AI Authorship: An Analysis of Short Story Writing

by Sofia Giannuzzi

sgiannuzzi@college.harvard.edu

Harvard University

Introduction

Generative AI and Large Language Models (LLMs) like ChatGPT are beginning to shape creative writing genres, even at times serving as full or co-authors of published works. One such genre that AI has been impacting is the short story genre. Interestingly, the short story, like other literary forms, is not a genre made up of hard and fast rules, but rather is shaped by the common conventions of contributing authors. Some of the conventional features that are recognized by literary theorists to make up the short story genre are brevity, distinctive endings, and a unity of impression (see Poe, Matthews, Patea, Pasco, May, Algee-Hewitt, et al.). With the rise of AI authorship, a pressing question emerges: how well can LLMs replicate these established genre conventions?

In creating a small corpus consisting of 25 human-written, traditionally popular short stories and 25 short stories generated by ChatGPT-4o (created with the prompt "You are a creative writer. Write a short story of 1000-7500 words"), we propose the use of the following digital humanities techniques to evaluate how effectively this widely-used GPT model replicates the genre conventions of the short story.

"Features" of the Short Story

"Brevity and conciseness" speaks to how a short story is written with precise, intentional language.

"Distinctive ending" refers to how a short story has a conclusive ending that carries weight and meaning.

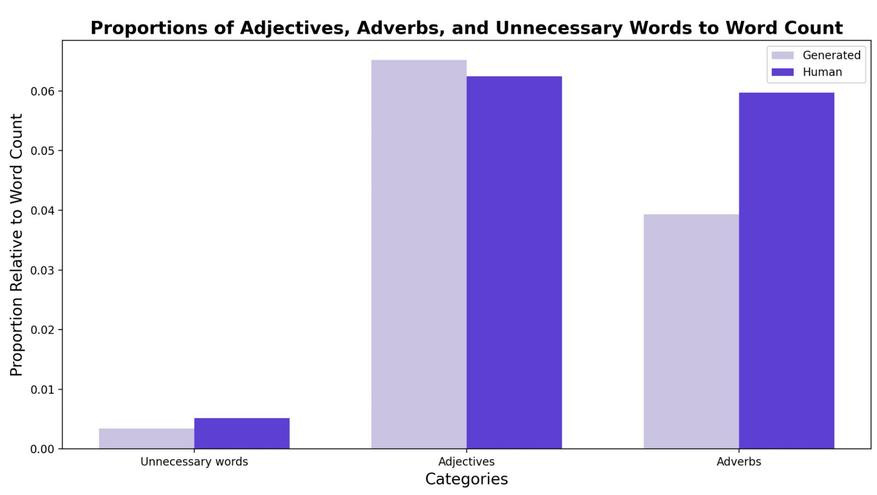
"Unity of impression" means that a short story only includes details pertaining to one central idea or takeaway.

Brevity and Conciseness

There are simple ways to measure the brevity of a short story (i.e., word or character count), though conciseness in language is tricker. To measure the "conciseness" of a short story, we counted the use of parts of speech such as adverbs and adjectives using the spaCy NLP library, as well as words widely regarded as "unnecessary" in creative writing works.

	Min Word Count	Median Word Count	Max Word Count
Human	1,625 words	15,583 words	5,623 words
Generated	1,010 words	1,119 words	1,374 words

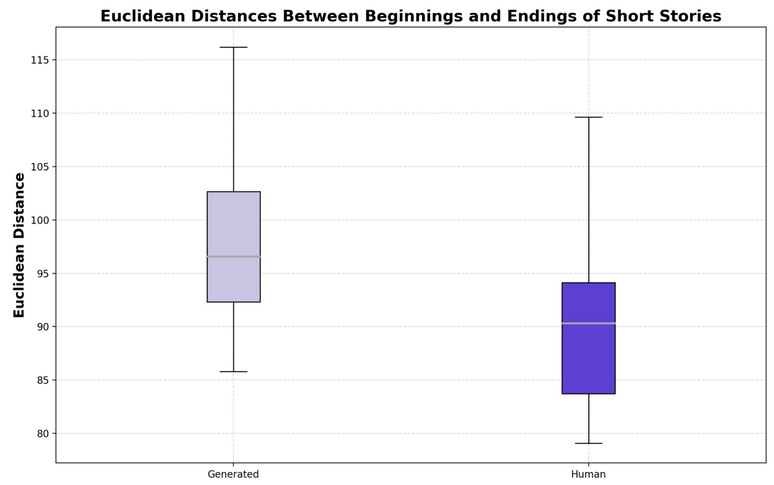
Our findings showed that, despite prompt engineering specifying a desired word count range of 1,000-7,500 words for the generated short stories, the majority of outputs averaged approximately **1,124.48 words.** The 25 human-written short stories, on the other hand, averaged **6,652.64 words**. This means that human-written short stories are on average 5.92 times longer than generated works, while around 6.32 times longer in number of sentences.



