DramatVis Personae: Visual Text Analytics for Guiding Character Development and Avoiding Stereotypes in Creative Storytelling

MD NAIMUL HOQUE, University of Maryland, College Park, USA

BHAVYA GHAI, Stony Brook University, USA

NIKLAS ELMQVIST, University of Maryland, College Park, USA

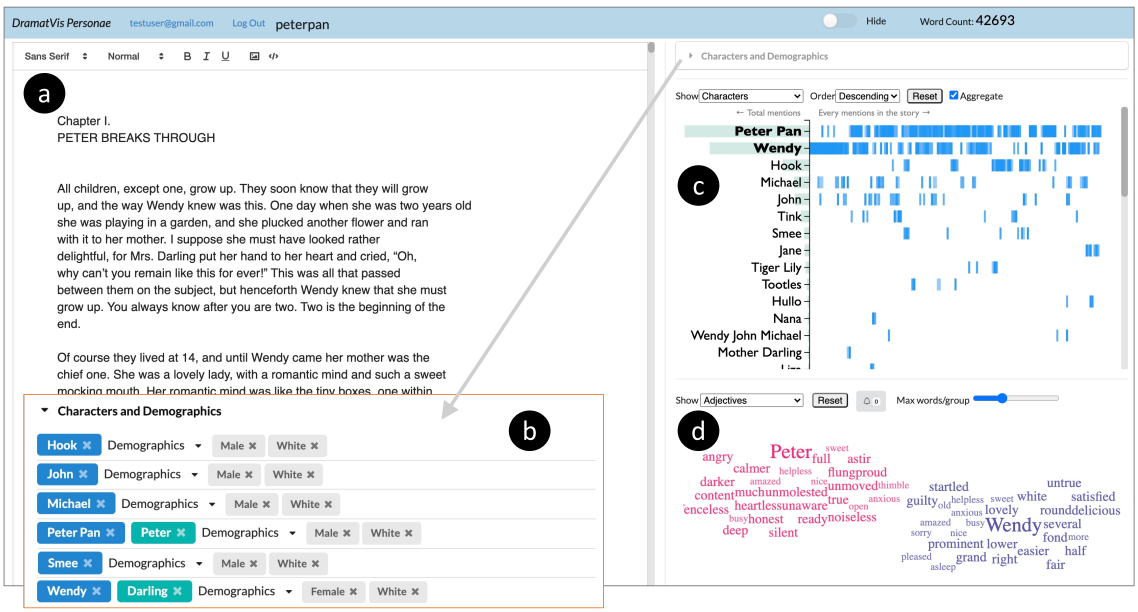

Fig. 1. **Overview of DramatVis Personae.** (a) Rich text editor. (b) Characters and Demographics panel for tracking characters, merging aliases, and assigning demographics and social identities to characters. (c) Timeline representation of the story (*Peter Pan* by J. M. Barrie (1911)) showing every mentions of a character as well as the total number of mentions of a character. (d) Word zone [31] showing sample adjectives used for the selected characters (Peter Pan and Wendy).

Implicit biases and stereotypes are often pervasive in different forms of creative storytelling such as novels, screenplays, and children's books. To understand the kind of biases writers are concerned about and how they mitigate those in their writing, we conducted formative interviews with nine writers. The interviews suggested that despite a writer's best interest, tracking and managing implicit biases such as a lack of agency, supporting or submissive roles, or harmful language for characters representing marginalized groups is challenging as the story becomes longer and complicated. Based on the interviews, we developed DramatVis Personae, a visual analytics tool that allows writers to assign social identities to characters, and evaluate how characters and different intersectional social identities are represented in the story. We conducted follow up study and case studies with three writers and found that DramatVis Personae is easy-to-use, naturally integrated into the writing process, and helps writers in several critical bias identification and mitigation tasks.

CCS Concepts: • **Human-centered computing → Visualization**; **Empirical studies in visualization**; **Visualization design and evaluation methods**.

Additional Key Words and Phrases: Creativity, creative writing, visualization, natural language processing, fictional characters.

ACM Reference Format:
Md Naimul Hoque, Bhavya Ghai, and Niklas Elmqvist. 2022. *DramatVis Personae*: Visual Text Analytics for Guiding Character Development and Avoiding Stereotypes in Creative Storytelling. In *CHI '22: Proceedings of the ACM Conference on Human Factors in Computing Systems, April 30–May 6, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 24 pages. https://doi.org/xxxx

1 INTRODUCTION

Gandalf. Elizabeth Bennet. Hermione Granger and Ron Weasley. Atticus Finch. Anomander Rake, Lord of Moon's Spawn and Son of Darkness. Holly Golightly, Lisbeth Salander, and Hannibal Lecter: literature lives and dies by its characters. Heroes and anti-heroes, villains and bad guys, innocent bystanders or willing accomplices—fiction is arguably about conjuring more or less complete humans out of whole cloth and then providing audiences with the emotional release of *catharsis*, often by having these characters go through hell and high water. Therein lies also the secret of great literature: creating believable, nuanced, and multidimensional characters that spring out of the written page and come to life in the reader's mind. Achieving this is no mean feat, particularly when considering that truly great fiction often requires a diverse, inclusive, and just treatment of its cast of characters; its *dramatis personae*. For years untold, this creative process has largely been conducted in the mind and the scratchbooks and the libraries and the writing nooks of the authors themselves. But is there a better way?

Significantly, while modern literature is becoming increasingly diverse and inclusive, there is still considerable room for addressing stereotypes and bias. Implicit bias is just that, implicit and unconscious and difficult to wheedle out even for the best and most self-reflective of authors. As a result, literature is filled with tired, unoriginal, and sometimes harmful stereotypes related to race, gender, sexuality, ethnicity, and age, such as the trope of the angry African-American woman, the studious Asian person, or the helpless damsel in distress [12, 47, 47]. For example, prolific horror author Stephen King is often accused of tone-deaf and problematic language, particularly in his older work, and Mark Twain's often lauded stories about Tom Sawyer and Huckleberry Finn are notorious for their use of racial slurs. Many of these stereotypes and slurs can now be recognized automatically using modern natural language processing techniques to guide authors to avoid them—or to embrace them, if their art so dictates.

We present DramatVis Personae, a web-based visual analytics system to support character development for creative writing. DramatVis Personae—DVP for short—is designed to integrate smoothly with the writer's own creative process, allowing them to upload new content as it becomes available or even write in the tool itself, and then having its text analytics and visualizations update in real time. Using entity identification, dependency parsing, and other sophisticated text analytics methods, DVP will automatically detect characters in the text and collect data about them as the story progresses, including their aliases, demographic information, and representation throughout the story. The author can also furnish additional information for each character, such as their age, ethnicity, gender, etc. The DVP dashboard uses this continually growing dataset to visualize the plot of the story as well as the character development over time. In particular, DramatVis Personae will help authors identify harmful stereotypes in their writing.

Creative writing is a notoriously personal and individual experience, with each author's idiosyncratic process being their uniquely own. We make no claim that DVP is a one-size solution that will fit all authors, stories, and settings. However, we designed DVP based on in-depth formative interview with nine creative writers with published stories in their portfolios. These interviews guided the design and development of the DramatVis Personae tool. After our initial design and implementation, we then approached these original writers and conducted follow-up expert reviews using the tool. During a hands-on evaluation session conducted over videoconferencing, one writer was asked to use the tool to write a short story given a specific writing prompt. Other participants used the tool to evaluate their own existing stories, or stories written by others. We observed their performance and then interviewed them with regards to their experience. All participants expressed positive sentiment about the DVP tool, claiming that it helped them get a better grip of their characters and their story arcs throughout the process. In particular, all participants appreciated that the tool managed and visualized character demographics, enabling them to write a more nuanced and equitable story.

In sum, we claim the following contributions with this work: (1) findings on how to support the creative writing process, in particular in mitigating implicit bias, from an interview study involving nine fiction writers; (2) a visual analytics tool called DramatVis Personae for supporting both online creative writing as well as offline analysis of fiction based on the interview findings; and (3) results from two separate case studies of deploying DramatVis Personae with creative writers in both story generation as well as story analysis settings.

2 BACKGROUND

Our work in this paper intersects several areas of research, including creative storytelling, bias in storytelling, natural language processing (NLP), and text and literary visualization. Below we describe each of these topics in turn.

2.1 Creative Storytelling and Writing

We use the umbrella "creative storytelling" in this paper to encompass the production of all forms of fictional narratives, including oral stories, written novels, short stories, and plays, and scripts for live or screen movies, TV, or radio shows. Furthermore, we narrow our work here to focus on "creative writing," or the production of the written artifacts capturing the narrative, such as the book manuscript, screenplay, or short story.

Speaking informally, there are two schools of thought for creative writing: *pantsing* and *plotting* [37]. *Pantsing*, or seat-of-your-pants writing, entails the writer establishing the world, characters, and premise of the story, but mostly avoiding to plan out the plot prior to actually writing it. Instead, the writer essentially sets the world and its characters "in motion" and then let their personality and interactions determine the outcome. Famous American horror writer Stephen King is a noted pantser [43]; he typically finishes a book in approximately three months ("a season"), forcing himself to write every day in order to keep the characters and the story fresh in his mind. This often results in the organic, unpredictable, and realistic stories that are one of King's hallmarks.[1] Other notable pantsers include Margaret Atwood, Pierce Brown, and Chuck Wendig. Atwood, for example, is said to start her writing with *"an image, scene, or voice... I couldn't write the other way round with structure first. It would be too much like paint-by-numbers."*

Plotting, on the other hand, takes a more structured approach to creative writing by having the writer carefully plan out the plot of the book well in advance prior to actually writing it. Some plotting methods, such as the Snowflake

[1]On the other hand, King is often accused of bad endings to his books, which is a common problem for pantsing because of the limited ability to plan ahead.

method [37], stipulates iteratively working on an increasingly detailed outline until the outline eventually becomes the manuscript. For this reason, plotters tend to spend significant time and resources *prior* to writing on research (for realistic settings) and world building (for fantasy or science fiction settings). Notable plotters include J.K. Rowling, E.L. Stein, Joseph Heller, and John Grisham.

A logical outcome of this clear dichotomy is that a tool for supporting creative writing must support both modes of working: either when writing by the seat of your pants, or when fastidiously planning the story in advance.

