Thus, while y is always one of the eight story elements, x can be a composite of elements resulting from a chain of operations.

³A non-academic example

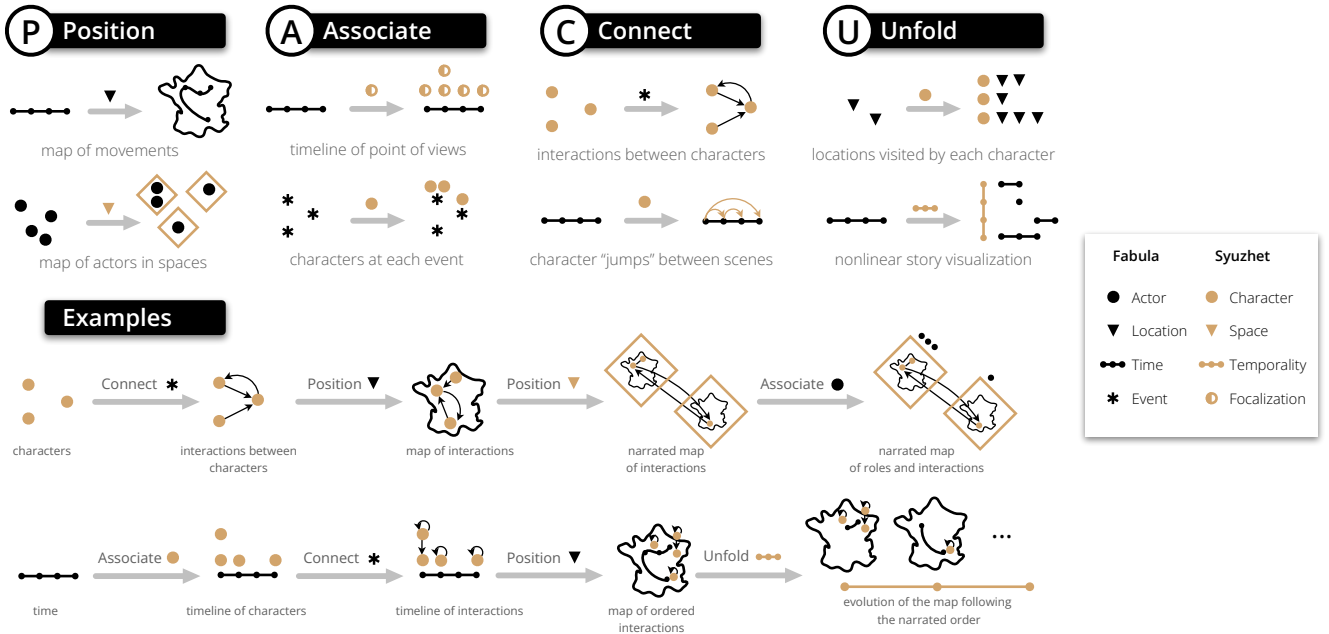


Figure 2: We propose a framework grounded in narratology to generate story constructs that can support visual story-writing. First, select an element from the fabula or syuzhet and then apply operators sequentially: (P) POSITION to place elements spatially; (A) ASSOCIATE to add new elements; (C) CONNECT to add edges; and (U) UNFOLD to duplicate and organize elements.

3.5 Recommending Interactions and Visuals

The framework can be further expanded to prescribe interactions and visual representations. For example, applying the POSITION operator suggests that the placed elements can be dragged to express a change of location; the ASSOCIATE operator suggests a detail-on-demand type of interaction to view the associated elements; the CONNECT operator suggests an interaction to remove and create links between elements; and the UNFOLD operator suggests elements can be dragged to be reordered or reassigned. Similarly, the framework could also prescribe visual representations by mapping operations to specific visual channels.

4 The Visual Story-Writing Prototype

To test the potential of visual story-writing, we developed a prototype system, shown in Figure 3, with three visual representations informed by our framework: a diagrammatic view of the entities and actions, a spatial view of the locations and entities, and a timeline of the events (section 3.4). The focus of the prototype was to help writers explore variations and perform spatial or temporal edits via visual story-writing. As such, the system allowed specifying edits through direct manipulations of the graphical elements [38, 68] and followed the design principles of creativity support tools to support exploration [69]. Below, we detail the system design and features (marked with **READ** when the feature helps review the story, **WRITE** when it helps edit the story, and **R&W** when it helps for both editing and reviewing). All screenshots were taken using the system on the first three paragraphs of the story *Alice's Adventures in Wonderland* by Lewis Carroll.

4.1 Entities and Actions View

This representation shows all the entities in the story as nodes with their name and an emoji representing them (Figure 3.b). The actions between the entities are represented as directed edges with a label describing the action in one or two words. By default, the nodes are arranged using a force layout to prevent overlaps, but users can rearrange the position of the entities by dragging them.

4.1.1 [R&W] Editing the Traits of an Entity. Selecting an entity opens a menu with traits similar to a list of personality traits which would be present on a character sheet (e.g., “curious”, “adventurous”, “chatty”). The intensity of each trait is rated on a scale of 1 to 10. Moving the slider beneath the trait results in editing the text of the story to reflect the new intensity of the trait (Figure 4.a).

4.1.2 [WRITE] Adding and Removing an Entity. Double-clicking on the canvas opens a text input box, which, once filled, creates an entity with the given name at the location of the mouse pointer. Selecting an entity and then pressing “delete” or “backspace” on the keyboard removes the entity. This results in updating the text of the story so that the entity is effectively removed (Figure 4.b).

4.1.3 [WRITE] Adding and Removing Actions between Entities. Creating an edge between two nodes opens a text input field to enter a new action. The direction of the edge indicates the source and target of the action. Connecting a node to itself means the action is for the same character. This helps represent actions done by an entity with no target (e.g., walking or thinking). Double-clicking the label of an action edits it. Similarly, selecting an action and pressing “delete” or “backspace” removes it. As soon as the edge is created, edited, or deleted, the story text is updated to reflect the change (Figure 4.c).

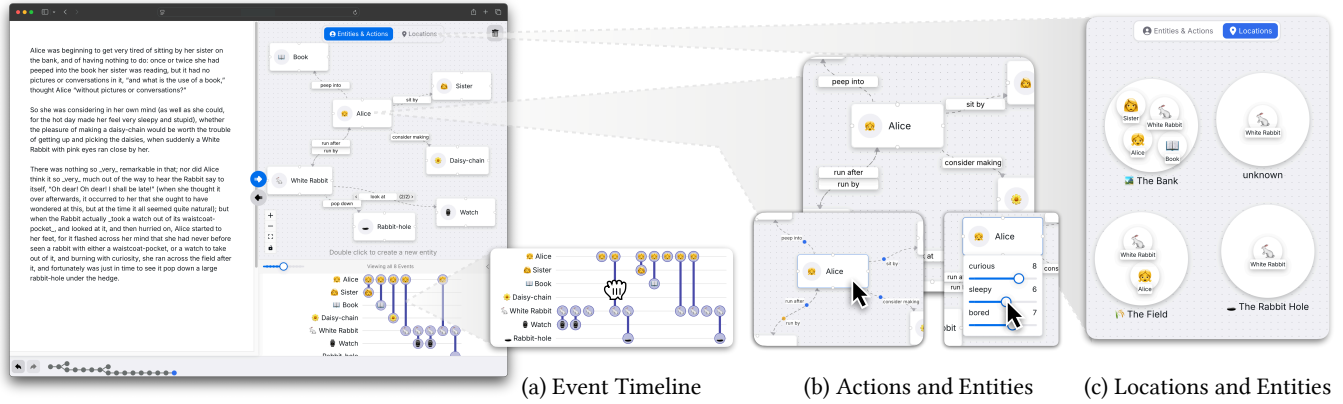


Figure 3: The visual story-writing system used on the first three paragraphs of *Alice's Adventures in Wonderland*: (a) the event timeline allows reordering the events; (b) the actions and entities view allows editing the characters' traits and adding or removing entities and actions; (c) the locations and entities view allows moving entities and creating new locations

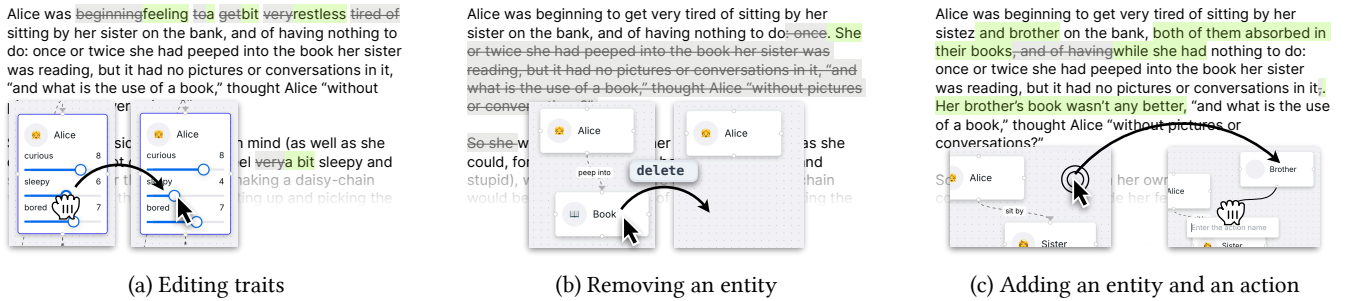


Figure 4: The entity view allows modifying entities: (a) selecting an entity opens up character traits that can be modified to edit the story; (b) an entity can be removed in which case the story is edited so that the entity and its actions are removed; and (c) a new entity can be created and connected to other entities to create actions

4.1.4 READ Overlapping Edges and Animations. To declutter the view, overlapping edges are grouped so that only the first action is shown alongside a counter. Users can cycle through the events by clicking the next and previous buttons. Alternatively, hovering over an entity shows animated dots coming and going from the entity, with the name of the action beside them.

