

# **By AI: Authorship, Literature, and Large Language Models**

by

Imani Ekemini Finkley

A Thesis

Presented to the Department of Literatures in English

for a degree with honors of Bachelor of Arts

Cornell University

April 2024

Supervised by Dr. Jane Juffer

*This thesis and the research behind it would not have been possible without the exceptional support of my supervisor, Jane Juffer. Her enthusiasm to dive into a new frontier — as is the case for research in AI, even more so when approaching it as a humanities scholar — kept my work on track from the moment I suddenly proposed this topic to the final draft. Matthew Wilkens, my second reader who hails from the Information Science department, has also been an important resource as a digital humanist thinking about the intersection of computation and literature. I can confidently say that had I not taken his text mining course, one of these chapters would not have been possible. I am also grateful for the insightful comments from peers, friends, faculty, graduate students, and those in the Humanities Scholar Program. The generosity and expertise of one and all have improved this thesis in innumerable ways.*

*Lastly, I hope this answers questions from those who were wondering why a literature major was taking all of those computer science courses.*

**Table of Contents**

1.1 Introduction	3
1.2 Overview of Generative AI	6
1.3 Benjamin & Mechanical Reproduction of Literature	7
2.1 Smith & Contingent Literary Values	17
2.2 Case Study: Poetry in the Age of Mechanical Reproduction	29
3.1 Foucault & Generative AI Authorship	38
3.2 Politics of the Generative AI Author	44
4. Conclusion	48
Works Cited	50
Appendix	52

## 1.1 Introduction

Literature can be imagined as an intricate web, each thread a narrative or documentation of human thought, history, and discovery. While other art forms have integrated technological innovations within their medium — photo editing technology, for example — the silent tradition of the human literary artist has largely remained. Yet, Aidan Marchine's *Death of an Author: A Novella* shatters that notion. In the mystery novel, Marchine tells of detective Augustus Dupin as he clears his name and finds out who, or what, killed his favorite author. However, Aidan Marchine does not exist, or at the very least, is not human. Novelist and essayist Stephen Marche prompted an artificial intelligence system to write 95% of *Death of an Author* (hence the pseudonym combining the two authors) while he edited and revised the product, making the novel one of the first known published AI-authored literary works (Garner). Understandably, the AI-assisted novel sparks conversations about the broader implications for the literary canon and introduces the possibility for technology to intervene in the literary tradition. Suddenly, literature finds itself on the precipice of a paradigm shift, questioning whether the literary canon can transcend the confines of human authorship to embrace works generated by artificial intelligence. In the shadow of Marchine's novella, the boundaries of literary interpretation seem to blur, inviting a reevaluation of traditional notions of authorship and the enduring legacy of human creativity. In fact, if literature and the literary canon are to survive the integration of artificial intelligence into our everyday routines, perhaps the time has come to clarify our perspective on what constitutes literature.

The literary canon is an evolving collection of works considered culturally and artistically significant, shaping the foundation of a given literary tradition and defining literary values. It includes works of distinct quality, originality, and influence, representing a range of genres, time

periods, and perspectives. Canonical texts often are products of a single author; they tackle themes and provoke critical discussion, ensuring their relevance. However, canon formation has been critiqued for its historical biases and underrepresentation of marginalized voices. In response, there's been a growing effort to diversify canons, incorporating more diverse perspectives and recognizing previously overlooked works (Bean). Hence, the literary canon is an ongoing dialogue among scholars, critics, and readers, reflecting evolving societal values and the ongoing reevaluation of what constitutes literature. That said, should the literary canon now consider works whose authors are not even human?

Once considered mere science fiction, artificially intelligent and machine learning systems are constantly and exponentially evolving to emulate human tasks and behavior. The release of ChatGPT not only cemented generative artificial intelligence systems within the social discourse, but also introduced large language models as a viable tool in our increasingly technology-dependent society. ChatGPT demonstrates the capacity of machines, once prompted, to write, or compose language in a grammatically correct and coherent format. OpenAI admits that GPT-3 can “generate several paragraphs of synthetic content that people find difficult to distinguish from human-written text” (Brown et al. 35). So, since ChatGPT is already producing poems, essays, and speeches, it follows that soon these large language models will regularly return larger literary works like Marchine's novel. I will refer to these works, which are produced with specific, abstract prompting from human users (i.e. ‘write a poem’, ‘write a limerick’, ‘write a short story’, etc.), as generative AI literature. At this moment in time, we as readers and authors, must question where generative AI literature resides within the literary tradition and the expanding definition of the literary canon.