2.2 Creative Writing Processes

The individual writing processes varies between each author, and idiosyncrasies are common. Well-known fantasy writer G.R.R. Martin, for example, famously writes his first drafts in WordStar 4.0 on a computer disconnected from the internet and running only Microsoft DOS. The computer, he explains, can do one very important thing: "it can type words." Other authors have similar peculiar processes: Mark Twain wrote in bed; Ernest Hemingway wrote standing up; Victor Hugo wrote naked; Vladimir Nabokov used index cards; Honoré de Balzac reportedly binged on coffee; and Aaron Sorkin acted out dialogue in front of a mirror (once breaking his nose when he inadvertently headbutted it during a particularly emotional passage). In sum, this means that there are few common denominators for creative writing.

Beyond the every-day-use text processors such as Microsoft Word, there are several professional and open-source software available to writers for helping them in guiding character development. Scrivener [5] is a paid service for organizing stories. It allows flexible page breaking, adding synopsis and notes to each section, and easy merging or swapping between sections. It also has a distraction-free writing mode where everything else on the computer is tuned out, similar to G.R.R. Martin's MS-DOS binges. Granthika [2] is a similar sort of paid service that helps writers in tracking characters and events in a story. It lets users to integrate knowledge to the system as they write, and then use those knowledge for tracking in a timeline. It also allows writers to apply causal constraint to the events and people in a story (e.g, "The inquest must happen after the murder"). While there are certain similarities between our tool and these paid services (e.g, interactively adding characters for tracking), none of these tools have features to add social identities to characters, and investigating potential biases against marginalized groups. Additionally, these tools primarily use textual description for summarizing and communication. While that is helpful, we utilize data visualization, a visual communication medium, for quick and fast digression of information extracted by the tool.

2.3 Bias in Storytelling

Creative writing is often considered a reflection of society [70]. Stereotypes in the form of art often mirror the problems, issues, thinking, and perception of different social groups in society [70, 77]. The presence of biases and stereotypes, especially gender and racial bias, has been reported ubiquitously in different forms of creative writing. We provide a brief overview of research in this area below.

Many researchers have shown the prevalence of gender stereotypes in children's books, dating back to early 1970s [22]. Since then, several studies have reported that males are often portrayed as active and dominating, while females are instead described as passive and soft [44, 56, 63]. Other studies have found the presence of racial bias [47], stereotypes against disability [12], and occupation [30] in children's books. Researchers have argued that the presence of such stereotypes in children's book is severely problematic as children are susceptible to inherit stereotypes at an early age [44, 77]. While the situation is improving (i.e., females are portrayed with more active roles in recent children's books) due to increased social awareness, the improvement is not significant [65], and there are still reports of prevalence of different stereotypes in children's books [7, 23, 30].

Another form of creative storytelling medium that has been heavily criticized for promoting stereotypes is movie scripts. The *Geena Davis Institute* regularly publishes reports of gender and racial representation in Hollywood movies and is a valuable resource for current representational problems. The institute has found under representation and misrepresentation of females [58], and Black or African American females in Hollywood [60]. It is worth noting here that there are reports of better representation and inclusivity in recent years [36, 59, 61], indicating an increased awareness among content creators. Beyond Hollywood, researchers have found similar sort of biases in television shows and movies in other countries. Emons et al. [19] found stereotypes in gender roles of males and females in U.S.-produced Dutch TV shows, misrepresenting females in Dutch society. Madaan et al. [50] has shown existence of gender biases in Bollywood movie scripts.

Finally, newer forms of writing such as blogs, online writeups, and social media posts are rife with harmful stereotypes. Fast et al. [20] analyzed fictions written by novice writers in the online community Wattpad and found it to be rampant with common gender stereotypes. Joseph et al. [38] analyzed forty-five thousand Twitter users who actively tweeted about the Michael Brown and Eric Garner tragedies. Their method can quantify semantic relations between social identities. Other works discussed the impact of stereotypes in Reddit [21], Facebook [51], and U.S. history books [49].

All the above-mentioned research has been instrumental in raising awareness among writers, directors, and general audience, a critical step towards equality. Motivated by this line of work, we focus on supporting the unwanted bias mitigation strategies for self-aware writers.

2.4 NLP for Story Analysis

Over the years, NLP research has developed and refined different techniques such as coreference resolution, named entity recognition (NER), dependency parsing, sentiment analysis, etc., which can be instrumental in analyzing stories. For example, NER helps extract named entities such as person names, organizations, locations, etc. from text. In our context, this can help identify different characters of a story [18]. Dependency parsing deals with finding relationships between words. This can help identify different adjectives/verbs linked to a character in a story [55]. Similarly, coreference resolution can help track the representation of different characters and their demographic groups across the story line [45]. Other studies in the NLP literature have specifically focused on analyzing stories. This includes segmenting stories by predicting chapter boundaries [66], recognizing flow of time in a story [41], analyzing emotional arc of a story [67], extracting character networks from novels [46], etc. Using such methods, researchers have conducted large-scale analysis of book, and stories [20, 34, 38, 46, 57, 64].

While NLP techniques and models are predominantly used for corpora analysis, we leverage different NLP techniques to iteratively analyze text in our interactive tool. It is also worth noting here that the existing writing software discussed in Section 2.2 also uses some NLP techniques. For example, Grammarly supports users in finding grammertical errors and finding appropriate deliveration of speech. However, none of them have features (using NLP) to support bias identification from text.

2.5 Visualization for Text and Literature

Harking back to some of the original approaches to visualizing "non-visual" text documents [80], data visualization has long been proposed as an alternative to reading through large document corpora [8, 25, 72]. For example, the investigative analytics tool Jigsaw [73] is often styled as a "visual index" into a document collection; while it is not a replacement for reading, it provides a linked collection of entities and documents for easy overview and navigation.

This generally also true for document and text visualization as a whole; the goal is to be able to "see beyond" the raw text into content, structure, and semantics [8].

Beyond simplistic text visualization techniques such as word clouds, data visualization can become particularly powerful when applied to entire documents [25]. These ideas can also be used for literary analysis of fiction and poetry [16]. For example, Rohrer et al. [68] used implicit 3D surfaces to show similarities between documents, such as the work of William Shakespeare. Similarly, Keim and Oelke propose a visual fingerprinting method for performing comparative literature analysis [40]. McCurdy [52] present a organic linked visualization approach to scaffolding close reading of poetry. The literary tool Myopia [15] uses color-coded entities to show the literary attributes of a poem for readers. Abdul-Rahman et al. [6] also apply data visualization to poetry. Our work here is inspired by, if not the design and implementation, then at least the motivation of these tools; however, in comparison, our goal is to support the creative processing while focusing on identifying and mitigating implicit bias.

Finally, visualization can also be applied to the stories themselves rather than the actual text. XKCD #657,[2] titled *Movie Narrative Charts*, shows temporal representations of plotlines in five movies, including the original *Star Wars* trilogy (1977–1983), *Jurassic Park* (1993), and the complete *Lord of the Rings* movie trilogy (2001–2003). Liu et al. [48] propose an automated approach to generating such storyline visualizations called StoryFlow. Tanahashi and Ma discuss design considerations for best utilizing storylines [74]. Some effort has also been directed towards minimizing crossings in storyline visualizations [29]. VizStory [35] takes a very literal approach to visualizing stories by identifying segments and themes and then searching the web for appropriate representative images. TextFlow [17] and ThemeDelta [24] and related topic modeling visualizations can be used to automatically extract and visualize evolving themes in a document (or document collection) over time. StoryPrint [79] shows polar representations of movie scripts and screenplays in fashion similar to Keim and Oelke's fingerprints; the authors note that this approach could also be used to support the creative writing process. Finally, StoryCurves [42] shows non-linear narratives in movies on a timeline representation. Common for all of these storyline and plot visualization tools is that they are designed mostly for retrospective analysis and not for on-line creative writing. Nevertheless, we draw on all of these tools in our design of the DramatVis Personae character arc visualization tool.

Finally, Story Analyzer [55] shows several visualizations representing summary statistics of a story. Similar to the above-mentioned tools, Story Analyzer [55] does not have any feature for investigating presence of social identities in the story. Additionally, the tool is not a writing tool, rather primarily a post processing tool for analyzing already written stories. It is not clear how users could benefit from the tool, and how it can be used for larger documents.

3 INTERVIEWS WITH CREATIVE WRITERS

To inform the design of our creative writing tool, we conducted semi-structured interviews with 9 creative writers. Through this formative study, we aimed at understanding the writing process for creative writers, their use of technology, different modalities of stereotypes they are concerned about, and how they address stereotypes in their stories.

3.1 Participants

We initiated the recruitment process by posting advertisement to social media such as Twitter and Facebook, local mailing lists, and university mailing lists. After initial pre-screening, we selected 12 participants out of 54 responses. Among the 12 participants, 9 participants responded for the final interviews.

[2]https://xkcd.com/657/

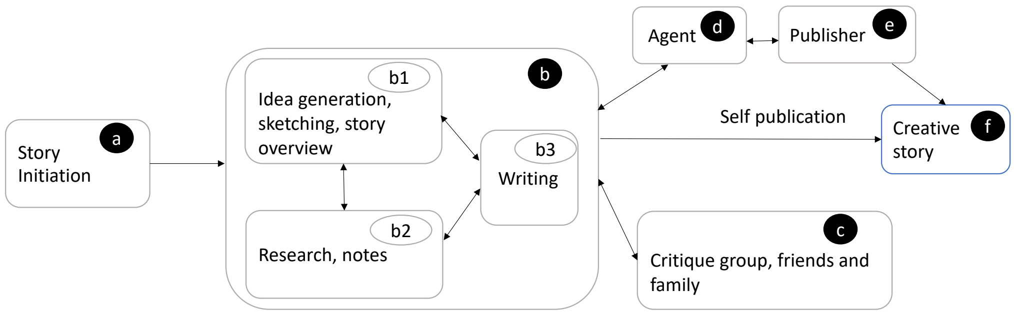

Fig. 2. Writing Cycle of Creative Storytelling as perceived from the formative interviews.

Our inclusion criteria included prior experiences in creative writing such as novels, short stories, screenplays, etc, and familiarity with writing in text editors. The participants varied in Gender ($Female = 5, Male = 3, Non\text{-}binary = 1$), Race ($White = 4, Asian = 4, Black$ or $African\text{-}american = 1$), and Age (19-44, $M = 26.5, SD = 5.6$). Five participants identified themselves as professional writers while others identified writing as a hobby or secondary profession. All participants reported a basic understanding of commonly used visualizations such as bar chart and line chart. We referred to the participants as W1-W9 below.