4.2 Locations and Entities View

The location view is a spatial representation accessed through a tab (Figure 3.c). It displays all the locations in the story as nodes with their name and an emoji representing them. Similar to the entities and actions view (Section 4.1), the view shows all entities as smaller nodes, but their position depends on their order of appearance in the story. Entities that go to different locations in the story are duplicated to be represented in all their locations.

4.2.1 WRITE Adding a Location. Double-clicking anywhere on the canvas opens a text input field to create a new location (Figure 5.a). Once the location is created, entities can be moved to it.

4.2.2 WRITE Moving an Entity. Entities can be dragged and moved around the view. If the entity is released on top of a location, then that entity is moved to that location (Figure 5.b-c). Otherwise, the

entity goes back to its original location. Once the entity is moved, the story text is updated to reflect the change in location.

4.3 Timeline of Events View

The timeline view is a temporal representation showing the events in the story represented as vertical lines with the emojis of the entities involved in the event on either side of the line (Figure 3.a). These lines are organized one after the other based on the order in which they are presented in the text (*temporality*). The timeline view is always visible. It allows selecting one or multiple events, in which case the other views are updated to fade out the entities and locations not involved in the selected events. Selecting events also forces subsequent modifications to impact only the selected events.

4.3.1 R&W Selecting Events. The timeline also allows finding specific scenes and selecting them to allow precise editing. Hovering over events highlights the corresponding sentence in the story (Figure 6.a) and the corresponding entities and locations in the event and location views. Clicking selects the event, and dragging selects multiple events. Once events are selected, future visual edits will target these selected events. For example, selecting an event and then moving an entity in the location view updates the story text so that the entity moved exactly during the selected event.

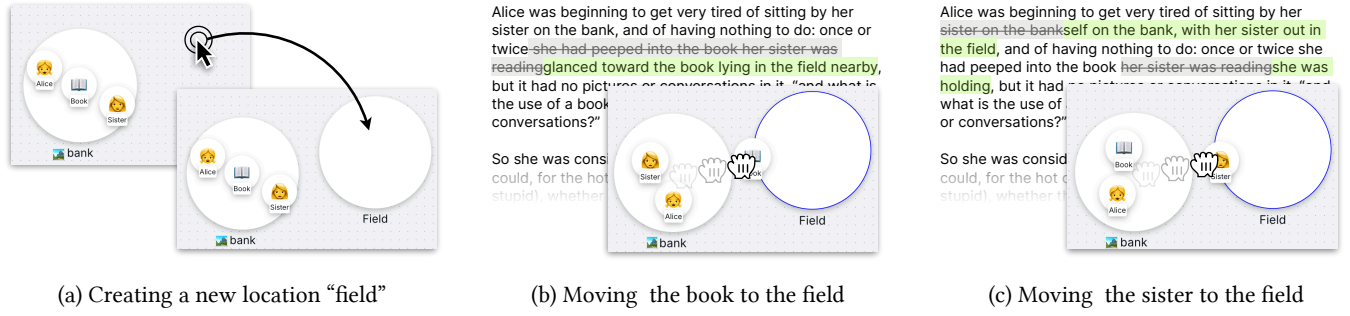


Figure 5: The location view allows moving entities: (a) a new location “field” is created through a double click; (b) the entity “book” is moved to the field through a drag-and-drop; and (c) the sister is moved to the field instead

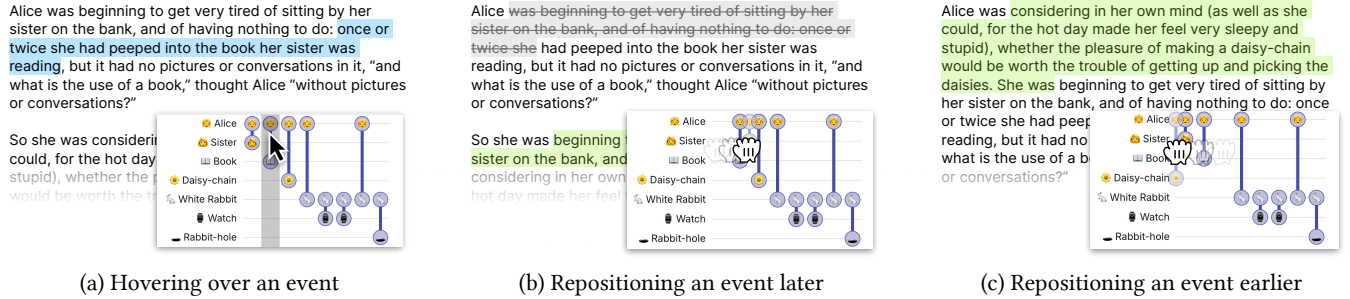


Figure 6: The timeline view allows finding and reordering the events in the story: (a) hovering over an event highlights the corresponding text; (b) the event about Alice sitting beside her sister is moved later in the story; (c) the event about Alice considering making daisy-chains is moved earlier in the story

4.3.2 [WRITE] Reordering Events. Dragging an event slides it horizontally (Figure 6.b-c). Once moved, the story text updates so that the event happens when indicated in the timeline.

4.4 Bi-directional Editor

The rest of the interface consists of a text editor on the left, a history tree at the bottom, and interface buttons at the centre to refresh the visuals or rewrite the story (Figure 3).

4.4.1 [READ] Highlighting Visual Elements and Sentences on Hover. Similar to how hovering over events in the timeline highlights the corresponding sentences, hovering over the sentences in the text editor highlights the corresponding events, entities, and locations in the different views. This helps find scenes. Similarly, placing the caret in one sentence will act as if the events in that sentence were selected in the timeline. As such, subsequent edits to the visual representations will be restricted to that sentence.

4.4.2 [READ] Updating the Visual Representations from the Text. When the text is updated manually, the visual representations might become desynchronized, showing the story as it was before the modification. In this case, the refresh button becomes blue to indicate that the visual model might be outdated. Clicking it re-extracts the information from the text and refreshes the views.

4.4.3 [WRITE] Writing the Story from the Visual Representations. By clicking the “refresh from visuals” button (arrows between the two views in Figure 3), the story is regenerated from scratch using only the visual representations as reference. The regenerated story

preserves the same events, locations, and entities, but everything else is rewritten. This allows for specifying the skeleton of a story and quickly exploring different ways of phrasing it.

4.4.4 [WRITE] Track changes. Changes made by editing the visual representations are tracked, such that removed text is struck while additions are highlighted in green. This allows writers to locate and review the changes before accepting (or rejecting) them.

4.4.5 [WRITE] History Tree. To make edits easily reversible and facilitate the exploration of alternatives, the system implements a history tree to store different versions of the story and return to them at any time. When a modification happens, be it by editing the visual representation or editing the text, a snapshot of the story and visual representations is added to the tree. When the writer selects a previous snapshot, the story is reverted to that snapshot. If, after selecting a previous snapshot, a modification is made, then a new branch is created. This allows writers to navigate between branches and versions by clicking on the nodes of the history tree.

4.5 Implementation

We implemented the visual story-writing system using TypeScript and React [63] for the interface, NextUI [55] for the graphical components, Slate.js [70] for the text editor, and the OpenAI library for prompting large language models [58]. Modifications to the visual representations are turned into engineered prompts sent to OpenAI’s “GPT4-o” model. A live demo and the source code are accessible online: <https://github.com/m-damien/VisualStoryWriting>

4.5.1 Extracting Entities, Locations, and Events. In our tests, the naïve approach of asking an LLM to extract the necessary information results in slow and incomplete extractions. Instead, we first ask the model to extract the entities and locations in the story. We then split the text into sentences and parallelize requests by giving the model the whole story but asking it to extract only the information within each sentence (appendix A.1). As such, a text with 30 sentences sends 30 smaller prompts in parallel. This leads to faster and more exhaustive extractions. It also helps associate sentences with events. All requests use a structured JSON output.