However, though the age of a new machine is upon us, I argue that literature need not fear the presence of generative AI literature. Rather, literary work generated by large language models offer a different perspective on the literary tradition and the function of (human) authors in relation to how literature is consumed and produced. In response to the rise of technological innovation, Walter Benjamin and other contemporaries warn that the rise of this technology will ultimately erase traditional literary values and artistic tradition. Though I acknowledge Benjamin's concerns — especially considering that he was writing when technology recast art as a propaganda tool in Nazi Germany — I insist that innovation ultimately transforms our relationship with artistic tradition, bringing artistic practices closer to their audience. I will use Barbara Hernstein Smith's notion of the contingency of value to understand the history of literary evaluation and generative AI potential to prompt literary values to once again evolve. I will pair this with an empirical analysis of ChatGPT literature to reveal that generative AI literature actually stands closer to human-authored literature than one might assume. Finally, I will refer to Michel Foucault's idea of the author function and Cornell University's regulation of generative AI authorship to reinforce the idea of large language models as authors, not just writing tools. Specifically, I insist that since the author is a societal construct, there is room to embrace generative AI authorship; even university policy unintentionally recognizes the generative AI author, despite claiming that large language models are writing partners at best. In doing so, we as literary audiences can sooner engage with this new form of literature that invites us to revisit the literary tradition and inspires a new understanding of how to continue to document diverse human experiences.

## 1.2 Overview of Generative AI

A large language model (LLM) is a type of machine learning model designed to perform various natural language processing tasks such as generating and classifying text, answering questions in a conversational manner, and translation. LLMs are trained on massive amounts of text data, allowing them to learn how to understand and generate human-like text. This is possible because of neural network architecture, which mimics the human brain to process and analyze information. A model is trained by ‘reading’ an immense number of books, articles, and other written content to identify the nuances of a given language. During training the models learn to predict the next word in a sentence, understand context, and generate larger bodies of text.

Generative artificial intelligence, or generative AI, is a type of artificial intelligence system that enables users to generate new text, images, and other types of media based on a variety of inputs. Released in June 2020, GPT-3, or Generative Pre-trained Transformer 3, is a LLM developed by OpenAI. The key innovation of GPT-3 — popularly known as ChatGPT — and its successors is its pre-training methodology. The model undergoes pre-training on a diverse and extensive dataset that includes a significant portion of Internet-scraped sources. Given a prompt or input, the model can generate coherent and contextually relevant text, completing sentences, answering questions, or even composing entire larger literary works such as essays.

### 1.3 Benjamin & Mechanical Reproduction of Literature

In recent discussions of generative AI literature, an issue has been whether the mechanical production of literature aims to erase the human author from the literary function. In “The Work of Art of Mechanical Reproduction,” Benjamin explores how the technological innovations of photography and film change the tradition of fine art, the artist, and the audience. For Benjamin, “even the most perfect reproduction...is lacking in one element: its presence in time and space, its unique existence at the place where it happens to be” (3). In his essay, Benjamin examines the definition and value of art amidst technological revolution and introduces his notion of aura as the unique and authentic presence of the original work of art. The aura of an art piece encompasses the sense of time, history, and tradition embedded within an artwork, creating a deep connection to its historical and physical origins. His exploration of the concepts of aura and artistic value foreshadows voices today that want to remain with traditional literary works and are hesitant to consider generative-AI authors.

Admittedly, art has always been reproducible. From the style of an artwork to the piece itself, there are numerous ways to mimic art and the artist. That is why there are policies, patents, and laws to regulate the reproduction of a work of art. Benjamin does not deny that “man-made artifacts could always be imitated by men,” but he argues that the introduction of the “mechanical reproduction of a work of art, however, represents something new” (2). After all, for centuries literature has been produced by the mechanical reproduction technology of printing. Before machines, printing was a craft, a precise methodology, to reproduce a single work of literature. And yet, while the introduction of the printing press and eventually the printing technology seen today did render the manual printing industry obsolete, it never reproduced the



author. Large language models, on the other hand, are a mechanical reproduction of the entire literary function, representing something entirely different.