3.2 Procedure

Interviews were semi-structured and lasted around an hour. One author of this paper conducted the interviews over Zoom. Each interview was divided into three parts. First, after gathering informed consent, we asked the writers to describe how they write creative materials. Our goal at this part was to understand the domain of creative writing. After that, interviewees shared their perspective on bias in creative writing, and several personal encounters of biases in their own writing, and/or in others' writing. In the final part, we asked the writers about the challenges they faced in addressing stereotypes in their writing, their current approach to overcome stereotypes, and how they thought digital tools could help them in this regard. At the end of the interview, each participant received a $15 worth Amazon gift card.

3.3 Analysis

Audio was recorded for all interviews, and anonymized transcripts were made for each. One author of this paper analyzed the interview data, following a thematic analysis process [14]. Throughout the analysis process, themes were refined, hypotheses developed, and relevant passages excerpted. We present the findings next.

3.4 Findings

Our findings relate to topics of general creative writing, implicit bias, mitigating bias, and the potential for computational support for such issues.

Understanding Creative Writing. Our first goal in this formative step was to understand the writing process for writers. The writing process heavily varied among the writers we interviewed. W1, W2, and W5 mentioned that they do not follow any particular workflow while W4 and W9 mentioned that they often employ different workflows for different stories. However, we were able to identify different components and stages of story writing (Figure 2). While we present the workflow in a sequential manner and there are a few natural orders among the components (e.g, writers can not get publish before they have a draft), the sequence is not rigid. For example, writers often reach out to their peer critique group (Figure 2c) before they actually start writing to understand the validity of the idea (W8). Thus, the components in Figure 2 are interchangeable, and may not even be present for some writers.

All writers mentioned that a story often starts from an imagination of a character, a scenario, or a personal experience (Figure 2A). Three writers (W1, W3, W8) considered themselves as *pantsers*, as they immediately start writing after the initial stage without any significant planning. In their case, the plotting of the story is closely intertwined with the writing. In contrast, other writers mentioned that they spend significant amount of time in plotting, idea generation, and research after which they start writing the story (Figure 2B). It is worth noting here that although pantsers do not usually plan at the beginning of a story, they may employ significant amount of time for researching and planning at the later stages of writing. The other components in Figure 2 (c, d, e) are part of an iterative process until the story is self-published or published by a publisher.

3.4.1 Types of Biases. The lack of diversity among the authors and characters was a recurring theme in the interviews. Participants mentioned that there are not enough authors from the marginalized groups and their stories are not told enough:

> *One recent study*[3] *found that there were more characters that feature animals, then there were of all of the minorities. I mean, it is horrendous, to be honest with you, in 2021... and it is really sad.* (W5)

However, several participants mentioned that they believe the situation has been improving, especially because of increasing awareness among writers and publishers. Beyond the lack of diversity among authors and characters, participants mentioned two nuance type of biases that they are concerned about while writing stories (described below).

Bias 1: Lack of agency for minority characters. One critical challenge for writers is to maintain agency for minority characters which ensures that these characters have impacts in the story and are not sidelined.

> *I believe it is Gail Simone. She is a comic book writer and she started this website called "Women in Refrigerators" because there is a very famous storyline from a Green Lantern comic (vol. 3, #54) where the Green Lantern comes home to find that his girlfriend has been murdered and shoved in his refrigerator. And the only reason that the woman really existed was to be killed. So it has become a name for this type of killing of female characters to advance the man's plot, which is unfortunately very common.* (W3)

W3 also referred to the *sexy lamp test*, which tests if one replace the female character of a story with a sexy lamp whether the story still make sense. W4 mentioned the *Bechdel* test, which asks whether a work features at least two women who talk to each other about something other than a man. Overall, a common theme among our creative writers was the sentiment that ensuring agency for minority characters is critical in creative storytelling.

Bias 2: Stereotypes in portraying minority characters. Another form of bias that the writers were concerned about is how minority characters are portrayed in the story. This includes how characters are described in the story and what actions they take.

> *One example would be Truman Capote's Breakfast at Tiffany's. I read the book as a teenager, and it's very beautiful and well-written overall. But one example of a stereotyped character is Holly's neighbor Mr. Yunioshi.*

[3]https://www.theguardian.com/books/2020/nov/11/childrens-books-eight-times-as-likely-to-feature-animal-main-characters-than-bame-people

> *He is a side character, but is portrayed simply as the irritable neighbor who always tells Holly off for forgetting her keys. This depiction of East Asian characters as grumpy old people is very stereotypical and common in Western literature. The image is further perpetuated by the Breakfast at Tiffany's movie where Mr. Yunioshi is played by Mickey Rooney (a white man). Rooney wears dramatic, stereotypical "East asian" clothing and makeup and runs around the top floor of the apartment yelling at Holly (played by Audrey Hepburn) in a terrible imitation of a Japanese accent.* (W7)

W2 and W5 mentioned that Male African-Americans characters often appear in comical scenarios in Western movies. W1 mentioned that female characters are often described as homely, beautiful, and sacrificial.

Methods for addressing bias in creative writing. Writers mentioned that acknowledging that their own filters and stereotypes may be propagated in their writing, despite their best intentions, is critical for them. They are self-aware of any potential stereotypes and misrepresentation in their writing.

> *I am very concerned about my own writing and it is not that I am afraid people will not like my writing. It is more that I do not want to be misrepresenting people cause I know how that feels.* (W7)

> *You have to be open to being wrong and to listening that you are wrong and accept that and then find out why. As a human, I have my own filters. The best I can do is try to overcome them, and admit those filters to the person who is reading my work.* (W4)

> *That whole aspect of write what you know is honestly very limiting, cause if you only write what you know then you are writing about a very narrow area. You have to write what you wanna know and you have to go out and look for information and look for ways that you can portray what you do not know in a way that is actually accurate.* (W3)

The process for avoiding stereotypes often starts from the research (Figure 2B) they conduct at the beginning of writing (W3-4, W6-7).

> *If I feel that I do not know how to write a character of a specific race, for instance, then I will probably find other books that are written by authors who are of that race and have characters of that race and see how they have written them. I find that helps.* (W7)

During writing, writers employ several self-evaluation techniques. After writing for a period of time, they evaluate the words they used for characters (W1-4, W6, W9), specific interactions between characters (W2-5, W7-9), and a part of the story (W1-9). However, this process becomes challenging and tiresome as the stories become longer and complicated (W1-3, W5, W7, W9).

Finally, after having a satisfying draft, writers seek feedback from critique groups (Figure 2C) about the overall story. A critique group may be informal, consist of one person, and be among the friends and family of a writer (W1-3, W6-7). More formal critique groups, especially for professional writing, include focus groups, and "Sensitivity Readers", a group that especially look for harmful misrepresentation (W3-4).

Potential and requirements for an interactive tool. At the final part of the interviews, writers brainstormed with the study administrator (one of the authors) about how interactive tools can help them in mitigating stereotypes. While critique groups and sensitivity readers provide important feedback to writers, they are usually available at the post-writing stage. During writing, an analytic tool for investigating a writer's own work for potential stereotypes could be helpful (W1, W3-4, W6-8). Writers expressed their enthusiasm as well as a few requirements for such a tool.

First, writers wondered whether the tool would support character development and story writing since the various bias mitigation strategies (Section 3.4.1) are part of the overall process (W2, W4-5, W8). Additionally, the tool would be ideal if it is not overwhelming and distracting since most of the times writers only want to concentrate on writing (W2-5, W8).

Second, writers expressed their concerns on how the tool could support bias mitigation:

> *I use Grammarly Pro where it sometimes recommends to avoid some words. I agree with it sometimes. And then sometimes it says a word is incorrect, and I am like this is my identity and I am allowed to say this word this way. Eventually, I stopped using that feature.* (W3)

> *I am wondering how you could support us. Language around minority groups is fluid and continuously evolving. And it is also important to be inclusive. There is a full spectrum of how a person could identify themselves, even within a community (such as LGBTQ+).* (W4)

Additionally, creative storytelling can be subjective to a writer's own experience, or imagination (W2), and may elicit intended stereotypes for the purpose of the plot (W1, W3, W4). Thus, writers suggested that rather making decisions for them, the analytic tool should support them in making informed decisions.

4 DESIGN GUIDELINES

Based on the findings of the formative interviews, we derive the following design guidelines.

DG1: Explore Agency for Characters and Social Identities. The formative interviews revealed that writers are concerned about the lack of agency for characters that represent marginalized groups. To ensure agency for characters from marginalized groups, the writer often keep track of the characters as well as their social identities. Thus, our tool should support writers in this task. Based on Section 3.4.1, the nature of this support should be exploratory, rather automated. The interactions should be designed for easy navigation to find specific passages of interest.

DG2: Explore Stereotypes in Describing Characters. Another important concern raised during the formative interviews was the stereotypes used for characters from marginalized groups. Writers mentioned that they search for possible stereotypes in the words which describe characters (adjectives), and the action characters take in the story (verbs). As such our tool should help writers in finding relevant words for a character, or a group of characters.

DG3: Integrated into the Character Development and Self-evaluation Process. Writers employ significant planning, and research for character development and various bias mitigation strategies discussed in Section 3.4.1 are integrated into this process. Thus, the bias exploration process should not be a separate process, rather be integrated into the natural character development process.

DG4: Unconstrained in Defining Identity. The formative interviews suggest that writers use diverse and intersectional identities for characters that may not conform to any pre-defined categories. Thus, our design should be inclusive, and allow authors to assign identities without any restrictions. Our tool should also avoid any identity detection algorithms/models as these models are often limited to binary identities.

DG5: Unobtrusive and Easy-to-understand Visualization. Too many visual components can hinder the natural process of writing, irritate authors, and make the tool overbearing (Section 3.4.1). Thus, the writing component of the tool should have the central focus with minimal extraneous visual components for bias exploration. Additionally, to

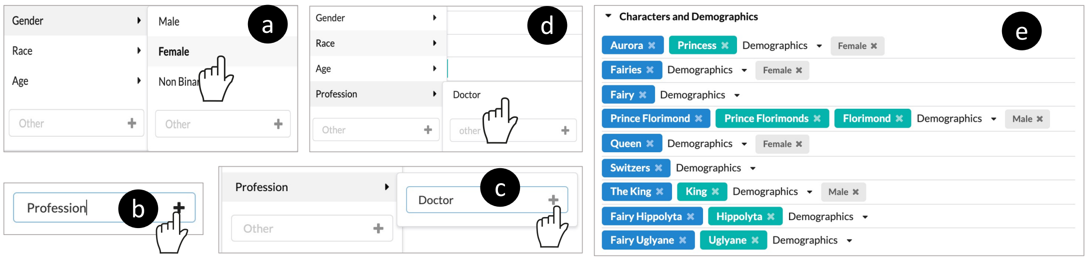

Fig. 3. **Demographics interface.** (a) A dropdown menu with different social identities is available for each character in the Character and Demographics Panel. (b) The dropdown can be extended dynamically. User can add a new identity (e.g., Profession). (c) User can then add categories under the newly added identity (e.g., Doctor). (d) An example demographics panel for the children’s book, *Sleeping Beauty*, retold by Arthur Quiller-Couch and Charles Perrault, freely available under Project Gutenberg [10].

ensure accessibility for writers who are not familiar with complex data visualization paradigms, our tool should use easy-to-understand visualizations.