4.5.2 Engineered Prompts to Edit the Story. The system relies on engineered prompts such as splitting requests to make them faster, giving the model only a subset of the text to force localized edits, and giving the model the previous abstract representation of the story and the new one to help it understand the edits. See the appendix and/or the source code for more details.

5 Study 1: Planning Using the Visualizations

In our initial user study, we assessed how story visualizations aid the planning and reviewing phases of writing [25]. To ensure the study focused solely on visualization, the prototype only supported the **READ** features. We then compared this read-only version to a baseline text-only interface, asking participants to answer high-level planning and reviewing questions. Because no text had to be produced, we recruited participants with various creative writing experiences to reflect the many potential target users of the tool

5.1 Method

5.1.1 Participants. We recruited 12 participants with various writing experiences from our social network and a mailing list of writers (age range: 20 to 37, M=25; 7 self-identified as male, 4 as female, and 1 preferred not to disclose). On a 5-point Likert scale ranging from 1-never to 5-always, they reported how often they write (Mdn=4), how often they use visuals while writing (Mdn=2), and how often they find visuals to help when writing (Mdn=3). For visual aids, some mentioned using timelines (N=5) and diagrams (N=4). They also reported different writing contexts, including academic (N=9), professional (N=7), and creative (N=6) writing.

5.1.2 Apparatus. Participants joined the study remotely via the Zoom conference software and accessed the experiment website using their own computers. The sessions lasted 1 hour and were audio and screen-recorded. Additionally, interactions with the experiment website were also recorded (e.g., features used, commands executed, and answers to questions). Participants received a CAD\$20 gift card for their time. Our institutional ethics board approved the study.

5.1.3 Tasks. Tasks simulated what writers ask themselves when reviewing their craft. We extracted questions from best practices on revising fiction [44, 57] and then revised them with a professional creative writing instructor. This led to two open-ended questions per story aspect: characters (“*Could characters be combined without changing the outcome of the story?*” and “*Is any character too passive?*”); locations (“*Are there any locations which could be removed?*” and “*Are there moments where the spatial logic is broken or unclear?*”); time and rhythm (“*Is there a large gap between two actions that make the story progress?*” and “*Is there a scene you think could take place*

later/earlier in the story?”); and focalization (“*If told from another character’s perspective, how would the story change?*” and “*What does the main character mention that no other character would?*”).

5.1.4 Stories. We picked two short stories from the *Tell Me A Story* dataset [37], which consists of human-written stories. Stories were screened to be of similar length (~700 words) and featuring more than two speaking characters and location shifts.

5.1.5 Procedure. After signing the consent form and completing a demographic questionnaire, participants started with one of two conditions: visual story-writing or text-only. For the visual story-writing condition, they watched a one-minute tutorial and then tried the tool. For both conditions, they read the story fully before starting tasks. Participants had to think and answer questions aloud within five minutes, and then explain how they came up with the answer. After each condition, they completed a raw NASA TLX. At the end, they engaged in a semi-structured interview to share impressions, preferences, and strategies. Tasks, stories, and conditions were counterbalanced across participants.

5.2 Results

Participants were able to complete the tasks without difficulties in both conditions. However, how they accomplished the task varied. Themes are based on their relevance to our research questions, as per an interpretivist approach to reflexive thematic analysis [11].

Visualizations helped confirm, analyze, and explore. All participants mentioned using the visualizations to confirm their intuitions. Visualizations gave the “*reassurance that my hunch is correct, and I’m not missing anything when answering the questions.*” (P6) as they felt “*it was mostly a confidence thing where I felt more sure in what I was saying*” (P2).

The visuals also helped “*have conversations [about] deeper analysis [...] and narrative structures*” (P4). For example, participants used visuals to “*remove this location simply because nobody was there*” (P2), “*identify who can know about each other*” (P5), “*figure out if [an event] was playing a part in the story*” (P6), and see “*how the entities interact with each other*” (P12). P8 added it could be particularly useful “*if you were editing or doing some kind of work on a piece of text where you’re like [...] has everyone been given as much time to speak as I want them to.*”

Sometimes, visuals led to discoveries and reflections. For example, P11 mentioned “*it kind of dawned on me by looking at the chronology that [...] we actually don’t really get [character]’s motivations*”. Similarly, P6 used the location view to “*reflect on if [location] were to be changed to a different location, would that really have a significant impact on the story? And I realized that’s probably no.*” P7 was “*paging through the timeline and seeing how the characters jumped around and kind of weren’t even represented [...] it actually showed me [...] that it’s like, oh, a lot of moments in this story are sort of disconnected from one another*”. P3 added “*it does make me curious [...] how would a story that features less locations look like*”.

Visual-driven search vs. skimming. Participants praised the visuals for their help finding passages. With the visuals, “*it was way more tidier, similar kind of information were in similar places [...] or visually looked similar*” (P5) and “*if I just click [an entity], it’ll just highlight the parts where this entity is being talked about*

[...] I don't have to look at other useless portions" (P12). In particular, participants contrasted this to their workflow with text, where *"if I just remember, like vaguely, [...] then I can probably pattern-match to the visualization faster than trying to figure out like, what is [...] the keyword I could search to find the sequence of text."* (P3). In the text condition, participants mentioned how they *"tried to hunt for the dialogues [...] where there were quotations, and tried to see how many of them were said by [character]"* (P4), and were finding locations by *"looking for capitalized nouns [...] so there's always the chance that there was something else [...] like, oh, the coast!"* (P7).

Impact on cognitive load. Results were inconclusive. Wilcoxon signed-rank did not find a significant difference for mental demand ($p=.21$), physical demand ($p=1$), temporal demand ($p=.49$), performance ($p=.38$), effort ($p=.13$), nor frustration ($p=.16$).

Participants' comments were also mixed. Some mentioned how the visualization reduced their perceived cognitive load because *"[with text] I was trying to visualize the things in my head. But with [visuals], it kind of takes away that mental strain"* (P12) and *"it has all of the characters and locations listed out. If I were to use command F, I would have to put in whatever I can still remember, which oftentimes might not be the complete set"*. Others appreciated being able to focus on other aspects because *"the visualization [...] lets you offload some of the more concrete stuff so that you could explicitly focus on things like character motivations, whereas I found, for the text-based one, I was constantly reskimming in order to refresh this mental schema"* (P11). However, others mentioned that the system added to their cognitive load because *"trying to learn [the tool] [...] almost added this extra layer where I wasn't engaging with the text"* (P7).

Mental model matches and mismatches. The visuals made sense to some participants because *"I feel like my brain does this, [...] but maybe not this clearly"*. In fact, P9, who had the text condition after the visual condition, mentioned how they *"laid down a timeline in [their] head, something akin to what we've seen in [the visual condition]"* to accomplish the tasks.

Yet, sometimes, there was a clear model mismatch. P8 was the most critical, explaining *"I'm sure there's people who just think like this, and this would be a really useful tool for them, because it is an extension of the way they process information. But for me, it's like really not"*. In fact, P8 mentioned often using visuals when writing but explained their process was different: *"I will write those scenes on sticky notes and then I can start to arrange them in an order that gives me a plot [...], and then I sometimes will use [...] picture boards [...] to remind me of a visual idea or concept"*.

Some other times, the mismatch happened at the semantic level: *"calling it an event for me makes it feel like there's a plot line that's happening. But it wasn't really that."* (P2) and *"the timeline is very fine grained [...] I wasn't looking at fine-grained scenes [...] but just in terms of more large-scale scenes"* (P3).

Complementarity of text and visuals Participants viewed the visualization as complementing the text because *"you still need the text, the words, to fully appreciate the story"* (P1). In fact, when asked how they did the task, participants reported using the views for 74% of tasks, their memory for 46%, and reading for 43% (adds up to more than 100% because two or more ways could be used). One reason is that the visuals pointed to text that then had to be

verified. P4 explained *"I was using text for the verification, because there were cases when the visualization was capturing something that could be potentially misleading"*. Another reason is that the visualization could not capture everything. In particular, the visuals only represented explicit information: *"It is really helpful in terms of managing kind of explicit things, like characters, entities, locations [...]. But it doesn't quite capture things like motivations"* (P11).

6 Study 2: Editing and Free-form Writing

The vision of visual story-writing is to complement the existing writing workflow. As visuals have shown to be helpful for reflecting on the text, we further explore how visual story-writing can assist writers during the "translating" and revision processes of the cognitive model of writing [25]. To explore this, we recruited experienced creative writers so that they could give us insights on the possible impact the tool would have on their writing workflows. This exploratory study used the fully-featured prototype and comprised two parts: the first part isolates the use of the three different views in a controlled setting, and the second part is a free-form creative writing task. We were interested in the following questions: Are the visual representations understood and helpful? Can people express their editing intents using visual story-writing? What is the impact of visual story-writing on exploration and creativity?