We can further consider the implications of Benjamin's argument by looking at art. Consider the *Mona Lisa* being displayed in the Louvre compared to an image of the *Mona Lisa* in an art history textbook. The former carries the weight of history, while the latter is a mere reproduction. The aura of *Mona Lisa* originates from its presence in Florence, Italy at the height of the Renaissance. The portrait of the mysterious woman in a three-quarter pose is renowned for its novelty during a continental revolution of contemporary art. Conversely, the image of this artwork in a textbook is incapable of accessing that aura. Still, this shift towards mechanical reproduction has both positive and negative implications. On one hand, it leads to greater accessibility, which even Benjamin acknowledges: "technical reproduction can put the copy of the original into situations which would be out of reach for the original itself" (4). By reproducing the painting, it is liberated from the constraints of time and space: budding artists and historians are able to examine the minute details of the work, copies can be hung in one's personal space, etc. This commodification broadens the audience for art and challenges traditional notions of what art is and who it is for. A *Mona Lisa* poster hangs in a classroom, one of millions scattered far away from the original housed in Paris. If not for the reproduction, the students in the classroom would have to rely on description and imagination of one of the world's most iconic paintings.

On the other hand, mechanical reproduction, such as printing or photography, diminishes the aura of an artwork — copies lack the same historical and physical connection to the original. Further, "by making many reproductions it substitutes plurality of copies for a unique existence" (Benjamin 4). In other words, for Benjamin, distancing an artwork from its aura is ultimately

diminishing its value until it is indistinguishable from its reproductions. He explains, “...the reflected image has become separable, transportable. And where is it transported? Before the public.” (Benjamin 11). The reproduction makes the original work a separate entity from the original. Note that, in doing so, the artist no longer has control. Granted, Leonardo da Vinci no longer has a say in how his artwork is transported. However, the color choice, brush strokes, size of the painting — in other words, the details of the work that contribute to its aura and novelty — are lost due to reproduction. With each print and photograph, the artwork is no longer fixed and the aura becomes ever more distant.

Literature is a work of art where language is the medium. Like a sculpture or a painting, writing is a form of creativity and self-expression. Specifically, literature is composed of writing that is valued in the canon for its unequivocal aura. The literary canon, while ever-expanding, aims to establish and reflect the literary tradition. For example, the uniqueness of Harper Lee’s *To Kill A Mockingbird*, is “inseparable from its being embedded in the fabric of tradition” (Benjamin 6). The tradition is the cultural significance of a literary work produced by a human author, dependent on its aura, that elevates it to historical reverence.

Now, consider a poem generated by ChatGPT. Here, the author is being mechanically reproduced. Traditionally, the (human) author drafts, edits, revises, and publishes their poetry before an audience. The literary form is used to display an aspect of the human experience and evoke an emotion or understanding from its readers at a specific point in time. The author’s, or literary artist’s, ability to speak to an audience with their poetry and instill an enduring significance in their work is part of the aura that Benjamin values. However, the generative AI poem is removed from that. LLMs cannot participate in the tradition of creating literature, they merely manufacture a semblance of the process. As a result, due to being a reproduction of an art

form, large language models distance the art of literature from its tradition. Similar to the *Mona Lisa* in the textbook, the authorship function in literature is put into situations out of reach from traditional authorship that has contributed to the artistic canon. For Benjamin, this would suggest that the audience can no longer resonate with the author as a fellow human, diminishing a core aspect of the aura of literary art. In doing so, the value of human authors in literature is questioned. Benjamin explains that “the situations into which the product of mechanical reproduction can be brought may not touch the actual work of art, yet the quality of its presence is always depreciated” (Benjamin 4). Benjamin posits that the existence of mechanical reproduction of art immediately devalues the original art form. In other words, the technological imitation of the author in generating literary works can only serve to disparage human authorship and the literary canon that they have formed. And as audiences and users continue to engage with this manufactured authorship, both the human author and human-authored literature will grow more distant.

The aura relies on distance. The author must stand outside of its audience as a literary genius who bestows their masterpieces on the masses. An artist must observe its subject to capture it — like a painter before their muse. Similarly, the literary artist must distance themselves from their readers, crafting their work in seclusion before presenting it before the public. Benjamin defines the aura as “the unique phenomenon of distance” which mechanical reproduction, both implicitly and explicitly seeks to diminish (5). In a mechanical reproduction, the painter is no longer before their subject with an easel maintaining their boundary. Rather, specifically for large language models, the subject sits alongside the painter to guide their brush to craft something new. Generative AI invites intimacy between the literary artist and the reader,

allowing them to become one. This detracts from the aura, removing the author from its high ground above the public.