DG6: Scalable Computation and UI. Our tool should support stories of different length and provide feedback to the writers in a short response time. However, analyzing a large textual data can be computationally expensive. Thus, our tool should have the mechanism to process large textual data in a scalable way. Similarly, the visual components should be able to show information extracted from large textual data.

5 DRAMATVIS PERSONAE

DramatVis Personae (DVP) is a visual analytics system to support off-line and on-line creative writing and analysis, particularly in character development, story arcs, and mitigating stereotypes and bias. We present the visual components and analysis pipeline for DramatVis Personae below.

5.1 Visual Interface

In this section, we demonstrate four visual components (Figure 1) of the interface. We link the design guidelines from Section 4 wherever applicable. We also discuss our design rationales and design alternatives considered during the development of the tool. The interface supports analysis for three different entities: (a) Characters, (b) Social Groups (e.g., Male, Female), and (c) Intersectional Social Groups (e.g., Male Doctors). We use the term “entity” or “entities” to refer to the three entities together whenever applicable for brevity. For demonstration purpose, we use several well-known Western stories.

Text Editor. The central component of DramatVis Personae is a text editor (**DG3, DG5**). The text editor is equipped with traditional formatting features such as selecting font, font size, font weight, etc. We use QuillJS [4] as a rich text editor which has been widely used in the web including social media apps such as Slack, and Linkedin. The unit of measure for QuillJS is single character (symbol) and Delta [1], a JSON formatted data structure, stores changes to the text editor. This helps us in tracking changes to the text editor conveniently.

Characters and Demographics Panel. The Characters and Demographics Panel lists all the named entities identified by our NLP pipeline. Additionally, a user can also manually highlight a character missed by our pipeline in the text editor to be added to this panel. The panel provides several validation functionalities such as deleting a character from

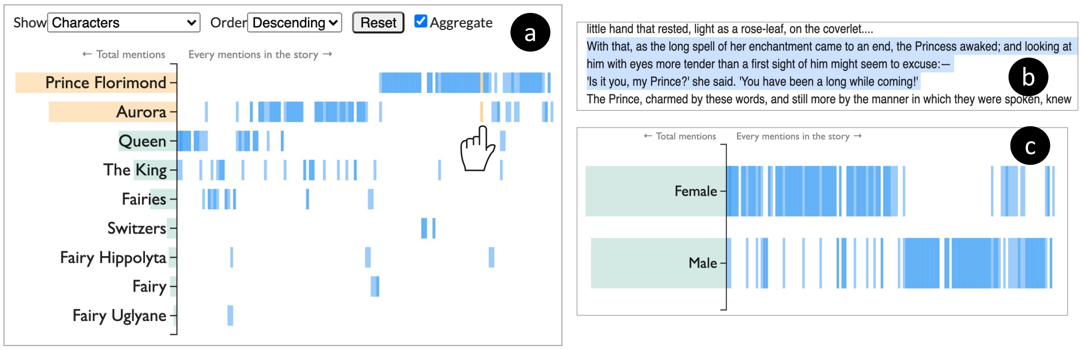

Fig. 4. **Timeline interface.** (a) An example timeline based on the characters of *Sleeping Beauty*. Note that even though the story is about Aurora, the sleeping beauty, it was Prince Florimond who had the most number of mentions. Note also the lack of overlaps between the timelines of two characters. (b) Upon hovering over a tile from the timeline, a user can see the relevant passage in the editor. The relevant passage in this case is the first scene after the Princess wakes up. (c) A user can also visualize the presence of different demographics in the timeline. We only used Male and Female for this example. Note that when aggregated, the female presence in the story is slightly larger than the male presence.

the list, and merging different names of the same character (i.e., aliases) together. A user can use drag and drop for merging characters. The blue labels show the name of a character while labels colored with Teal represent aliases.

Each character in this panel has a dropdown named *Demographics*. Using this dropdown, a user can add multiple social identities to each character (Figure 3a). We populate the dropdowns with commonly used identities. However, to support **DG4**, we made the dropdowns dynamically extendable. A user can add any number of new identities in these dropdowns. For example, Figure 3b and 3c show how a user can add Profession as a new identity and Doctor as a profession in the dropdowns.

Timeline. The timeline visualization is divided into two parts. On the left side of y-axis we encode the total number of mentions for entites using bars encompassing the axis labels. The y-axis can be sorted either in descending or ascending order. A user can choose the sort order from the *Order* dropdown. The dropdown defaults to descending order for showing the prominent entities at the top. On the right side of y-axis we show individual mentions for each character. The x-axis represents a linear scale with a range $(1, S)$ where S is the total number of sentences. For a mention of an entity in sentence s, we draw a tile (rectangle) with width $(pos(s) + 0.5) - (pos(s) - 0.5)$ where $pos(s)$ represents the position of s in the x-axis. We used linear scale instead of ordinal scale to make the adjacent tiles connected and smoother.

For larger documents, due to space constraints, the tiles can become extremely small, making it difficult to interact with them. To scale the visualization to larger documents (**DG6**), we added an *Aggregate* option in the toolbar of the timeline (Figure 4a). In the case of a document with more than 500 sentences ($S > 500$), the option is automatically triggered, and the document is binned together to restrict the x-axis to range $(1, 500)$. At that point, the x-axis represents passages, instead of sentences. The user can disable the aggregate option and see the timeline in terms of single sentence. As an alternative, we considered making the x-axis scrollable. The tile size would have remained same for any documents in that design. However, we noticed that this design does not provide full overview of the timeline together and it becomes difficult to compare different parts of the timeline as the user scrolls left and right in the timeline.

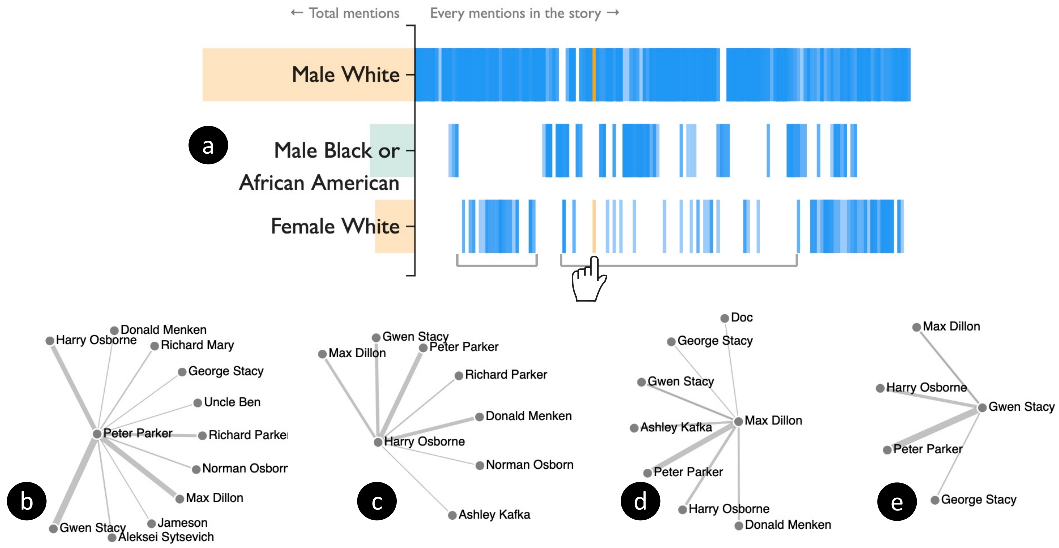

Fig. 5. **Timeline interface and impact graphs** (a) An example timeline showing the presence of three intersectional identities (*Male White, Male Black or African American*, and *Female White*) in the Movie *The Amazing Spiderman 2*. Male White characters are present throughout the storyline. Further, there appears to be a lack of interactions (identified by the grey encompassing x-axis lines) between Male Black or African American and Female White characters except for a few aberrant ones. However, both groups have interactions with Male White characters. The orannge bars show one such interaction between Male White and Female White characters. (b-e) Example impact graphs for the characters Peter Parker, Harry Osborne, Max Dillon, and Gwen Stacy.

Figure 4a shows an aggregated version of the timeline for the children’s book, *Sleeping Beauty*, retold by Arthur Quiller-Couch and Charles Perrault, freely available under Project Gutenberg [10]. The opacity of a tile represents the number of times an entity was mentioned in a particular passage. A user can hover over any tile and see the relevant passage, highlighted in the text editor (Figure 4b). Based on the identities defined by a user in the Characters and Demographics Panel, the timeline can also show the presence of different identities in the timeline (Figure 4c).

Using the *Show* dropdown in the timeline (Figure 4a), a user can choose to show the timeline for characters or demographics. A user can also choose to show the presence of intersectional groups in the timeline. For example, Figure 5a presents a timeline showing the presence of intersectional groups in the Movie *the Amazing Spiderman-2*. Note that for the sake of brevity, we do not show the Characters and Demographics Panel for this movie. For the relevant Characters and Demographics Panel, please refer to the supplemental video. We also note that the race assigned for the characters of this movie is based on the perception of the authors and may be different in reality.

In summary, the goal of the timeline visualization is to facilitate writers investigating agency for characters and social identities (**DG1**). Since the task is subjective to a story and a writer’s perspective, our intention was to expose potential gaps in the story and let a writer explore and navigate the story easily (**DG1**). We figured visualizing the mentions of the characters in a timeline can be a starting point for this task and will allow users to see gaps easily. The user can then use the timeline to further investigate any part of the story (**DG1**). To aid this process, other visual components are also connected to the timeline, as described below.