6.1 Method

6.1.1 Participants. We recruited 8 creative writers from our social network and word-of-mouth (age range: 20 to 31, $M=24$; 7 self-identified as female and 1 as male). None participated in the first study. All had several years of experience with creative writing, including personal writing ($N=7$), fiction ($N=4$), scriptwriting ($N=2$), and poetry ($N=2$). On a scale from 1-never to 5-always, they reported how often they use visuals (e.g., mind maps, timelines) ($Mdn=2$), how often they find visuals to help ($Mdn=3$), and how frequently they use AI services ($Mdn=1$) during writing. They reported prior use of visuals to help their writing, including mood boards ($N=4$), timelines ($N=3$), mind maps ($N=2$), and reference imagery ($N=2$). Participants are labelled W1 to W8 in the rest of the paper.

6.1.2 Apparatus. Same as 5.1.2 except the sessions were about 45 minutes long and participants received a CAD\$15 gift card.

6.1.3 Tasks. The first part had three task blocks for each view: (1) *entities and actions view tasks* required removing an entity then adding a new entity and action involving the main character; (2) *location and entities view tasks* required moving an entity to an existing location, creating a new location, and move another entity to it; (3) *timeline view tasks* required reordering events. Instructions were phrased as questions, forcing participants to engage with the story. For example, a timeline task asked *"What if Jack loses his hat when he flies into the sky?"*. This required participants to understand the story's events and move the action to the right location. Tasks were counterbalanced across participants. The second part involved a single writing task using a writing prompt.

6.1.4 Procedure. After completing the consent form and demographics questionnaire, participants watched a 30-second video introducing the tool. Before a task block, they watched a short

Scale	Avg. Score (SD)	Avg. Counts (SD)
Enjoyment	15.73 (1.84)	3.50 (1.00)
Immersion	12.18 (3.04)	<u>3.13</u> (1.76)
Results Worth Effort	15.28 (2.26)	3.00 (1.50)
Exploration	13.48 (2.91)	2.88 (1.27)
Expressiveness	13.70 (2.86)	2.50 (0.50)
Overall CSI Score	71.50 (21.12)	

Table 2: Results for the Creativity Support Index (CSI) sorted based on the average factor count (i.e., the factors considered most important by participants). The highest value is in bold, and the second highest is underlined. Like previous work [17, 71], we omit the *Collaboration* factor to avoid confusion.

video demonstrating the relevant features. After a block, participants rated their perceived success on a 5-point semantic differential scale and answered five questions related to usability, understandability, and workflow integration. Following all tasks, participants engaged in a free-form exploration to continue a story with no instructions other than to “explore what could happen next.” At the end, participants completed an exit questionnaire (including the creativity support index (CSI [18]) and the raw NASA TLX) and engaged in a semi-structured interview to gather their experience and perceived advantages and disadvantages of the system.

6.2 Results

Participants completed the tasks and the free-form exploration without major difficulties. Below, we summarize the quantitative findings using descriptive statistics. Following recommendations for similar study designs [48, 85], confidence intervals use the studentized bootstrap with 10,000 repetitions. Themes are based on their relevance to our research questions, as per an interpretivist approach to reflexive thematic analysis [11].

Finding scenes and keeping track of events, entities, and locations. As a search tool, the results of this study corroborated those of the first study. Most participants agreed that “The visual representation was easy to understand” for all three views (Figure 7). W5 appreciated that you can “click on a person, click on a location, and then see stuff pop up” instead of “having to look through [...] many paragraphs [...] or doing keyword search”. W4 explained “Just having a visual representation helps, you know, just with the different perspectives”. W3 added “Moving around the timeline was a good way of, kind of figuring out different events”. W1 and W5 also mentioned how they usually struggle to keep track of everything while writing their stories and that the system could help them. For instance, W5 details that currently they “have booklets everywhere, and [they] have it annotated weird [...] it’s a mess and this is so much easier. You got like, you know, your locations, your people [...] keeping track of things. Organizing who’s where, what’s going on, being able to scroll through, like highlighting part of the text shown on the timeline [...] making it easy to understand”. This was also reflected in participants reporting low mental demand on the NASA-TLX ($M=7.49$ $SD=1.95$).

Specifying temporal, spatial, and entity-related edits. On a 5-point semantic differential scale from 1-unsuccessful to 5-successful,

participants rated their success $M=4.25$ (95% CI: [3.8, 4.6]) at accomplishing the tasks. They were most successful at the task about adding and removing entities ($M=4.5$, 95% CI: [3.8, 5]), changing the locations of entities ($M=4.25$, 95% CI: [3.4, 4.8]), and reordering events ($M=4$, 95% CI: [2.8, 4.7]). Participants used an average of 2.42 visual edits to accomplish a task. In the free-form part, the most common visual edits were adding actions (36%), editing actions (26%), and editing traits (19%). Specifically, participants appreciated that the edits produced text as a good starting point. W2 mentioned “If I add in an extra character, it gives you just enough text to build that story with that new character”. W4 appreciated that it fixed inconsistencies when editing locations: “the fact that it also updated the sentence where the hat flew off his head that is extremely useful because that’s the sort of inconsistencies that can easily creep in when you’re making modifications like that”. Similarly, W6 expressed, “I really liked the ability to create relations between characters”.

On the NASA-TLX, they also reported high-performance ($M=7.49$ $SD=1.95$) for low effort ($M=3.39$, $SD=2.17$), frustration ($M=2.5$, $SD=1.2$), temporal demand ($M=1.18$, $SD=1.04$), and physical demand ($M=1.18$, $SD=1.04$). For example, W6 further expressed that “[the tool] could be useful as a touch-up tool for when I want to make larger sweeping changes to a very long existing story [...] I can confirm whether or not I want to keep each of these edits” stating that the system “made it a lot easier than any sort of other systems that I’ve used”. The idea of reviewing the changes was important for participants, as W3 explained “You’d still have to review it, but it made enough changes that it [the story] was able to flow well.”

Helping explore and be creative. Participants generated an average CSI score of 71.5 ($SD=21.12$) (Table2). They felt the tool was useful as a creative assistant, even when the text produced was not up to their standards. W6 explained “A lot of times, I am interested in how making a change to a given detail in a story would look [...] I was appreciating [the tool] as sort of an exploratory tool where even if it doesn’t end up producing something that I would personally publish, I think it’s really good to give an example that I can then engage with”. Similarly, W5 mentioned “You can just like, throw in a verb or the general action, interaction between them and it was kind of spitballing ideas”. Regarding the timeline, W7 felt that “it’s an easy way to explore your options of what it sounds like if something happens here or if it sounds better when it’s there.”. They also mentioned how just looking at the visualization could help, for instance: “In terms of ideas, I could see how looking at the map, something could pop up into your head [...] once you have, like, those objects laid out on a map, then that can help you spark inspiration”.

Impact on the writing workflow. Perhaps due to the study setting, all participants heavily used the system during the free-form part. Most participants used the system to explore, “just to see what would happen” (W2), then read the suggested changes and kept them or undid them either by editing an event (W2, W6), undoing using the history graph (W3), deleting actions and characters (W1, W4, W6), or rewriting passages manually (W4, W8). When exploring, participants mostly used the actions and the location views. The timeline was only used to select specific scenes to edit. Participants also had different strategies when creating events, for example, W1, W4, and W7 wrote longer and detailed actions such as “Interrupts conversations, she says that she saw it too, and she is also

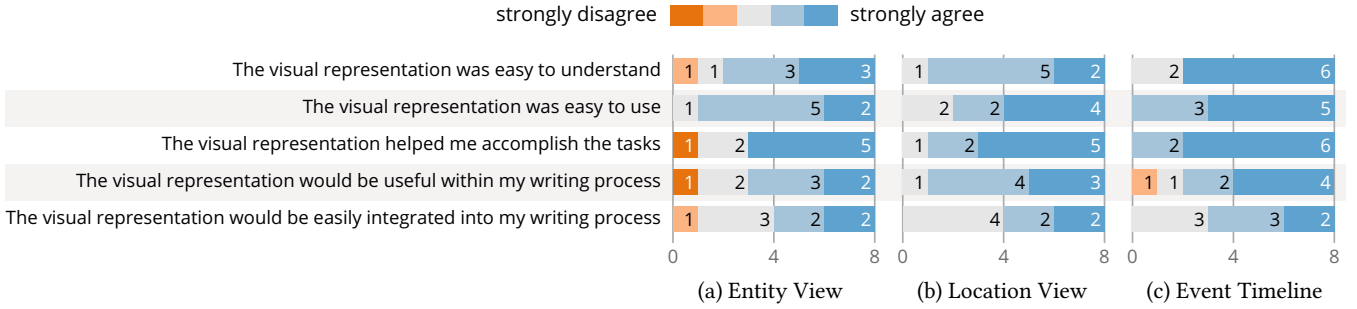


Figure 7: Participants’ response for study 2 when rating the 5-point scale statements for (a) the entities and actions view; (b) the locations and entities view; and (c) the event timeline.

confused” (W1) or explicitly writing a dialogue “Says ‘what are you doing here?’” (W7), whereas others wrote one or two-word actions or descriptions. W8 had the most unique workflow in that they used the visuals to edit the text only once. Instead, they spent most of the session writing their story manually and would periodically refresh the visualization based on what they wrote and then reflect. They explained “*In my most focused state I would just be focused on writing, but then when I was done, I was out of it, I would scan back and forth between text and the [visuals]*”.