Proportionally, however, we see that **fewer "unnecessary" words and adverbs** are used in generated works, while the **use of adjectives is slightly higher**. This is particularly interesting in the case of adverbs—adverbs are often considered expendable in creative writing practices, a principle that ChatGPT seems to follow. Our findings indicate that AI-generated short stories adhere to the "brevity and conciseness" criteria in an exceptionally strict manner.

Distinctive Ending

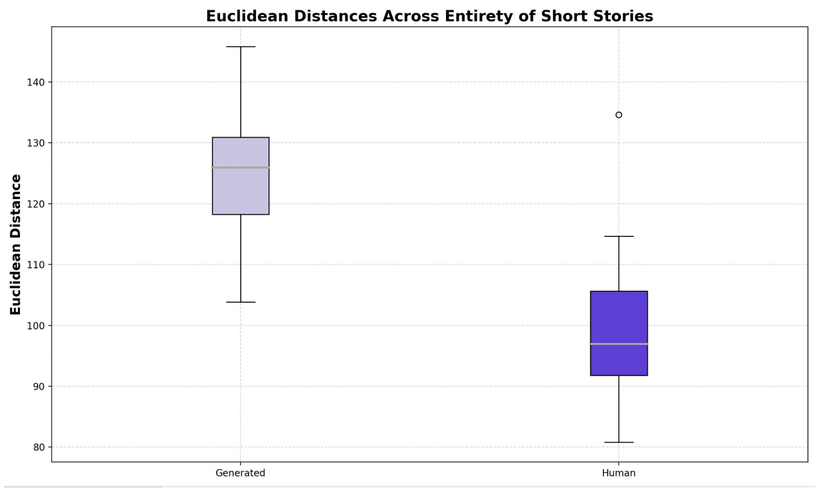
A method for assessing how "distinctive" a short story's ending is was introduced in the paper "Affordances of Mere Length." Following a similar approach to the one described in their study, we divided each story into two sets: the portion up to the last ten sentences and the final ten sentences of the story. By generating word embeddings for each word in these two sets, we can observe how distinct the language is from the beginning and middle of each story compared to the ending by finding the average Euclidean distance between these vectors.



Our findings mark that there is **an average larger Euclidean distance** between the beginning and ending of generated short stories than in human-written short stories. In other words, the endings of short stories that are AI-made are more distinctive than the endings of popular short stories.

Unity of Impression

In a similar approach to the distinctive ending methodology, we can utilize word embeddings to calculate the average Euclidean distance between all words in a short story, assessing whether the vocabulary converges toward a central theme. This allows us to measure the unity of impression.



Our data clearly shows that **generated short stories, on average, exhibited less unity of impression**, as the average Euclidean distance between words in generated short stories was significantly larger than that of human-written short stories.

Implications

These findings provide a small example of how we can assess the presence of genre conventions in LLM-generated texts. Though genres naturally evolve over time, we now confront a new reality where non-human authors contribute to these changes. It is essential to assess the characteristics of these forms and how they are likely to shift in response to these new influences.

Moving forward, our research will apply these algorithms to larger datasets across several Generative AI models, while also incorporating methodology for additional genre conventions found in the short story genre.

Sources

Algee-Hewitt, Mark, et al. "The Affordances of Mere Length: Computational Approaches to Short Story Analysis." *Digital Humanities Quarterly*, vol. 12, no. 1, 2018.

Cook, Claire Kehrwald, and Modern Language Association of America. *Line by Line: How to Edit Your Own Writing*. Houghton Mifflin, 1985.

May, Charles E., editor. *Short Story Theories*. Ohio University Press, 1976.

Matthews, Brander. *The Philosophy of the Short Story*. Longmans, Green, and Co., 1901.

Pasco, Allan H. "On Defining Short Stories." *New Literary History*, vol. 11, no. 2, 1980, pp. 407-423.

Patea, Viorica, editor. *Short Story Theories: A Twenty-First-Century Perspective*. Rodopi, 2012.

Poe, Edgar Allan. "The Philosophy of Composition." *Graham's Magazine*, vol. 20, no. 4, 1846, pp. 163-167.