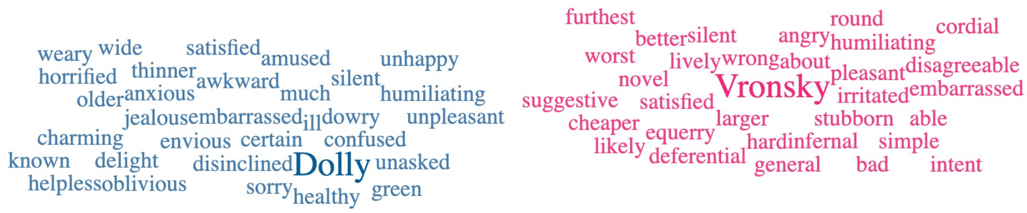

Fig. 6. Word Zone representation for Dolly and Vronsky from *Anna Karenina* (1877) by Leo Tolstoy, publicly available under Project Gutenberg [75].

Impact Graph. The timeline visualization primarily shows the presence of entities in a story. While a user can identify the interactions between entities by observing overlaps in their presence in the x-axis, timeline does not give a definitive answer to how an entity interacted with the other entities. Additionally, multiple entities can be mentioned in a passage of a story; however, that does not necessarily mean they interacted with each other.

To overcome this shortcoming, we introduced the *Impact Graph*, a force-directed network visualization that shows interactions between entities (**DG1**). We consider an interaction between two entities if they are mentioned together in a sentence. The impact graph for an entity is available whenever a user clicks on a y-axis label in the timeline.

Figure 5b-d show impact graphs for Peter Parker, Harry Osborne, Max Dillon, and Gwen Stacy from the movie *The Amazing Spiderman 2*. The selected entity is placed on the center of the network. The edge-widths represent the strength of the interactions. For reducing clutter, only edges with at least five interactions are shown. We observe that the impact graph for Gwen Stacy is much smaller than the other three male characters. Further, Gwen Stacy has significant interactions with Peter Parker only.

Word Zones. Word clouds are a popular visual representation for showing a collection of words. They have aesthetic value to lay users, and are fun, and engaging [31, 78]. In contrast, researchers have shown that they are not well-suited for analytic tasks such as finding a word and comparing the frequencies of words [31]. To balance the utility and aesthetic value of word clouds, Hearst et al. [31] proposed Word Zones, a variation of word clouds, where words are grouped based on predefined labels/categories. Since our users will most likely be non-experts in terms of visualization expertise, we decided to use Word Zones, thus opting for a visualization that is expected to be well-known to the writers as well as has better representation for analytic purposes (**DG5**).

A writer can add an entity for seeing words used with the entity in the word zone visualization whenever a user clicks on a y-axis label in the timeline (**DG2**). A writer can add as many entities as they want in the word zone. A user can control the number of words to show for each entity in the word zone using a slider. User also has a dropdown to see relevant adjective, verbs, or both. This helps exploring questions such as: *How the females are described in the story? What actions females are taking in the story?* (**DG2**).

We considered each entity as a document and the story as a corpora of documents (entities). Based on that, the weight of a word (w) for an entity (e) is calculated as:

$$weight(w, e) = tf(w, e) * (1/df(w)) \tag{1}$$

where $tf(w, e)$ is the frequency of w in e and $df(w)$ is the frequency of w in the whole story. It is essentially a normalized version of *tf-idf*, popularly used for filtering out common and stop words, finding words of interest. However,

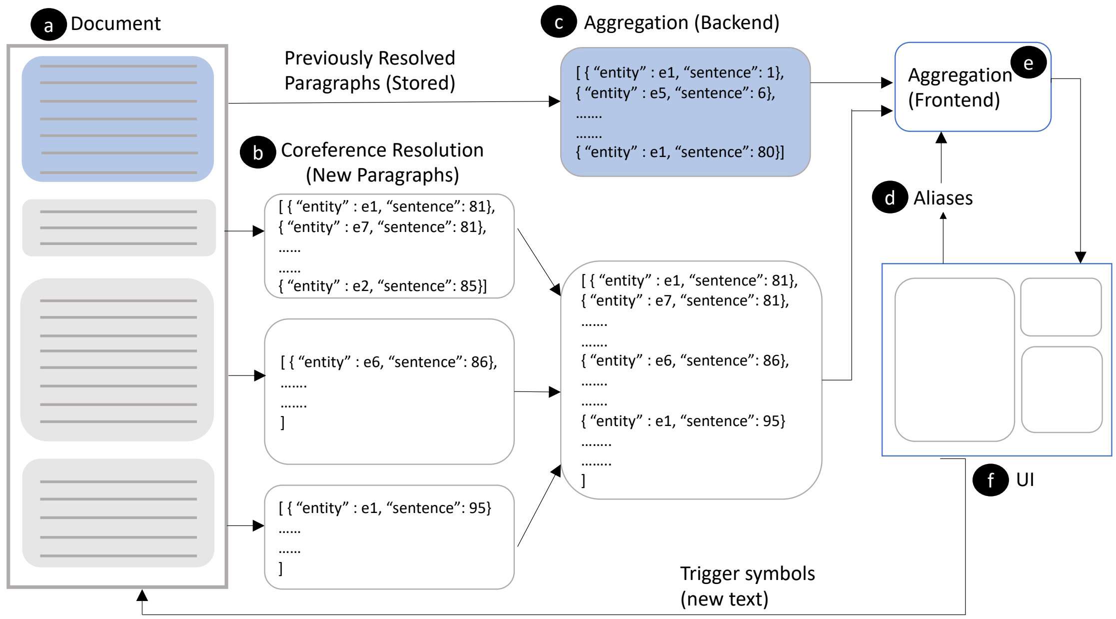

Fig. 7. **Example workflow.** a) the system detects three new paragraphs in the text editor. The blue box represents the previously written text while the grey boxes represent the newly added paragraphs. b) the system runs co-reference resolution and entity extraction independently on the newly added paragraphs. c) They are then aggregated together with the stored entity information of the previous text. e) A second aggregation on the front end is applied based on the aliases (d) identified by the user. f) Finally the JSON formatted data is fed to the UI. The UI listens for the trigger symbols such as ".", "?", "!" or copy/paste event in the text editor for updating the UI.

tf-idf is often applied on large corpora. In our case, the number of entities is limited to at most a few hundreds. We also consider adjectives and verbs only. Thus, we opted for a simple normalization.

Figure 6 shows an example word zone showing adjective used for Dolly (a female character), and Vronsky (a male character) from *Anna Karenina* (1877) by Leo Tolstoy. Note the words such as "charming", "envious", "jealous", "helpless", "oblivious" and others in Dolly's word zone.

Finally, while word zone provides a easy way to investigate adjective and verbs, the search space to explore all the combination of entities can be large. To aid writers in this process, we introduced a word embedding based approach to find potential candidate pairs of entities for investigation. Let $\vec{g}1$ be the mean vector representing the words from entity $ent1$ and $\vec{g}2$ be the mean vector representing the words from entity $ent2$. Then, the cosine distance between $\vec{g}1$ and $\vec{g}2$ can indicate how different the words from $ent1$ and $ent2$ are. This method closely matches the *relative norm difference* [26, 28], previously used to quantify gender bias in word embeddings. Any pair having a cosine distance more than 0.5 are fed to the interface as a notification in the word zone control panel.

5.2 Entity Extraction and Coreference Resolution

We used NeuralCoref [3] along with Spacy [32] for extracting named entities, and their mentions in a document (Coreference Resolution [39]). Both packages are considered state-of-the-art and widely used in different applications. However, coreference resolution for long documents such as a book can take a lot of time and system memory [76].

This could hinder the usability of the tool as our tool is targeted as a writing tool that provides instant feedback to writers as they write (**DG5, DG6**).

To increase the scalability of computation and reduce processing time, we adopted a divide-and-conquer method and run the coreference resolution model paragraph wise, instead of the full document together. We noticed that it is unlikely that a character would be referenced with a pronoun without the actual noun of the character in a new paragraph. We reached out to the participants from the interview studies and confirmed that this is the usual case. Figure 7 presents an example scenario of our method. On detecting a new trigger symbol such as ".", "?", "!" or a copy/paste event, the UI sends the document, and the current Delta [1] object. The Delta object contains information about how many and what characters (symbols) are inserted, deleted, or retained since the last update.

Upon receiving the document, the server splits the document in paragraphs. Based on the Delta object, the server then determines which paragraphs are retained (same as the previous update), and which paragraphs have new contents (insert or delete). The paragraphs are identified by double newlines or "\n\n". In Figure 7, the server detects three new paragraphs (grey boxes). The server then runs entity recognition and coreference resolution models on the newly detected paragraphs. Although our current implementation processes the paragraphs sequentially, they can be processed in a parallel fashion since the paragraphs are mutually independent. The information extracted from the newly added paragraphs are then aggregated together with the stored information of previously processed paragraphs. The client side then performs another aggregation to combine aliases together.

5.3 Implementation Notes

DramatVis Personae is a web-based authoring tool. We used Python as the back-end, and JavaScript as the front-end language and D3 [13] for interactive visualization. Semantic-ui and Bootstrap was used for styling various visual objects. We employed MongoDB for storing data structures and user information for SignIn and SignUp. The anonymized version of the source code is currently available here: https://osf.io/e42y7/?view_only=65adc608060f417c821b433aa5361679. If accepted, we will make the tool open-source and share the Github page.

6 EXPERT EVALUATION

We conducted follow-up study with three creative writers from our earlier study to evaluate DramatVis Personae. For this first evaluation, our goal was to find usability issues in different use cases so as to answer questions such as: Is the visualizations understandable? Can experts gain insights about their stories from this interface? Does the interface help them find potential stereotypes? To achieve that we wanted participants to have hands-on experience with our tool. We designed two case studies for this purpose: (1) writing a new story using our tool; and (2) analyzing previously written story either by the participant, or other writers.

6.1 Participants

We contacted 5 participants independently who participated in the formative interviews via email, requesting them to participate in a follow up study. Three participants (W1, W2, W3) agreed to participate in the study. Their self-reported demographics are: W1: male (he/him), white, in his 40s, with writing experience for more than 20 years; W2: female (she/her), asian, 25 years old, and has been writing since early childhood. W3: female (she/her), black or african american, 44 years old, with five years of experience in writing children's books.