Mismatches between the system and participants’ existing workflows. Questions related to integrating the system in the writing process were rated the lowest (but still with an $Mdn=4$). Similar to the findings from the first study, the most critical participants explained that the visuals did not match their workflow and the visualizations they already used. W4 mentioned “*I really really like having a visual representation of the story, but this is not the visual representation I would choose [...] I would like a combination of those two things [timeline and entity view] where those interlinks and connections appear within the timeline.*”⁴. Similarly, W6 explained how a single timeline would not be enough for them: “*my brain interprets timelines in stories where I enjoy having things happening in multiple places at the same time and having a single linear approach makes it very hard to have that sort of interesting kind of development*”⁵. Other than visual representations, W4 also mentioned they “*don’t want the system overwriting [...] because then it’s not me expressing myself through text*” and instead, they would prefer “*just highlighting the places where things need to change, or the things are inconsistent*”.

Participants would have liked more control in using the visual representations to edit the story. While participants could express their intents with the tool, there are nuances in how an idea can be expressed in terms of writing style and voice, which is not currently supported by our system. For example, when discussing adding actions by connecting two entities, W8 mentioned “*I feel like that’s very limiting how you wanted to express what you want versus just typing it out in a sentence because it kind of format it in a specific manner*”. Similarly, W7 explained “*in terms of expressing myself, I would say that it does help [...] but I don’t think it allows*

for like total freedom of expression [...] if you had something, in your mind of how you wanted to say it, it might not come out that way”.

7 Discussion

We propose visual story-writing, an approach that uses visual representations of story elements to support writing and revising narrative texts. Despite how little writing support tools have used visualizations [83], our findings show that they hold great promise to help review and ideate stories. The space of possibilities is large, and our framework of story constructs is an initial attempt at structuring it. In our study, we looked at a subset of this space, using simple representations that did not cover all story elements. However, we encourage others to continue this line of research.

One question with tools like this is how it might change the role of writers. AI-assisted writing tools tend to make us write less and review and edit more [14]. We believe visual story-writing can accommodate both traditional and “editorial” writing roles: in its simplest form, visual story-writing is automatizing the creation of visualizations writers were *already* using [1, 3, 54, 66, 83], with the added bonus of serving as input, if the writer desires it.

Another question is how visual story-writing could be integrated with other writing support tools. We view visual story-writing as complementary to tools that rely on textual prompts (e.g., ChatGPT or Grammarly). Like work on direct manipulation with LLMs has shown [49], we expect visual story-writing to be preferable when writers have a clearer idea of the edit they want to perform (i.e., “move this character there”) because it helps refer to story elements unambiguously. Whereas textual prompts are preferable when the goal is vague (e.g., “make this story sound better”).

Finally, our work targets the creative writing of narratives that involve tangible elements like characters and locations. It applies not only to novels but also to screenwriting for films, TV shows, video games, and theatre plays. However, it is unclear how visual story-writing could work in domains like academic writing, where the elements manipulated are not so tangible.

7.1 Future Work

7.1.1 Extending the framework and exploring new visual interactions. Our framework (§3) operates at a different conceptual level than typical visualization frameworks, such as the grammar of graphics [77]. It combines story elements into constructs that are then linked to visuals without prescribing how to extract story

⁴While this visualization was not implemented in our system, it can be modelled with our framework: start with characters, connect by events, position by locations, and unfold by time

⁵Again, this can be modelled by our framework by unfolding by both time and temporality to visualize nonlinear narratives

elements or what the visualizations should look like, allowing more freedom in implementation. For example, our system represents locations as bubbles instead of a traditional map. Alternatively, images could indicate what an object, character, or location looks like, allowing manipulation (e.g., changing a character's hair colour) or emphasis on certain descriptive points. More work is needed to identify which constructs can best serve writers within visual story-writing workflows, and how to best represent them. For example, inter-character relationships are important in a story [9] and a view constructed by connecting actors would help visualize and edit these relationships. Additionally, the framework can be extended to include story elements beyond those highlighted by Bal [4]. For example, "emotions" could be a basic building block, as used in StoryPrint [76], and "motivations" could be another, as suggested by P11 during the first study. In fact, other narratology models, such as Propp's morphology of the folk tale [61], Todorov's principles of narratives [73], or the contemporary concepts on narrative time by Hume [36] would lead to a different framework.

7.1.2 Visual editing of writing style and story plots. Creative writing encompasses more than just events, actors, locations, and time. The approach could also help edit other aspects, such as writing style. As proposed by Kempen et al. [39], manipulating the syntax tree can restructure sentences. Alternatively, tone, sentence length, or structure could be represented with histograms or text embellishments, allowing style adjustments through visual manipulation. At a higher level, visual story-writing can aid revising the plot. Similar to work on manipulating outlines [23, 82], main plot events can be visually represented to be reordered or changed.

7.1.3 Suggestions and creativity support in the visual space. The mental processes involved in interpreting visuals differ from those involved in processing language [59]. Future work could explore the impact on the creativity of writing stories through visuals rather than with text. Similarly, typical writing support features would be interesting to adapt to the visual space. For instance, there are many benefits of phrase suggestions [7, 15], and future work could investigate if there are similar benefits to suggestions in the visual space. These visual suggestions would essentially try to "auto-complete" an interaction initiated on a visual element.

7.1.4 Supporting long stories. For long stories, the visuals have to find the right balance between overview and details. Our prototype attempts to show the details of all the events and entities and rely on interactions to declutter the views, by merging overlapping actions and allowing users to hover or select a passage to filter what is shown. If details are prioritized, this view could be unfolded by temporality to "spread out" the interactions over time. Unfolding again by time would visualize complex nonlinear narratives. Alternatively, interaction techniques such as details-on-hover or semantic zooming could also help manage longer stories.

7.1.5 Visualization-builder to support custom visuals. The studies showed creative writers have various preferences in terms of visuals. One way to support these workflows would be to help users construct their own visuals. Our procedural framework could inform such a view-builder: the eight story elements would be the

basic building blocks, and users could construct a view by applying the different operators. Once the view is constructed, it could become interactive and automatically populated from the text.

7.2 Limitations

7.2.1 Participants' experience and reservation towards AI might have impacted our results. In study 1, half the participants had no creative writing experience, meaning they were possibly at a disadvantage to accomplish the tasks. Conversely, all participants in study 2 had years of creative experience, and, while they could comment on how visual story-writing could be integrated in their workflows, the results could not tell us about the benefits of the tool for beginners. Similarly, some participants expressed reservations towards AI and reported rarely, if ever, using it. W4 explained "I write because I want to, so if I want to be writing then there's ideas that I have, and I want to express them. I don't want a system to tell me what it is that I am going to be expressing". Considering writers' have strong opinions about AI [8, 30], this might have biased the results.

7.2.2 The system might not have been robust enough to test the full potential of visual story-writing. The system relied on an LLM that occasionally responded unexpectedly, such as modifying unintended elements or refusing to do certain modifications. Additionally, some manipulations were slow to execute (up to 10-15 seconds) despite our efforts to parallelize requests. These issues may have affected our results as interactions should be seamless and paired with immediate effects to promote user exploration [50, 67, 69].

8 Conclusion

We proposed visual story-writing, an approach to support writing by reviewing and manipulating visual representations. In doing so, we defined a framework to help inform the design of story visualizations. By applying this framework, we implemented a prototype system demonstrating one possible design for visual story-writing tools. Two studies covering the different aspects of the writing process showed the potential of creative writing to help keep track of story elements, rapidly specify edits, and explore story variations in a way that encourages creativity. Broadly, our work advocates for a new generation of writing support tools that embed visualization to help review and edit text narratives.

Acknowledgments

We thank Daniel Aureliano Newman for his help revising the questions from study 1. This project was undertaken thanks to funding from IVADO, the Canada First Research Excellence Fund, and NSERC Discovery Grant RGPIN-2018-05072.