The shrunken distance between author and reader due to mechanical reproduction does dismantle the inherent hierarchy required for an artist to establish the aura of their work and contribute to the literary canon. It also raises a question of authenticity once the line between artist and subject blurs. Benjamin explains, “The authenticity of a thing is the essence of all that is transmissible from its beginning, ranging from its substantive duration to its testimony to the history it has experienced” (4). Amidst infinite literary publications, the authenticity of the human author in the literature speaks to an enduring literary tradition that is lost in translation during technical reproduction. This is because the authenticity stems from practicing the ritual of a literary work, beginning from its first draft and extending to the publication; that process deems it canonical, or at least literary.

Generative AI introduces work separate from this methodology. Benjamin defines two concepts of how art is received and valued: cult value and exhibition value. The cult value emerges from the aura of authenticity and novelty of an artwork, “demand[ing] that the work of art remain hidden” (Benjamin 7). The original artform’s cult value requires that it be revered for its historical significance and cultural importance. In other words, the cult value of literature expects a canonical work to have ritualistically followed the literary tradition. The ritual lends itself to maintaining the authenticity of an artwork. Therefore, the cult value of the original literary art form includes the distance between author and reader as well as the writing to publication process.

Clearly, large language models remove the literary function from this ritual. And “with the emancipation of the various art practices from ritual go increasing opportunities for the

exhibition of their products (Benjamin 7). As such, instead of holding cult value, a poem produced by ChatGPT leans into the exhibition value of an artwork. That is, due to mass distribution and wider accessibility, literary art becomes detached from its original context. Instead of being acclaimed for their aura, literary products produced by generative AI are readily consumed as commodities. Exhibition value emphasizes the reproducibility and decreased aura caused by mechanical reproduction. Similar to the establishment of photography into an art form, “by the absolute emphasis on its exhibition value the work of art becomes a creation with entirely new functions, among which the one we are conscious of, the artistic [and writing] function, later may be recognized as incidental” (Benjamin 7). So, as large language models like ChatGPT and their literary products stand before the literary tradition, it invites a new form of the literary function, signaling a new art form that can become substantial.

The roles of the artist and the audience both undergo a transformation in this new paradigm. The aura of art is intrinsically linked to its historical context, and as mechanical reproduction flattens this connection, our relationship with history and the past is altered. A generative AI author’s understanding of the past is limited to its dataset, thus it is not an active participant in ongoing historical events. Readers of generative AI literature cannot look toward the author for its awareness of historical and current nuances. Benjamin’s ideas provide a lens through which we can evaluate the changing landscape of art, whether produced by human or generative-AI authors, as it forces us to reexamine the concepts of authenticity, accessibility, and the evolving nature of artistic value.

The relationship of the artist to their art also undergoes a transformation. The traditional relationship of an artist with their muse is based in a natural reality. For example, a painter sits before a subject. Or a literary artist surrounds themselves with notes and references. However,

generative AI is not an active participant in reality. Benjamin highlights this difference, saying, “The painter maintains in his work a natural distance from reality, the cameraman penetrates deeply into its web” (Benjamin 13). Similarly, the author exists at a natural distance from reality while generative AI aims to penetrate its dataset, find a pattern, and generate a statistically, optimal replica. The model acts on aspects of the literature corpus that it learned from, a breadth of literary archives, inaccessible to a literary artist; it is impossible for an individual to have read all literature. Benjamin goes on to compare a traditional artist and one who practices mechanical reproduction:

That of the painter is a total one, that of the cameraman consists of multiple fragments which are assembled under a new law. Thus, for contemporary man the representation of reality by the film is incomparably more significant than that of the painter, since it offers precisely because of the thoroughgoing permeation of reality with mechanical equipment, an aspect of reality which is free of all equipment. And that is what one is entitled to ask from a work of art. (Benjamin 14)

Again, Benjamin highlights the importance of distance between an artist and their work. The painter is an observer of its muse from afar whereas the cameraman attempts to intervene in the present moment for its artwork. Since large language models need not draft, edit, and finalize their artwork, they invite users to become instant literary artists. Their products have been fine-tuned and meticulously calculated for immediate publication. While the literature of a human author is that of (usually) one creative, that of generative AI consists of a large, potentially infinite, corpus from which its algorithm — the new law — assembles its artwork. Generative-AI literature offers, through a permeation of a dataset that represents reality, an aspect of reality which is free of all subjective, authorial inspection. This streamline production

is free of intervention of critics, publishers, editors, or even an author's doubts — in other words, it is everything that the contemporary audience cannot demand from a traditional work of literature. This simplified, more convenient form of literature hints at the possibility for traditional literature to experience a decrease in social significance for the sake of the ease of generative AI. That is, the authorial status is diminished and reduced into an efficient, commercial mechanism that manufactures optimal literary products. The author as an exalted genius and renown artist is at risk due to mechanical reproduction, especially since the quality of generative-AI literature is constantly improving as it receives feedback from its reinforcement learning framework and its dataset increases.