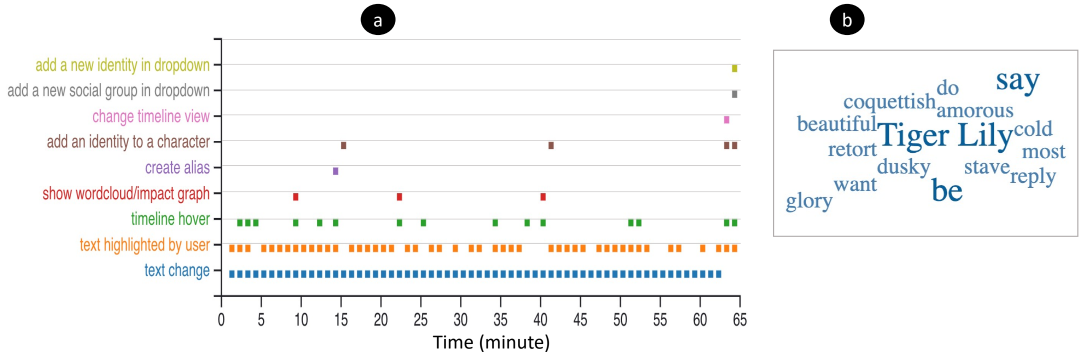

Fig. 8. **Study Results**. (a) Log analysis for case study 1. X-axis represents time in minutes while y-axis represents different types of interactions. (b) A word zone created by W2 for the character Tiger Lily in *Peter Pan*.

6.2 Method

One author of this paper administrated the study. After consent, the administrator demonstrated the tool with a sample story. We encouraged participants to ask questions at this stage. Once participants felt comfortable with the tool, we asked them to write a story using our tool, or analyze already written stories. While using the interface, participants thought-aloud and conversed with the administrator regularly. The sessions were conducted in ZOOM. The sessions were audio and video recorded for post-session analysis. At the end of the interview, each participant received a $15/hour worth Amazon gift card.

6.3 Case Study 1: Writing a Story using DramatVis Personae

In this case study, we investigated how a creative writer can write a story using DramatVis Personae while observing the visual feedback from the interface. To realize this, we asked W1 to write a story using our tool in a two-hour long session. However, writing a story instantaneously is not a trivial task; it requires significant planning and considerable amount of time to complete. Considering this, we provided the following story prompt to W1 before the study session:

> "Write a short story of approximately 1,000 words on the theme of *redemption*: the main character overcoming flaws or accusations and redeeming themselves. The setting, premise, and plot is entirely up to you besides these basic parameters. The story should include at least three significant (named) characters and any number of supporting characters, at least two different scenes, and some dialogue between the characters. You are welcome to ideate and take notes on your story and characters, but please do not begin the actual writing process until the session."

We asked for a short story to reduce the study length to a reasonable time and reduce cognitive load for the participant. Note that we communicated with W1 multiple times during this pre-study process to make him comfortable.

Figure 8 shows usage logs from the study. As expected, W1 interacted with the text editor most for writing the story. W1 was curious whether our system could identify the mentions correctly, resulted in the periodical interactions with the bars in the timeline (green rectangles in Figure 8). Initially, W1 expressed concern about missing out a few mentions in the timeline. The functionality of adding characters and mentions manually reduced such concern. We clarified further that the system may still miss out a few references, and the interface is designed to provide an overview, not exact representations. As the story grows at length, the visualizations will provide a better overview. W1 was satisfied by this explanation.

Around 15 minutes into the session, W1 first used the Character and Demographics panel to create an alias, and assign identities to the characters. W1 mentioned this is an important feature for him as he usually plan very little before writing (he is a self-professed pantser), and this helped him plan the story early into the writing and made him think how the characters should be represented in the story. At this time, W1 hid the visualizations to concentrate only on writing. However, W1 periodically checked the timeline, word zones, and impact graphs (red rectangles).

W1 finished the first draft of his story in about an hour and half. The story is a 1,100-word flash fiction titled *Sacrifice* about a young female soldier saving her king and fellow soldiers during an ill-fated siege by inhuman invaders after initially failing to protect the gate. After finishing the story, W1 examined the timeline and changed the timeline view to Intersectional Demographics. He noticed a gap in the timeline (see supplemental for the relevant timeline). Immediately, W1 examined that section of the story. W1 explained that the particular section is the start of the second scene which describes the scene without any particular character in it. W1 suggested in a longer story, he would analyze the story after every chapter.

On the final notes, W1 suggested a few minor stylistic changes such as auto-indenting a new paragraph, adding a thesaurus, etc. Except for these suggestions, W1 expressed his admiration for the usability of the tool, suggesting it is easy to use and understand, and not overwhelming.

6.4 Case Study 2: Analyzing an Already Written Story using DramatVis Personae

The formative interviews suggested that reading others' work is an important part of research and help writers in understanding how writers from marginalized groups write about their own community. In this case study, we aimed at understanding how DramatVis Personae might be useful in this task. To realize this, we asked W2 to analyze a well-known Western story in a one-hour session. At the start of the session, we provided W2 with 5 well-known children's book. W2 chose *Peter Pan* for this analysis.

while analyzing, W2 used the word zones extensively in the session. W2 mentioned that this is a process she carries out regularly. She especially liked the option to see the adjectives, verbs, or both of them. The word choices are very important for avoiding stereotypes, according to W2. For example W2 found the Figure 8(b) to be an example of such biases. The usage of stereotyped words such as "amorous", "beautiful", "coquettish" for a Native American Character (Tiger Lily) is extremely problematic to her. W2 mentioned DramatVis Personae could become a handy tool for writers to quickly analyze stories, and perform case studies. W2 appreciated the design of the interface and did not find any major usability issues.

Following W2, we asked W3 to analyze a story that she have written previously. W2 explored different features of the interface for potential stereotypes. However, W3 did not find any significant bias in her writing. One of the animal character was not identified by our tool. W3 wondered whether our tool can identify animal characters as she particularly writes for children. She also mentioned that her sci-fi stories often have names that are unreal for a person. Finally, W3 suggested that the interface will be even better if she could see the entities highlighted in the text editor.

7 DISCUSSION, LIMITATIONS, AND FUTURE WORK

In this section, we discuss potential implications and limitations learned from the development and evaluation of DramatVis Personae.

Impact on Bias in Storytelling. Biases in creative writing is ubiquitous, as described in Section 2.3. Our tool is targeted towards reducing implicit biases in creative writing. However, we believe it is important to acknowledge that this is one of the modalities or reasons behind bias in creative writing. For example, there is a lack of writers from marginalized groups in the current writer's community [62]. We also acknowledge that beyond the two types biases explored in this work, our interface may not be helpful in addressing other sort of biases. Therefore, we consider DramatVis Personae as a part of the larger movement against biases in creative writing, a probe for reducing biases, and a catalyst for future efforts in this direction. We hope its adaptability to different forms of writing (novels, children's books, short stories) will attract diverse writers. While our current version is intended for use during writing, we envision several extensions of our interface.

First, DramatVis Personae can be used as a learning tool to raise awareness among novice writers. They can use it to study books or movies known for misrepresenting marginalized groups. Similarly, writing communities and literary schools can use it as an teaching material to quickly perform case studies, a common practice in these institutions.

Second, writers and publishers may use our tool to communicate summary statistics about the book to the readers. We believe this practice has the potential to increase accountability among authors and increase trust among writers and readers.

Finally, while our tool can analyze textual data, in its current form, it cannot analyze pictures that may be present in a children's book. A natural extension would be to add an image analysis pipeline in the existing system to find signals in pictures as well.

Balancing Automation and Artistic Freedom/Agency. From the beginning of this project, we were vigilant about the artistic freedom of writers, and how we can balance agency with automation. That is one of the reasons why we did not use any automation in detecting biases and stereotypes. Violating writers' freedom and suggesting wrong interpretation of their work can severely lower their trust in the system. However, this leads to a potential limitation of our tool: a writer can still write harmful stereotypes and there is no way to automatic way to perform corrective measures if the writer is not wiling to self-evaluate themselves.

Furthermore, there is no such thing as an unbiased algorithm. Our approach will not eliminate biases because not all biases will be flagged as such. However, we hope that it will help in reducing them.

It is also worth noting here that our intention was not to replace critique groups or sensitivity readers. We believe it is important to have subjective feedback about creative writing, especially from relevant marginalized groups. Our intention was to help writers during their writing process by offering another set of eyes—albeit mechanical ones.

Design Implications for Human-centered AI. AI systems can inherit harmful biases and stereotypes. These biases can impact social groups disparately, especially when used as a decision-making platform for critical resources [9, 71]. These systems may also lack inclusivity (e.g., lack of supports for non-binary identities [33]). These limitations motivated several design decisions of our system.

First, based on **DG4**, we made the *demographics* dropdowns unconstrained so that a writer could add any social identities required for the story. Second, we provided several functions in the interface for the writers to validate (e.g., merge, delete) the results returned by the NLP, a safety check against potential biases in coreference resolution [81], and dependency parsing [27]. Additionally, a writer can interactively add an entity for tracking in the case the entity was not recognized by the NLP pipeline. Thus, the design of our interface suggests accepting the limitations of the underlying AI early in the project, and including the targeted users in the design phase can reduce concerns around potential biases in a human-in-the-loop interface.

Design Implications for Visual Analytics. The design of DramatVis Personae offers several design implications for visual analytics research. Over the years, researchers have made enormous efforts towards inventing increasingly sophisticated forms of visualization techniques. Our research suggests that the simplest and well-known form of visualization may all that is required for a usable interface, especially for users with no specific expertise in visualization.

Similarly, dashboards are common in visual analytics research. However, we deliberately kept a minimalist design in terms of visual components and interactions between them for making it less overbearing. Of course these implications are dependent on the application domain and task at hand. Nevertheless, we hope our research will motivate design of future creativity interfaces to support writers and in a broader sense interfaces for text visualization.

Of Pantsers and Plotters. An important finding revealed early in our work is that creative writers can be organized into two camps depending on whether or not they started writing straight away or spent significant time preparing and planning: pantsers and plotters. Our goal when designing DramatVis Personae was to support both, but based on the NLP algorithms that we used, the tool admittedly works better if it is fed writing immediately. In other words, it is currently a better match for a pantser who begins to write without much planning, or perhaps only at a later stage in the process for a plotting when some text has already been written. Of course, the manual creation and editing of characters will also help a plotter, but as with any visualization tool, DramatVis Personae is data-driven and works best when there is actual data to be had.

Limitations of the Underlying AI. NLP research has made big strides in recent years but still we cannot expect perfect results for problems like coreference resolution. In our context, this might result in certain character(s) and/or their mention(s) being unaccounted for in the tool. The case studies also found concerns around appropriate finding of the non-traditional characters and characters that appear as animals. It is also worth noting that different NLP models like coreference resolution, named entity recognition might exhibit biases based on social categorizations like gender [53, 81]. This might result in skewed performance against minority groups [53, 54]. Moreover, the NLP models used in this work were trained on a set of news articles, weblogs, etc. Such data might differ from literary texts like novels books, etc. as they might contain longer sentences, more sophisticated language, etc. [11, 69]. Future work might employ fairness aware NLP models which are trained on domain specific dataset like books, novels, etc. for better performance.