References

- [1] Angela Ackerman. 2017. *The Efficient Writer: Using Timelines to Organize Story Details*.
- [2] Kenneth C Arnold, April M Volzer, and Noah G Madrid. 2021. Generative Models Can Help Writers without Writing for Them. In *Joint Proceedings of the ACM IUI 2021 Workshops*. CEUR, College Station, USA, 8.
- [3] Articy Software GmbH & Co 2014. *Articy: narrative design for interactive projects*. Articy Software GmbH & Co. <https://www.articy.com>
- [4] Mieke Bal. 1997. *Narratology: Introduction to the Theory of Narrative* (2nd edition ed.). University of Toronto Press, Scholarly Publishing Division, Toronto.
- [5] Mieke Bal. 2021. *Narratology in Practice*. University of Toronto Press, Toronto.

- [6] Eden Bensaid, Mauro Martino, Benjamin Hoover, and Hendrik Strobelt. 2021. FairyTailor: A Multimodal Generative Framework for Storytelling. doi:10.48550/arXiv.2108.04324 arXiv:2108.04324 [cs]
- [7] Advait Bhat, Saaket Agashe, Parth Oberoi, Niharika Mohile, Ravi Jangir, and Anirudha Joshi. 2023. Interacting with Next-Phrase Suggestions: How Suggestion Systems Aid and Influence the Cognitive Processes of Writing. In *Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23)*. Association for Computing Machinery, New York, NY, USA, 436–452. doi:10.1145/3581641.3584060
- [8] Oloff C. Biermann, Ning F. Ma, and Dongwook Yoon. 2022. From Tool to Companion: Storywriters Want AI Writers to Respect Their Personal Values and Writing Strategies. In *Designing Interactive Systems Conference*. ACM, Virtual Event Australia, 1209–1227. doi:10.1145/3532106.3533506
- [9] Arthur P. Bochner, Carolyn Ellis, and Lisa M. Tillman-Healy. 1997. Relationships as Stories. In *Handbook of Personal Relationships: Theory, Research and Interventions, 2nd Ed.* John Wiley & Sons, Inc., Hoboken, NJ, US, 307–324.
- [10] Richard Brath. 2021. Surveying Wonderland for Many More Literature Visualization Techniques. <https://arxiv.org/abs/2110.08584v1>.
- [11] Virginia Braun and Victoria Clarke. 2022. *Thematic Analysis: A Practical Guide*. SAGE Publications Ltd, Los Angeles.
- [12] Nathalie Bressa, Jordan Louis, Wesley Willett, and Samuel Huron. 2024. Input Visualization: Collecting and Modifying Data with Visual Representations. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (Honolulu, HI, USA) (CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 499, 18 pages. doi:10.1145/3613904.3642808
- [13] Margaret M. Burnett and David W. McIntyre. 1995. Visual Programming. *Computer* 28, 03 (March 1995), 14–16. doi:10.1109/MC.1995.10027
- [14] Daniel Buschek. 2024. Collage Is the New Writing: Exploring the Fragmentation of Text and User Interfaces in AI Tools. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference (DIS '24)*. Association for Computing Machinery, New York, NY, USA, 2719–2737. doi:10.1145/3643834.3660681
- [15] Daniel Buschek, Martin Zürn, and Malin Eiband. 2021. The Impact of Multiple Parallel Phrase Suggestions on Email Input and Composition Behaviour of Native and Non-Native English Writers. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3411764.3445372
- [16] Alex Calderwood, Vivian Qiu, Katy Ilonka Gero, and Lydia B. Chilton. 2020. How Novelists Use Generative Language Models: An Exploratory User Study.. In *IUI '20 Workshops*. CEUR, Cagliari, Italy, 5.
- [17] Erin A. Carroll, Celine Latulipe, Richard Fung, and Michael Terry. 2009. Creativity factor evaluation: towards a standardized survey metric for creativity support. In *Proceedings of the Seventh ACM Conference on Creativity and Cognition* (Berkeley, California, USA) (C&C '09). Association for Computing Machinery, New York, NY, USA, 127–136. doi:10.1145/1640233.1640255
- [18] Erin Cherry and Celine Latulipe. 2014. Quantifying the Creativity Support of Digital Tools through the Creativity Support Index. *ACM Trans. Comput.-Hum. Interact.* 21, 4 (June 2014), 21:1–21:25. doi:10.1145/2617588
- [19] John Joon Young Chung, Wooseok Kim, Kang Min Yoo, Hwaran Lee, Eytan Adar, and Minsuk Chang. 2022. TaleBrush: Sketching Stories with Generative Pretrained Language Models. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–19. doi:10.1145/3491102.3501819
- [20] John Joon Young Chung, Melissa Roemmele, and Max Kreminski. 2025. Toyteller: AI-powered Visual Storytelling Through Toy-Playing with Character Symbols. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–23. doi:10.1145/3706598.3713435
- [21] Nadiyah Bremer-Visual Cinnamon. 2017. All Fights in Dragon Ball Z. <https://dragonballz.visualcinnamon.com>.
- [22] David Comberg, [n. d.], Kurt Vonnegut on Shape of Stories. <https://www.youtube.com/watch?v=oP3c1h8v2ZQ>
- [23] Hai Dang, Karim Benharrak, Florian Lehmann, and Daniel Buschek. 2022. Beyond Text Generation: Supporting Writers with Continuous Automatic Text Summaries. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. ACM, Bend OR USA, 1–13. doi:10.1145/3526113.3545672
- [24] Hai Dang, Sven Goller, Florian Lehmann, and Daniel Buschek. 2023. Choice Over Control: How Users Write with Large Language Models Using Diegetic and Non-Diegetic Prompting. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–17. doi:10.1145/3544548.3580969
- [25] Linda Flower and John R. Hayes. 1981. A Cognitive Process Theory of Writing. *College Composition and Communication* 32, 4 (1981), 365–387. doi:10.2307/356600 jstor:356600
- [26] Gustav Freytag. 2004. *Technique of the Drama: An Exposition of Dramatic Composition and Art*. University Press of the Pacific.
- [27] Gérard Genette, Nitsa Ben-Ari, and Brian McHale. 1990. Fictional narrative, factual narrative. *Poetics today* 11, 4 (1990), 755–774.
- [28] Camille Gobert and Michel Beaudouin-Lafon. 2023. Lorgnette: Creating Malleable Code Projections. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23)*. Association for Computing Machinery, New York, NY, USA, 1–16. doi:10.1145/3586183.3606817
- [29] Brian Hempel, Justin Lubin, and Ravi Chugh. 2019. Sketch-n-Sketch: Output-Directed Programming for SVG. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. Association for Computing Machinery, New York, NY, USA, 281–292. doi:10.1145/3332165.3347925
- [30] Aaron Hertzmann. 2018. Can Computers Create Art? <https://arxiv.org/abs/1801.04486v6>.
- [31] Md Naimul Hoque, Bhavya Ghai, Kari Kraus, and Niklas Elmqvist. 2023. Portrayal: Leveraging NLP and Visualization for Analyzing Fictional Characters. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. ACM, Pittsburgh PA USA, 74–94. doi:10.1145/3563657.3596000
- [32] Ting-Yao Hsu, Yen-Chia Hsu, and Ting-Hao (Kenneth) Huang. 2019. On How Users Edit Computer-Generated Visual Stories. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. doi:10.1145/3290607.3312965
- [33] Ting-Yao Hsu, Chieh-Yang Huang, Yen-Chia Hsu, and Ting-Hao Huang. 2019. Visual Story Post-Editing. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Anna Korhonen, David Traum, and Lluís Màrquez (Eds.). Association for Computational Linguistics, Florence, Italy, 6581–6586. doi:10.18653/v1/P19-1658
- [34] Ting-Hao Kenneth Huang, Francis Ferraro, Nasrin Mostafazadeh, Ishan Misra, Aishwarya Agrawal, Jacob Devlin, Ross Girshick, Xiaodong He, Pushmeet Kohli, Dhruv Batra, C. Lawrence Zitnick, Devi Parikh, Lucy Vanderwende, Michel Galley, and Margaret Mitchell. 2016. Visual Storytelling. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Kevin Knight, Ani Nenkova, and Owen Rambow (Eds.). Association for Computational Linguistics, San Diego, California, 1233–1239. doi:10.18653/v1/N16-1147
- [35] Golina Hulstein, Vanessa Peña-Araya, and Anastasia Bezerianos. 2023. Geo-Storylines: Integrating Maps into Storyline Visualizations. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (Jan. 2023), 994–1004. doi:10.1109/TVCG.2022.3209480
- [36] Kathryn Hume. 2014. *Fantasy and Mimesis: Responses to Reality in Western Literature*. Routledge, Abingdon New York.
- [37] Fantine Huot, Reinald Kim Amplayo, Jennimaria Palomaki, Alice Shoshana Jakobovits, Elizabeth Clark, and Mirella Lapata. 2025. Agents' Room: Narrative Generation through Multi-step Collaboration. doi:10.48550/arXiv.2410.02603 arXiv:2410.02603 [cs]
- [38] Edwin L. Hutchins, James D. Hollan, and Donald A. Norman. 1985. Direct Manipulation Interfaces. *Human-Computer Interaction* 1, 4 (Dec. 1985), 311–338. doi:10.1207/s15327051hci0104_2
- [39] Gerard Kempen, Gert Anbeek, Peter Desain, Leo Konst, and Koenraad De Smedt. 1987. Author Environments: Fifth Generation Text Processors. In *ESPRIT'86: Results and Achievements*. Elsevier Science Publishers, Brussels, Belgium, 8.
- [40] Nam Wook Kim, Benjamin Bach, Hyejin Im, Sasha Schriber, Markus Gross, and Hanspeter Pfister. 2018. Visualizing Nonlinear Narratives with Story Curves. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (Jan. 2018), 595–604. doi:10.1109/TVCG.2017.2744118
- [41] Amy J. Ko and Brad A. Myers. 2006. Barista: An Implementation Framework for Enabling New Tools, Interaction Techniques and Views in Code Editors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. Association for Computing Machinery, New York, NY, USA, 387–396. doi:10.1145/1124772.1124831
- [42] Mina Lee, Katy Ilonka Gero, John Joon Young Chung, Simon Buckingham Shum, Vipul Raheja, Hua Shen, Subhashini Venugopalan, Thiemo Wambsgans, David Zhou, Emad A. Alghamdi, Tal August, Avinash Bhat, Madiha Zahrah Choksi, Senjuti Dutta, Jin L. C. Guo, Md Naimul Hoque, Yewon Kim, Simon Knight, Seyed Parsa Neshaei, Agnia Sergeyuk, Antonette Shibani, Disha Shrivastava, Lila Shroff, Jessi Stark, Sarah Sterman, Sitong Wang, Antoine Bosselut, Daniel Buschek, Joseph Chee Chang, Sherol Chen, Max Kreminski, Joonsuk Park, Roy Pea, Eugenia H. Rho, Shannon Zejiang Shen, and Pao Siangliulue. 2024. A Design Space for Intelligent and Interactive Writing Assistants. doi:10.1145/3613904.3642697 arXiv:2403.14117 [cs]
- [43] Genevieve Liveley. 2021. *Narratology*. Oxford University Press, Oxford.
- [44] David Madden. 1988. *Revising Fiction: A Handbook for Writers*. Plume, New York.
- [45] Suraj Maharjan, Sudipta Kar, Manuel Montes, Fabio A. González, and Thamar Solorio. 2018. Letting Emotions Flow: Success Prediction by Modeling the Flow of Emotions in Books. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, Marilyn Walker, Heng Ji, and Amanda Stent (Eds.). Association for Computational Linguistics, New Orleans, Louisiana, 259–265. doi:10.18653/v1/N18-2042
- [46] Marcel Marti, Jodok Vieli, Wojciech Witoń, Rushit Sanghrajka, Daniel Inversini, Diana Wotruba, Isabel Simo, Sasha Schriber, Mubbasir Kapadia, and Markus Gross. 2018. CARDINAL: Computer Assisted Authoring of Movie Scripts. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces (IUI '18)*. Association for Computing Machinery, New York, NY, USA, 509–519.