It is not surprising, thus, that the possibility of generative-AI literature has been met with resistance by those who practice the literary tradition. Even though technology does increase accessibility and invite new art forms, Benjamin warns that we may lose a clear sense of what is original and authentic in a world inundated with copies. For example, the Authors Guild calls for protection of human authorship. As America's oldest and largest professional organization for writer, the Guild acknowledges that:

These generative AI technologies, which can generate new content because they incorporated millions of existing literary and artistic works into the fabric of their software, will have a significant impact on the future of the literary arts and the writing profession, and it is imperative that we approach their development and use with respect for human creators and copyright. (Authors Guild)

In short, the Authors Guild warns that as generative AI grows into everyday use, we do not overlook the authentic artistry of literature and literary creatives. After all, “mechanical reproduction of art changes the masses towards art” (Benjamin 14). Large language models

allow for literary production that circumvents the rituals of the original art form, undermining how the literary tradition is valued in society.

Though perhaps generative AI is inevitable. Benjamin concedes that “the desire of contemporary masses to bring things ‘closer’ spatially and humanly...is just as ardent as their bent toward overcoming the uniqueness of every reality by accepting its reproduction” (5). He is pointing out the drive for convenience and ease that is achieved through mechanical reproduction. When printing by hand was arduous, the printing press reproduced writing to bring literature closer to home and the masses. When self-portraiture was too expensive, photography allowed for anyone to get closer to themselves and a representation of themselves. Generative AI brings the literary function much closer — ChatGPT is a commercial product that can be accessed on OpenAI’s website and mobile app. Large language models, acting as literary authors, attempt to overcome the novelty of the singular author whose genius is drafted, edited, revised, and published. Instead, a model can simply be prompted to create thousands of literary works almost instantly, flooding the market with work that is in much closer proximity. Whether or not the quality of the reproduction is equivalent is negligible in the goal to be close to the literary tradition. For, “to pry an object from its shell, to destroy its aura, is the mark of a perception whose ‘sense of the universality of things’ has increased to such a degree that it extracts it even from a unique object by means of reproduction” (Benjamin 5). Hence, the mechanical reproduction of the literary author essentially satisfies bringing the literary function and values closer to contemporary masses at the cost of the aura of the literary arts. Just as Benjamin feared photography and film for the future of art, such blind progress towards mechanical reproduction reveals the potential fragility of our hold on literary values and, apparently, the human author.



Still, there is hope for the human author whose works now stand alongside generative AI works like Marchine's *Death of an Author*. Yes, literature is being mechanically reproduced, altering the relationship between literary art, artist and audience. However, the human author is not necessarily condemned to oblivion because of this transformation, despite what some literary authority figures and critics seem to fear. Rather, we must understand what literary audiences expect from and value in literary art. That is, how do we evaluate literary products of large language models, mechanical reproductions of the literary function, given the historical values of the literary canon?

## 2.1 Smith & Contingent Literary Values

Currently, generative AI literature faces modern values and evaluations of literature. These values, admittedly, have a history of change: the most notable change resulted in the canon encompassing a broader, multicultural collection of literary work. But rather than simply representing a diverse array of authors and perspectives, generative AI is challenging the canon with a new form of literary production altogether. Large language models challenge the literary tradition's evaluation of the author. Furthermore, it calls for a reflection on how the literary tradition has been and continues to be evaluated. In "Contingencies of Value," Barbara Herrnstein Smith asserts that the values shaped by authors writing in the literary tradition have been influenced by the readers of that time and place. Recall that Benjamin explains that the aura of a literary work, defined by the context of its publication, is determined by the temporal context of the audience of its original readers who evaluate and exalt the work as a part of the literary canon. However, while Benjamin clings to the nostalgia of a literary work published by a single genius author in his notion of the literature's value, for Smith, value does not have a timeless essence. Smith believes that any apparently objective truth that evaluates a work as literature and canonical text can be challenged by a contrary truth at some other place and time. Thus, values and the judgments it instructs are relative in the "sense of the contingent rather than a subjective" (Welch). As such, in her book, Smith argues that literary value represented in the literary canon emerges from a dynamic process of multiple variables.