The time to process text increases proportionally with the amount of text. Higher processing time will add more latency on the front-end which might negatively impact user experience. In this work, we made an informed decision to choose NLP models which provide good performance while providing fast response time. Based on the current trend in NLP literature, future NLP models might be more computationally expensive and provide better performance. To incorporate such models, future work might prefer to conduct text analysis periodically instead of on-demand analysis to preserve user experience.

Language Support. The design and development process of this tool has been influenced by English-speaking study participants and NLP models trained over English language corpus. Hence, our tool can currently only deal with English language. Having said that, we know that social biases based on gender, etc., transcend societies and languages. We believe that ideas put forward in this work will help facilitate the development of similar tools for other languages as well. Future work might support others languages like French, Chinese, etc. by incorporating NLP models trained over different language corpora.

8 CONCLUSION

In this paper, we have presented an in-depth case study on supporting creative writing and combating implicit bias in fiction using interactive technologies, data visualization, and natural-language processing. In particular, we have reported on an interview study involving 9 creative writers where we asked them about their process, how they navigate harmful stereotypes, and how they think tool support could help in this work. Based on these interviews, we design DramatVis Personae, a visual analytics tool using natural language processing to visualize characters, their demographics, and their story arcs in an effort to mitigate implicit bias. The tool can be used both in an online manner while writing a story, as well as offline during analysis of an already written story. To test this premise, we conducted two case studies with creative writers using DramatVis Personae, one to generate *Sacrifice*, an original flash fiction story of 1,100 words, and one to analyze an existing story. The outcome of these studies support the utility of interactive technologies to support creative writing and storytelling, particularly in serving as an automated "critique group" for immediate feedback to the writer.

We believe that our work here suggests many interesting future avenues of research. For one thing, while creative writing is a notoriously individual and idiosyncratic process, and while the human touch is vital to true art, our moderate success with DramatVis Personae points to possible ways to augment this human touch to improve even such famously crooked processes. In particular, we think that our work shows how automatic machine eyes, while certainly less keen and discerning than human ones, can be helpful for certain applications such as mitigating bias if only because they—unlike human eyes—remain unblinking. We are not so foolish as to believe in the idea of an "unbiased algorithm"'—all algorithms are created by humans and thus intrinsically carry the biases of their creators—but we do believe in the virtue of training as many different lenses as possible on a creative artifact in the hope of uncovering yet another harmful stereotype or instance of implicit bias. Thus, we tend to think that our work here is in no way indicative of an end of art, but rather a new beginning.

ACKNOWLEDGMENTS

[Anonymized for double-blind review.]

POSTSCRIPT

You may shoot me with your words,
You may cut me with your eyes,
You may kill me with your hatefulness,
But still, like air, I'll rise.

– *Still I Rise* (excerpt), Maya Angelou (1978).

REFERENCES

[1] 2021. *Delta.* https://github.com/quilljs/delta/
[2] 2021. *Granthika.* https://granthika.co
[3] 2021. *NeuralCoref 4.0: Coreference Resolution in spaCy with Neural Networks.* https://github.com/huggingface/neuralcoref
[4] 2021. *QuillJS.* https://quilljs.com
[5] 2021. *Scrivener.* https://www.literatureandlatte.com/scrivener/overview/
[6] Alfie Abdul-Rahman, Julie Lein, Katherine Coles, Eamonn Maguire, Miriah D. Meyer, Martin Wynne, Chris R. Johnson, Anne E. Trefethen, and Min Chen. 2013. Rule-based Visual Mappings - with a Case Study on Poetry Visualization. *Computer Graphics Forum* 32, 3 (2013), 381–390. https://doi.org/10.1111/cgf.12125

[7] Anjali Adukia, Alex Eble, Emileigh Harrison, Hakizumwami Birali Runesha, and Teodora Szasz. 2021. What We Teach About Race and Gender: Representation in Images and Text of Children's Books. *University of Chicago, Becker Friedman Institute for Economics Working Paper* 2021-44 (2021).

[8] Aretha Barbosa Alencar, Maria Cristina Ferreira de Oliveira, and Fernando Vieira Paulovich. 2012. Seeing beyond reading: a survey on visual text analytics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 2, 6 (2012), 476–492. https://doi.org/10.1002/widm.1071

[9] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. Machine bias. *ProPublica, May* 23 (2016), 2016.

[10] Charles Perrault Arthur Quiller-Couch. 2016. The Sleeping Beauty and other fairy tales. https://www.gutenberg.org/files/51275/51275-h/51275-h.htm Accessed: 08/30/2021.

[11] David Bamman. [n.d.]. LitBank: Born-Literary Natural Language Processing. *Computational Humanites, Debates in Digital Humanities (2020, preprint), https://people. ischool. berkeley. edu/ dbamman/pubs/pdf/Bamman DH Debates CompHum. pdf* ([n. d.]).

[12] Angharad Beckett, Nick Ellison, Sam Barrett, and Sonali Shah. 2010. 'Away with the fairies?'Disability within primary-age children's literature. *Disability & Society* 25, 3 (2010), 373–386.

[13] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. 2011. D^3 data-driven documents. *IEEE Trans on Visualization and Computer Graphics* 17, 12 (2011), 2301–2309.

[14] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[15] Manish Chaturvedi, Gerald C. Gannod, Laura Mandell, Helen Armstrong, and Eric Hodgson. 2012. Myopia: A Visualization Tool in Support of Close Reading. In *Proceedings of the Annual International Conference of the Alliance of Digital Humanities Organizations*, Jan Christoph Meister (Ed.). Hamburg University Press, 148–150.

[16] Michael Correll, Michael Witmore, and Michael Gleicher. 2011. Exploring Collections of Tagged Text for Literary Scholarship. *Computer Graphics Forum* 30, 3 (2011), 731–740. https://doi.org/10.1111/j.1467-8659.2011.01922.x

[17] Weiwei Cui, Shixia Liu, Li Tan, Conglei Shi, Yangqiu Song, Zekai Gao, Huamin Qu, and Xin Tong. 2011. TextFlow: Towards Better Understanding of Evolving Topics in Text. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 2412–2421. https://doi.org/10.1109/TVCG.2011.239

[18] Niels Dekker, Tobias Kuhn, and Marieke van Erp. 2019. Evaluating named entity recognition tools for extracting social networks from novels. *PeerJ Computer Science* 5 (2019), e189.

[19] Pascale Emons, Fred Wester, and Peer Scheepers. 2010. "He Works Outside the Home; She Drinks Coffee and Does the Dishes" Gender Roles in Fiction Programs on Dutch Television. *Journal of Broadcasting & Electronic Media* 54, 1 (2010), 40–53.

[20] Ethan Fast, Tina Vachovsky, and Michael S Bernstein. 2016. Shirtless and dangerous: Quantifying linguistic signals of gender bias in an online fiction writing community. In *Tenth International AAAI Conference on Web and Social Media*.

[21] Xavier Ferrer, Tom van Nuenen, Jose M Such, and Natalia Criado. 2021. Discovering and categorising language biases in reddit. *International AAAI Conference on Web and Social Media (ICWSM 2021)* (2021).

[22] Vicki C Flerx, Dorothy S Fidler, and Ronald W Rogers. 1976. Sex role stereotypes: Developmental aspects and early intervention. *Child Development* (1976), 998–1007.

[23] The Centre for Literacy in Primary Education. 2020. *Reflecting Realities - Survey of Ethnic Representation within UK Children's Literature.* https://clpe.org.uk/research/clpe-reflecting-realities-survey-ethnic-representation-within-uk-childrens-literature Accessed: 08/30/2021.

[24] Samah Gad, Waqas Javed, Sohaib Ghani, Niklas Elmqvist, E. Thomas Ewing, Keith N. Hampton, and Naren Ramakrishnan. 2015. ThemeDelta: Dynamic Segmentations over Temporal Topic Models. *IEEE Transactions on Visualization and Computer Graphics* 21, 5 (2015), 672–685. https://doi.org/10.1109/TVCG.2014.2388208

[25] Qihong Gan, Min Zhu, Mingzhao Li, Ting Liang, Yu Cao, and Baoyao Zhou. 2014. Document visualization: an overview of current research. *WIREs Computational Statistics* 6, 1 (2014), 19–36. https://doi.org/10.1002/wics.1285

[26] Nikhil Garg, Londa Schiebinger, Dan Jurafsky, and James Zou. 2018. Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences* 115, 16 (2018), E3635–E3644.

[27] Aparna Garimella, Carmen Banea, Dirk Hovy, and Rada Mihalcea. 2019. Women's syntactic resilience and men's grammatical luck: Gender-Bias in Part-of-Speech Tagging and Dependency Parsing. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 3493–3498.

[28] Bhavya Ghai, Md Naimul Hoque, and Klaus Mueller. 2021. WordBias: An Interactive Visual Tool for Discovering Intersectional Biases Encoded in Word Embeddings. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–7.

[29] Martin Gronemann, Michael Jünger, Frauke Liers, and Francesco Mambelli. 2016. Crossing Minimization in Storyline Visualization. In *Proceedings of the International Symposium on Graph Drawing and Network Visualization (Lecture Notes in Computer Science, Vol. 9801)*. Springer, 367–381. https://doi.org/10.1007/978-3-319-50106-2_29

[30] Mykol C Hamilton, David Anderson, Michelle Broaddus, and Kate Young. 2006. Gender stereotyping and under-representation of female characters in 200 popular children's picture books: A twenty-first century update. *Sex roles* 55, 11 (2006), 757–765.

[31] Marti A Hearst, Emily Pedersen, Lekha Patil, Elsie Lee, Paul Laskowski, and Steven Franconeri. 2019. An evaluation of semantically grouped word cloud designs. *IEEE transactions on visualization and computer graphics* 26, 9 (2019), 2748–2761.

[32] Matthew Honnibal, Ines Montani, Sofie Van Landeghem, and Adriane Boyd. 2020. *spaCy: Industrial-strength Natural Language Processing in Python.* https://doi.org/10.5281/zenodo.1212303

[33] Md Naimul Hoque, Nazmus Saquib, Syed Masum Billah, and Klaus Mueller. 2020. Toward Interactively Balancing the Screen Time of Actors Based on Observable Phenotypic Traits in Live Telecast. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–18.