- doi:10.1145/3172944.3172972
- [47] Damien Masson, Young-Ho Kim, and Fanny Chevalier. 2025. Textshop: Interactions Inspired by Drawing Software to Facilitate Text Editing. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '25). Association for Computing Machinery, New York, NY, USA, Article 775, 14 pages. doi:10.1145/3706598.3713862
- [48] Damien Masson, Sylvain Malacria, G ry Casiez, and Daniel Vogel. 2023. Statslator: Interactive Translation of NHST and Estimation Statistics Reporting Styles in Scientific Documents. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23)*. Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3586183.3606762
- [49] Damien Masson, Sylvain Malacria, G ry Casiez, and Daniel Vogel. 2024. Direct-GPT: A Direct Manipulation Interface to Interact with Large Language Models. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–16. doi:10.1145/3613904.3642462
- [50] Damien Masson, Jo Vermeulen, George Fitzmaurice, and Justin Matejka. 2022. Supercharging Trial-and-Error for Learning Complex Software Applications. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3491102.3501895
- [51] Mika l Mayer, Viktor Kuncak, and Ravi Chugh. 2018. Bidirectional Evaluation with Direct Manipulation. *Implementation for: Bidirectional Evaluation with Direct Manipulation 2*, OOPSLA (Oct. 2018), 127:1–127:28. doi:10.1145/3276497
- [52] Randall Munroe. 2009. Movie Narrative Charts. <https://xkcd.com/657/>.
- [53] B. A. Myers. 1986. Visual Programming, Programming by Example, and Program Visualization: A Taxonomy. *SIGCHI Bull.* 17, 4 (April 1986), 59–66. doi:10.1145/22339.22349
- [54] C. M. Neuwirth and D. S. Kaufer. 1989. The Role of External Representation in the Writing Process: Implications for the Design of Hypertext-Based Writing Tools. In *Proceedings of the Second Annual ACM Conference on Hypertext - HYPERTEXT '89*. ACM Press, Pittsburgh, Pennsylvania, United States, 319–341. doi:10.1145/74224.74250
- [55] nextui-org. 2021. *NextUI*. nextui-org. <https://nextui.org/>
- [56] Chris North and Ben Shneiderman. 2001. A Taxonomy of Multiple Window Coordinations. (2001), 8.
- [57] Scott Norton. 2011. *Developmental Editing: A Handbook for Freelancers, Authors, and Publishers* (illustrated edition ed.). University of Chicago Press, Chicago London.
- [58] OpenAI. 2023. OpenAI Node API Library. OpenAI.
- [59] Allan Paivio. 1990. *Mental Representations: A Dual Coding Approach*. Oxford University Press, Oxford.
- [60] Tom Pope. 1998. *Good Scripts, Bad Scripts: Learning the Craft of Screenwriting Through 25 of the Best and Worst Films in Hi Story*. Three Rivers Press, New York.
- [61] V. Propp and Svatava Pirkova-Jakobson. 2010. *Morphology of the Folk Tale*. University of Texas Press.
- [62] Lu Qiang and Chai Bingjie. 2016. StoryCake: A Hierarchical Plot Visualization Method for Storytelling in Polar Coordinates. In *2016 International Conference on Cyberworlds (CW)*. IEEE, Chongqing, China, 211–218. doi:10.1109/CW.2016.43
- [63] Meta. 2013. *React – A JavaScript Library for Building User Interfaces*. Meta. <https://reactjs.org/>
- [64] Andrew J. Reagan, Lewis Mitchell, Dilan Kiley, Christopher M. Danforth, and Peter Sheridan Dodds. 2016. The Emotional Arcs of Stories Are Dominated by Six Basic Shapes. *EPJ Data Science* 5, 1 (Dec. 2016), 1–12. doi:10.1140/epjds/s13688-016-0093-1
- [65] Mohi Reza, Nathan Laundry, Ilya Musabirov, Peter Dushniku, Zhi Yuan "Michael" Yu, Kashish Mittal, Tovi Grossman, Michael Liut, Anastasia Kuzminykh, and Joseph Jay Williams. 2024. ABScribe: Rapid Exploration & Organization of Multiple Writing Variations in Human-AI Co-Writing Tasks Using Large Language Models. doi:10.1145/3613904.3641899 arXiv:2310.00117 [cs]
- [66] Blake Snyder Enterprises. 2014. *Save The Cat! Story Structure Software*. Blake Snyder Enterprises. <https://savethecat.com/save-the-cat-software>
- [67] Shneiderman. 1983. Direct Manipulation: A Step Beyond Programming Languages. *Computer* 16, 8 (Aug. 1983), 57–69. doi:10.1109/MC.1983.1654471
- [68] Ben Shneiderman. 1982. The Future of Interactive Systems and the Emergence of Direct Manipulation†. *Behaviour & Information Technology* 1, 3 (July 1982), 237–256. doi:10.1080/01449298208914450
- [69] Ben Shneiderman, Gerhard Fischer, Mary Czerwinski, Mitch Resnick, Brad Myers, Linda Candy, Ernest Edmonds, Mike Eisenberg, Elisa Giaccardi, Tom Hewett, Pamela Jennings, Bill Kules, Kumiyo Nakakoji, Jay Nunamaker, Randy Pausch, Ted Selker, Elisabeth Sylvan, and Michael Terry. 2006. Creativity Support Tools: Report From a U.S. National Science Foundation Sponsored Workshop. *International Journal of Human-Computer Interaction* 20, 2 (May 2006), 61–77. doi:10.1207/s15327590ijhc2002_1
- [70] Slate.js. 2016. *Slate.js*. Slate.js. <https://www.slatejs.org>
- [71] Sangho Suh, Meng Chen, Bryan Min, Toby Jia-Jun Li, and Haijun Xia. 2024. Luminate: Structured Generation and Exploration of Design Space with Large Language Models for Human-AI Co-Creation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '24). Association for Computing Machinery, New York, NY, USA, Article 644, 26 pages. doi:10.1145/3613904.3642400
- [72] Y. Tanahashi and Kwan-Liu Ma. 2012. Design Considerations for Optimizing Storyline Visualizations. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (Dec. 2012), 2679–2688. doi:10.1109/TVCG.2012.212
- [73] Tzvetan Todorov. 1971. The 2 Principles of Narrative. *Diacritics* 1, 1 (1971), 37–44. doi:10.2307/464558 jstor:464558
- [74] Jing Wang, Jianlong Fu, Jinhui Tang, Zechao Li, and Tao Mei. 2018. Show, Reward and Tell: Automatic Generation of Narrative Paragraph from Photo Stream by Adversarial Training. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence (AAAI'18/IAAI'18/EAAI'18)*. AAAI Press, New Orleans, Louisiana, USA, 7396–7403.
- [75] Michelle Q. Wang Baldonado, Allison Woodruff, and Allan Kuchinsky. 2000. Guidelines for using multiple views in information visualization. In *Proceedings of the working conference on Advanced visual interfaces*. ACM, Palermo Italy, 110–119. doi:10.1145/345513.345271
- [76] Katie Watson, Samuel S. Sohn, Sasha Schriber, Markus Gross, Carlos Manuel Muniz, and Mubbasir Kapadia. 2019. StoryPrint: An Interactive Visualization of Stories. In *Proceedings of the 24th International Conference on Intelligent User Interfaces (IUI '19)*. Association for Computing Machinery, New York, NY, USA, 303–311. doi:10.1145/3301275.3302302
- [77] Leland Wilkinson, D. Wills, D. Rope, A. Norton, and R. Dubbs. 2005. *The Grammar of Graphics* (2nd edition ed.). Springer, New York.
- [78] Eric R. Williams. 2018. How to View and Appreciate Great Movies.
- [79] Zihan Yan, Chunxu Yang, Qihao Liang, and Xiang 'Anthony' Chen. 2023. XCreation: A Graph-based Crossmodal Generative Creativity Support Tool. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. ACM, San Francisco CA USA, 1–15. doi:10.1145/3586183.3606826
- [80] Ann Yuan, Andy Coenen, Emily Reif, and Daphne Ippolito. 2022. Wordcraft: Story Writing With Large Language Models. In *Proceedings of the 27th International Conference on Intelligent User Interfaces (IUI '22)*. Association for Computing Machinery, New York, NY, USA, 841–852. doi:10.1145/3490099.3511105
- [81] J.D. Zamfirescu-Pereira, Richmond Y. Wong, Bjoern Hartmann, and Qian Yang. 2023. Why Johnny Can't Prompt: How Non-AI Experts Try (and Fail) to Design LLM Prompts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–21. doi:10.1145/3544548.3581388
- [82] Zheng Zhang, Jie Gao, Ranjodh Singh Dhaliwal, and Toby Jia-Jun Li. 2023. VISAR: A Human-AI Argumentative Writing Assistant with Visual Programming and Rapid Draft Prototyping. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. ACM, San Francisco CA USA, 1–30. doi:10.1145/3586183.3606800
- [83] Zixin Zhao, Damien Masson, Young-Ho Kim, Gerald Penn, and Fanny Chevalier. 2025. Making the Write Connections: Linking Writing Support Tools with Writer's Needs. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Yokohama, Kanagawa, Japan) (CHI '25). Association for Computing Machinery, New York, NY, USA, Article 78, 21 pages. doi:10.1145/3706598.3713161
- [84] Xiaojin Zhu. 2013. Persistent Homology: An Introduction and a New Text Representation for Natural Language Processing. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*. IJCAI, Beijing, China, 1953–1959.
- [85] Yaqian Zhu and John Kolassa. 2018. Assessing and Comparing the Accuracy of Various Bootstrap Methods. *Communications in Statistics - Simulation and Computation* 47, 8 (Sept. 2018), 2436–2453. doi:10.1080/03610918.2017.1348516