Regarding the aura of a literary work as a product of its historical context without recognizing the variable context of evaluation restricts literary authors and audiences from recognizing the evolutionary aspect of literary value. Smith agrees as she urges literary audiences to realize that "literary evaluation is not merely an aspect of formal academic criticism but a

complex set of social and cultural activities central to the very nature of literature.” Furthermore, obscuring that only serves to “forestall the exhibition and obviate the possible acknowledgement of divergent systems of value and thus to ratify, by default, established evaluative authority” (Smith 7). Those who rush to exclude generative AI literature from the literary tradition overlook the malleability of literary evaluation. Moreover, they isolate the literary artistic form from introspective discourse provoked by the introduction of a new author and way of producing literature.

In the debates between generative AI and the literary canon, one established evaluative authority can be found in the protests of the Authors Guild. The Guild appears adamantly against any notion of its traditional evaluation evolving, releasing an official statement warning that “the unregulated development and use of generative AI technologies will lower the quality of books journalism, and public discourse fundamental to democratic culture.” The Guild is emphasizing that the literature will be devalued should generative AI authors’ influence grow and enter the literary market. To this, Smith would retort, “all value is radically contingent, being neither an inherent property of objects nor an arbitrary projection of subjects but, rather, the product of the dynamics of an economic system” (12). Hence, rather than fearing the devaluation of the literary function, it is important to reflect on the history of literary evaluation to understand how these values will proceed in conversation with generative AI literature.

Evaluating literature involves comparing a text against a set of values. When approaching a literary work, especially to ascertain how it compares to the preexisting literary canon, “value judgments are context-dependent and shaped by the relation of the speaker to [their] audience and by the structure of interests that sustains the verbal transaction between them” (Smith 22). Smith highlights the subjective aspect in evaluating literature and the literary canon, since

readers and the establishing authority are appealing both to readers' values and to the "objective" values that define the literary canon at the time of evaluation. Involving personal biases highlights what Smith calls the contingency of value as a product of both individual and institutional factors.

Likewise, large language models, especially public generative AIs like ChatGPT, also rely on a series of human evaluation and personal sentiments to produce literary works. Large language models, given a corpus of approximately all available literature, identifies patterns and features to infer the values of literature. These values do not involve personal sentiments on the literary experience of a text, but rather common arrangements of words used in the dataset that have been accepted in literature. So when prompted to act as a literary artist, the model returns the statistically optimal combination of words based on what it has learned from the dataset. ChatGPT goes a step further, using human feedback to revise and optimize its generations, also known as reinforcement learning. Through words of affirmation or critique, the model is better able to predict what literary product has a higher probability of being accepted as literature. Given its mathematical representation of our values of the literary canon and a record of the feedback, ChatGPT ultimately aims to be rewarded with a thumbs up or flattering review, reinforcing its method of generating literary art. Conceptually, this is no different than an author revising their work after a series of test readers and referencing similar work that has already been accepted as literature in their genre. It is reasonable to say that an accolade or place on a bestseller list reinforces their literary style and informs how they will continue to produce work going forward. Generative AI is just able to do this at a large scale.

While some may refute this comparison due to the model's inference of literary value being reduced to probabilities, Smith would remind them that the "traditional discourse of

value...reflects an arbitrary, attesting, segmentation, and hypothesization of the continuous process of our interactions with our environments” (12). Basically, there is no inherent way to define or represent literary value since it is so subjective, so it is unreasonable to dismiss statistical representations solely because it does not readily lend itself to personal sentiments. This is especially clear when you consider that the number of sales and positive reviews are statistics that a human author might consider as they craft their literary art. As a mechanical reproduction of the author, generative AI is “all the while testing the local and global effectiveness of each decision by impersonating in advance [its] various presumptive audiences, who thereby themselves participate in shaping the work they will later read” (Smith 24). Similar to what Smith describes, the generative AI author produces work that is shaped by what its audience reads, or, in other words, impersonating existing authors of literature found in its dataset. This pandering toward the audience’s values is done by both human and generative AI authors alike to guide a literary product.

Smith elaborates on the economic features that determine a literary evaluation. She explains, “all these forms of evaluation...represent a set of individual economic decisions, an adjudication among competing claims for limited resources of time, space, energy, attention—or, of course, money—and also, insofar as the evaluation is a socially responsive act or part of a social transaction” (Smith 25). Basically, each evaluative act of literature represents a set of economic decisions for both the author and audience. This is reasonable, as a literary artist who intends to publish their work essentially is participating in a transaction with its audience: their profit and accolades for the readers’ literary experience. Similarly, large language models and generative AI operate based on cost efficiency. While perhaps not optimizing their generated works to what will sell well, each literary work generated by a large language model is returning

a product that has a reasonable similarity to literature observed in its dataset and has the highest probability, based on a history of human feedback loops, of being affirmed as acceptable literature. A model has limited time (a user expects an immediate publication upon request), attention (if the user is not satisfied it will leave), and energy (the computational costs involved in operating a model at a large scale), so each evaluative act, each decision of how to generate a text, is done so to minimize the cost of that action while maximizing the positive response of the user.