[34] Alexander Miserlis Hoyle, Lawrence Wolf-Sonkin, Hanna Wallach, Isabelle Augenstein, and Ryan Cotterell. 2019. Unsupervised Discovery of Gendered Language through Latent-Variable Modeling. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Florence, Italy, 1706–1716. https://doi.org/10.18653/v1/P19-1167

[35] Chieh-Jen Huang, Cheng-Te Li, and Man-Kwan Shan. 2013. VizStory: Visualization of Digital Narrative for Fairy Tales. In *Proceedings of the Conference on Technologies and Applications of Artificial Intelligence*. IEEE Computer Society, Piscataway, NJ, USA, 67–72. https://doi.org/10.1109/TAAI.2013.26

[36] Darnell Hunt and Ana-Christina Ramón. 2020. *Hollywood Diversity Report 2020.* https://socialsciences.ucla.edu/wp-content/uploads/2020/02/UCLA-Hollywood-Diversity-Report-2020-Film-2-6-2020.pdf

[37] Randy Ingermanson. 2009. *Writing Fiction for Dummies.* For Dummies.

[38] Kenneth Joseph, Wei Wei, and Kathleen M Carley. 2017. Girls rule, boys drool: Extracting semantic and affective stereotypes from Twitter. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*. 1362–1374.

[39] Daniel Jurafsky and James H Martin. [n.d.]. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition.

[40] Daniel A. Keim and Daniela Oelke. 2007. Literature Fingerprinting: A New Method for Visual Literary Analysis. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*. IEEE, Piscataway, NJ, USA, 115–122. https://doi.org/10.1109/VAST.2007.4389004

[41] Allen Kim, Charuta Pethe, and Steven Skiena. 2020. What Time Is It? Temporal Analysis of Novels. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 9076–9086.

[42] Nam Wook Kim, Benjamin Bach, Hyejin Im, Sasha Schriber, Markus H. Gross, and Hanspeter Pfister. 2018. Visualizing Nonlinear Narratives with Story Curves. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (2018), 595–604. https://doi.org/10.1109/TVCG.2017.2744118

[43] Stephen King. 2000. *On Writing: A Memoir of the Craft.* Charles Scribner's Sons.

[44] Richard Kolbe and Joseph C La Voie. 1981. Sex-role stereotyping in preschool children's picture books. *Social Psychology Quarterly* (1981), 369–374.

[45] Eve Kraicer and Andrew Piper. 2019. Social characters: the hierarchy of gender in contemporary English-language fiction. *Journal of Cultural Analytics* 1, 1 (2019), 11055.

[46] Vincent Labatut and Xavier Bost. 2019. Extraction and analysis of fictional character networks: A survey. *ACM Computing Surveys (CSUR)* 52, 5 (2019), 1–40.

[47] Heidi Layne and Amikeng A Alemanji. 2015. "Zebra world": The promotion of imperial stereotypes in a children's book. *Power and Education* 7, 2 (2015), 181–195.

[48] Shixia Liu, Yingcai Wu, Enxun Wei, Mengchen Liu, and Yang Liu. 2013. StoryFlow: Tracking the Evolution of Stories. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2436–2445. https://doi.org/10.1109/TVCG.2013.196

[49] Li Lucy, Dorottya Demszky, Patricia Bromley, and Dan Jurafsky. 2020. Content analysis of textbooks via natural language processing: Findings on gender, race, and ethnicity in Texas US history textbooks. *AERA Open* 6, 3 (2020), 2332858420940312.

[50] Nishtha Madaan, Sameep Mehta, Taneea Agrawaal, Vrinda Malhotra, Aditi Aggarwal, Yatin Gupta, and Mayank Saxena. 2018. Analyze, detect and remove gender stereotyping from bollywood movies. In *Conference on fairness, accountability and transparency*. PMLR, 92–105.

[51] Ariadna Matamoros-Fernández. 2017. Platformed racism: The mediation and circulation of an Australian race-based controversy on Twitter, Facebook and YouTube. *Information, Communication & Society* 20, 6 (2017), 930–946.

[52] Nina McCurdy, Julie Lein, Katherine Coles, and Miriah D. Meyer. 2016. Poemage: Visualizing the Sonic Topology of a Poem. 22, 1 (2016), 439–448. https://doi.org/10.1109/TVCG.2015.2467811

[53] Ninareh Mehrabi, Thamme Gowda, Fred Morstatter, Nanyun Peng, and Aram Galstyan. 2020. Man is to person as woman is to location: Measuring gender bias in named entity recognition. In *Proceedings of the 31st ACM Conference on Hypertext and Social Media*. 231–232.

[54] Shubhanshu Mishra, Sijun He, and Luca Belli. 2020. Assessing Demographic Bias in Named Entity Recognition. *arXiv preprint arXiv:2008.03415* (2020).

[55] Michel Mitri. 2020. Story analysis using natural language processing and interactive dashboards. *Journal of Computer Information Systems* (2020), 1–11.

[56] May M Narahara. 1998. Gender Stereotypes in Children's Picture Books. (1998).

[57] Cathrine Norberg. 2016. Naughty boys and sexy girls: the representation of young individuals in a web-based corpus of English. *Journal of English Linguistics* 44, 4 (2016), 291–317.

[58] Geena Davis Institue on Gender in Media. 2008. *Gender Stereotypes: An Analysis of Popular Films and TV.* https://seejane.org/wp-content/uploads/GDIGM_Gender_Stereotypes.pdf

[59] Geena Davis Institue on Gender in Media. 2019. *Historic Gender Parity in Children's Television.* https://seejane.org/research-informs-empowers/see-jane-2019/

[60] Geena Davis Institue on Gender in Media. 2019. *Representation Of Black Women in Hollywood.* https://seejane.org/research-informs-empowers/representations-of-black-women-in-hollywood/

[61] Geena Davis Institue on Gender in Media. 2020. *Historic Gender Parity in Family Films!* https://seejane.org/2020-film-historic-gender-parity-in-family-films/

[62] Geena Davis Institue on Gender in Media. 2021. *Behind the Scenes: the State of Inclusion and Equity In TV Writing.* https://seejane.org/research-informs-empowers/behind-the-scenes-the-state-of-inclusion-and-equity-in-tv-writing/

[63] Kelly Crisp Paynter. 2011. *Gender stereotypes and representation of female characters in children's picture books.* Liberty University.

[64] Michael Pearce. 2008. Investigating the collocational behaviour of MAN and WOMAN in the BNC using Sketch Engine. *Corpora* 3, 1 (2008), 1–29.

[65] Sharyl Bender Peterson and Mary Alyce Lach. 1990. Gender stereotypes in children's books: Their prevalence and influence on cognitive and affective development. *Gender and education* 2, 2 (1990), 185–197.

[66] Charuta Pethe, Allen Kim, and Steven Skiena. 2020. Chapter Captor: Text Segmentation in Novels. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 8373–8383.

[67] 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 (2016), 1–12.

[68] Randall M. Rohrer, John L. Sibert, and David S. Ebert. 1998. The Shape of Shakespeare: Visualizing Text using Implicit Surfaces. In *Proceedings of the IEEE Symposium on Information Visualization*. IEEE, Piscataway, NJ, USA, 121–129. https://doi.org/10.1109/INFVIS.1998.729568

[69] Ina Rösiger, Sarah Schulz, and Nils Reiter. 2018. Towards coreference for literary text: Analyzing domain-specific phenomena. In *Proceedings of the Second Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*. 129–138.

[70] Tara Ross. 2019. Media and stereotypes. *Journal: The Palgrave Handbook of Ethnicity* (2019), 1–17.

[71] Morgan Klaus Scheuerman, Jacob M Paul, and Jed R Brubaker. 2019. How computers see gender: An evaluation of gender classification in commercial facial analysis services. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–33.

[72] Artur Ŝilic and Bojana Dalbelo Basic. 2010. Visualization of Text Streams: A Survey. In *Proceedings of the International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (Lecture Notes in Computer Science, Vol. 6277)*, Rossitza Setchi, Ivan Jordanov, Robert J. Howlett, and Lakhmi C. Jain (Eds.). Springer, 31–43. https://doi.org/10.1007/978-3-642-15390-7_4

[73] John T. Stasko, Carsten Görg, Zhicheng Liu, and Kanupriya Singhal. 2007. Jigsaw: Supporting Investigative Analysis through Interactive Visualization. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*. IEEE, Piscataway, NJ, USA, 131–138. https://doi.org/10.1109/VAST.2007.4389006

[74] Yuzuru Tanahashi and Kwan-Liu Ma. 2012. Design Considerations for Optimizing Storyline Visualizations. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2679–2688. https://doi.org/10.1109/TVCG.2012.212

[75] Leo Tolstoy. 1998. Anna Karenina. https://www.gutenberg.org/files/1399/1399-0.txt Accessed: 08/30/2021.

[76] Shubham Toshniwal, Sam Wiseman, Allyson Ettinger, Karen Livescu, and Kevin Gimpel. 2020. Learning to Ignore: Long Document Coreference with Bounded Memory Neural Networks. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 8519–8526.

[77] Ya-Lun Tsao. 2008. Gender issues in young children's literature. *Reading Improvement* 45, 3 (2008), 108–115.

[78] Fernanda B Viegas, Martin Wattenberg, and Jonathan Feinberg. 2009. Participatory visualization with wordle. *IEEE transactions on visualization and computer graphics* 15, 6 (2009), 1137–1144.

[79] Katie Watson, Samuel S. Sohn, Sasha Schriber, Markus Gross, Carlos Manuel Muñiz, and Mubbasir Kapadia. 2019. StoryPrint: an interactive visualization of stories. In *Proceedings of the ACM Conference on Intelligent User Interfaces*. ACM, New York, NY, USA, 303–311. https://doi.org/10.1145/3301275.3302302

[80] James A. Wise, James J. Thomas, Kelly Pennock, David Lantrip, Marc Pottier, Anne Schur, and Vern Crow. 1995. Visualizing the non-visual: spatial analysis and interaction with information from text documents. In *Proceedings of the IEEE Symposium on Information Visualization*. IEEE, Piscataway, NJ, USA, 51–58. https://doi.org/10.1109/INFVIS.1995.528686

[81] Jieyu Zhao, Tianlu Wang, Mark Yatskar, Vicente Ordonez, and Kai-Wei Chang. 2018. Gender bias in coreference resolution: Evaluation and debiasing methods. *arXiv preprint arXiv:1804.06876* (2018).