A Appendix

A.1 Extracting Information from the Text

The visual representations are generated from information extracted from the text about the entities, locations, and events. The extraction follows three steps, in sequence: 1) extract the entities and their traits; 2) extract the locations; 3) extract the events for each sentence. For all steps, OpenAI's "GPT4-o" model was used.

A.1.1 Extracting Entities. The prompt below extracts all the entities (i.e., characters and inanimate objects) from a story.

<STORY TEXT>

Extract all the entities in this story. For each entity, extract its ‘name’, an emoji best visually describing the entity (e.g., use the emoji of a person if it is a person but avoid reusing the same emojis), and properties about the entity, if any (no more than 3). Properties have to be adjectives describing the entity and their value should represent the intensity of the property (on a scale from 1 to 10).

The model responds with a structured output using the following JSON schema:

```
entities: [{
  name: string,
  emoji: string
  properties: [{
    name: string,
    value: number
  }]
}]
```

A.1.2 Extracting Locations. The prompt below extracts all the locations from a story.

<STORY TEXT>

Extract all the main locations visited by the characters in this story. For each location, extract its ‘name’ and an emoji best visually representing the location

The output is structured, using the following JSON schema:

```
locations: [{
  name: string,
  emoji: string
}]
```

A.1.3 Extracting Events. Extracting the events is done after the entities and locations are extracted. First, the story is divided into sentences. Then, in parallel and for each sentence, the following prompt is used.

BEFORE: <TEXT BEFORE>
TEXT: <SENTENCE>

Extract the actions done by the characters in TEXT and only the actions in TEXT. Do not extract the actions from BEFORE. Only consider actions that are happening exactly at the moment of TEXT, ignore memories etc. If there are no actions fulfilling these criteria in TEXT, then return an empty array. Source and target should be characters from this list: <ENTITIES>. Here are some possible locations but there might be others: <LOCATIONS> If an action is done by a character to itself, then the source and target character should be the same. For each action, extract the ‘name’ of the action (no more

than 2 words), the source character (the character doing the action) and the target character (the character targeted by the action), and the location of the action (you can use ‘unknown’ if the location cannot be inferred from the text).

With <TEXT BEFORE> corresponding to the text before the sentence (to give the model some context), <SENTENCE> the text of the sentence, <ENTITIES> the list of entities extracted previously, and <LOCATIONS> the list of locations extracted previously.

The output is structured, using the following JSON schema:

```
actions: [{
  name: string,
  source: string,
  target: string,
  location: string
}]
```

All extracted events are associated with their respective sentences to support highlighting the text when manipulating the visual representations. Additionally, to make future extraction faster, only sentences that changed are re-extracted.

A.2 Editing the Story by Manipulating the Visuals

Below, we list the prompts used to suggest edits to the story when visual representations are modified.

A.2.1 Reorder Events in the Timeline. This prompt works by giving the list of events in the current and new order. Once the story is modified, the newly generated text is re-extracted fully to make sure the visual representations match the new text.

<STORY TEXT>

In this story, the current order of actions is:

<CURRENT ORDER>

Rewrite the story to EXACTLY follow this new order:

<NEW ORDER>

<STORY> is the full text of the story, <CURRENT ORDER> is a list of the events in the story, as extracted previously, and <NEW ORDER> is this same list but with the events reorganized. Events are represented in the form “<SOURCE> <EVENT NAME> <TARGET>”.

A.2.2 Adding, Changing or Removing An Action. All the prompts resulting from manipulating links between nodes on the event view follow the same format:

<STORY TEXT>

SOURCE: <SOURCE ENTITY>

TARGET: <TARGET ENTITY>

ACTION: <ACTION NAME>

<PROMPT>

<PROMPT> is either “Rewrite <blank> to add that SOURCE also <ACTION> TARGET” when adding an action, “Rewrite the story so that SOURCE also <ACTION> TARGET” when changing an action,

and “Rewrite the story so that SOURCE does not do ACTION to TARGET” when removing an action.

A.2.3 *Removing an Entity.*

<STORY TEXT>

Rewrite the story so that there is no <ENTITY NAME>

Note that no prompts are executed when adding an entity. This is only when adding an action that the text is modified.

A.2.4 *Moving an Entity.*

<STORY TEXT>

Rewrite the story so that <ENTITY NAME> never goes

to the <CURRENT LOCATION> but instead goes to the <NEW LOCATION>

A.2.5 Targeted Edits. Visual edits can modify only a specific passage or set of scenes, if those scenes are selected on the timeline. For this to work, the prompts mentioned above are slightly modified so that <STORY TEXT> includes the fully story text with the selected passage replaced by “TEXT_TO_REWRITE”. Then, a new line is added defining “TEXT_TO_REWRITE” as the selected text passage. The prompt is then modified to indicate that it should only rewrite the passage. The result is then integrated into the full story. This approach is inspired by the prompts used in DirectGPT [49].