In addition to economic factors and reader responses, previous literature within the literary canon is also taken into account when an author attempts to introduce a new literary work. Smith points to the “activities of the academy” as figuring “significantly in the production of literary value” (25). Smith refers to the academy specifically as an institution that formally shapes the creation, evaluation, and history of the literary canon in both the classroom and anthologies. It follows that “...the repeated inclusion of a particular work...goes some distance toward creating its value...[it will also] have the effect of drawing the work into the orbit of attention of a population of potential readers...they make it more likely both that the work will be experienced at all and also that it will be experienced as valuable” (Smith 25). Unsurprisingly, the value of a work of literature is also drawn from the prevalence of the work within an individual reader’s familiarity with the literary canon. This concept of repeated exposure correlating to value within a canon can also be applied to other art forms. For a painting, the strength of its aura endures due to fine art consumers’ constant return to the work as an exemplar of art. This is also similar to when a movie is constantly discussed in any given film course across media departments in various institutions. This relationship between prevalence and value translates to generative AI, specifically when looking at its corpus. Consider a large language

model tasked with producing a work that it deems is viable literature. Suppose that its literary corpus was a combination of multiple, smaller datasets of work in the literary canon. If all datasets included a significant number of references to George Orwell's *1984*, then the model will infer Orwell's literature as a primary reference of literary value and refer to the novel more often when evaluating its literary generations. So, like a novice literary author, the model will refer to what it has experienced in its corpus (its form of instruction and guidance about the values of the literary canon) as more valuable examples of literature that pass evaluation.

Smith takes these variables — economics, predecessors, and context — to imagine the literary function as a cycle between the artist's motivations, resources, active feedback during generating, and anticipated audience interests. Through the integration of the human feedback loop into its architecture and the dependency on its training dataset to understand literary values and features, generative AI seems to reproduce this function to an extent. Large language models do not necessarily anticipate the shift of the audience. That shift can be inferred through means outside the scope of the dataset it is trained on. For example, current social trends or events, if not already examined and reinforced, are highly unlikely to appear in a generative AI literary work. Unlike a human author, which has the resources and agency to know the consciousness and imagination of its audience since they are a consumer of the literature as well, a large language is not privy to that context; it cannot fully grasp the potential aura of its generation within social discourse.

Regardless, the factors that define the aura of a literary work—the repeated interaction with a canonical text and the response of the audience to a literary product— are not fixed. Smith states that a canonical text is the best of its kind at a given time for some community of subjects. Smith describes these texts as having fulfilled the function that the audience at the time was

searching for. However, if that function is no longer desired by current readers it risks becoming obsolete. This is not a definite sentence; Smith introduces two possibilities for a literary work that no longer meets its original function: oblivion or reproduction.

It is important to note that evaluation is critical in the longevity of a piece of literature. “Whether overt or covert, verbal or inarticulate, and whether performed by the common reader, professional reviewer, big-time bookseller, or small-town librarian” for literature to remain canonical readers must return to evaluate it, to cite it, to discuss it...to read it. Otherwise, “its visibility as well as interest will fade, and it will survive, if at all, simply as a physical relic” (Smith 27). Basically, if the values of a literary text diverge from modern forms of literary evaluation it becomes a mere historical artifact, essentially oblivion for a work of literature. The work no longer performs for the readers so it can only be remembered, if at all, as evidence of literary scholarship at a certain time and place (e.g. ‘Here lies a book from the mid-twentieth century written by a man named George Orwell’ or ‘Here is a poem written by OpenAI’s GPT-3.5’). There is no longer appreciation for its originality since audiences have unconsciously or explicitly avoided it due to the values its aura represents. However, there is still a chance for redemption, explains Smith, and a return as an “unjustly neglected masterpiece” if those values and functions again become desirable within the rubric of literary evaluation (27).

Alternatively, after losing its original function, a literary text, rather than fall to oblivion and only hope for redemption, instead can go on to be culturally reproduced even if the original environment where its properties were optimal no longer exists. As long as it “continues to perform *some* desired/able functions particularly well, even if not the same ones for which it was initially valued,” it will continue to be evaluated and read. Smith goes on to say that this subset of desirable functions safeguards the literature within the canon from oblivion. As a result,



“when the value of a work is seen as unquestionable, those of its features that would, in a noncanonical work, be found alienating...will be glossed over” (Smith 28). Smith is saying that even archaic sentiments and values will not condemn a work to oblivion so long as it is canon. That is, so long as it was repeatedly introduced as an exalted canonical text while it was performing all desired functions exceptionally well, the more its value is cemented and evades the path of oblivion. Not only that, but “it will also begin to perform certain characteristic cultural functions by virtue of the very fact that it *has* endured...and will be valued and preserved accordingly” (Smith 28). Here, Smith is asserting that the reputation of canonical literature for its high value coupled with admiration for the enduring nature of said value — which is not necessarily the same as the current values that determine the evaluations of the literary canon — can allow works that would have otherwise become obsolete to remain canonical.

The paths of oblivion and reproduction that Smith has presented for literature reveal a critical difference between the evaluation of the literary product as it is published and presented to the public. Even though both generative AI and human authors undergo similar evaluation to inform the generation of their literature, how that literature is then evaluated to withstand evolving literary values is different. Traditional, human-authored literature, regardless of the contingent values that evaluate it, retains a sense of permanence defined by the moment it is published before a specific audience at a specific time in a specific place. This contributes to the inherent novelty nature that Benjamin insists is irreplaceable and endangered by mechanical reproduction. Once presented to the public, a literary work is frozen as it is and subject to further scrutiny by audiences (ideally) for generations to come. But large language model frameworks cannot replicate the same sense of permanence with their generated texts. Unless taken by a third

party to participate in the publishing process, all publication of generative AI literature occurs intimately and instantaneously with the user who requests it. There is no review board or bestseller list, only the feedback of a single user of unknown personal sentiments who will evaluate the work produced by the model. The literary experience for both author and reader has been reproduced and reduced such that the latter engages and must evaluate the literature for just a moment. Unless the work is stored externally, once the reader returns to the author for that literature (i.e. gives the same prompt to the model) the literature likely has already changed to something else entirely, because the author evolves constantly to generate work that it is certain will be accepted as literature in that instance. If one were to prompt ChatGPT to “write a poem,” the literary work exists in that space at that moment. For, if the model is asked “What about that poem you wrote yesterday?,” it will return: “I’m sorry for any confusion, but I don’t have access to specific poems or texts I may have generated for you [the user] in the past” (ChatGPT, 7 Apr 2024). In other words, commercial generative AI models, as they are now, produce literature that is not meant to be retrieved for later use. Hence, the fall to oblivion for a generative AI literary text means that it only failed to fulfill the values of the user of that instance rather than a series of evaluations by a broad array of the literary consumers. It can follow the path to reproduction, yes, insofar as its user(s) choose to desire work written by generative AI — a function that it can never fail. This is all to say that generative AI is not inherently susceptible to the generations of evaluation post-publishing as its framework stands given the mode of its publications.

However, this does not mean that generative AI literature cannot be published. After all, *Death of an Author* demonstrates that it is possible for a large language model’s literary work to participate in traditional publication. So I want to emphasize that the difference lies in the inherent purpose of literary production between the two types of authors. Yes, a generative AI

author or generative AI co-author like Aidan Marchine can present work to an audience outside of the model's single user framework with external (human) assistance. But we cannot overlook that, originally, the model was prompted *by* Stephen Marche and wrote the novel *for* Stephen Marche. This is especially true when we consider that the LLM generations were edited and curated by Marche's artistic style and expectations of a mystery novel. *Death of an Author* is a direct product of Marche's literary values published to Marche's satisfaction. Marche then took the novel and presented it to the public. Meanwhile, a traditional human author may dedicate their work to a single individual or even write for a very specific group, but their writing process cannot act on all of the literary values of each member of their audience. Hence, a human author publishes based on a generalized idea of their readers' literary values, whereas a large language model publishes to satisfy every literary value of the reader who requested its work.

In conclusion, human-authored and generative AI authored literature are shaped by the precarious values of the literary tradition. Both are products of a "complex evaluative feedback loop that embraces not only the ever-shifting economy of the artist's own interests and resources...but also all the shift economies of [their or its] assumed and imagined audiences" and thus undergo evaluation during generation and once published (Smith 24). A mechanical reproduction of this evaluation loop, the architecture of a large language model, similar to the human-author function it is translating from, requires constant human feedback to understand what its audience expects and values in a literary generation. Smith mentions that "the value of a literary work is continuously produced and reproduced by the very acts of implicit and explicit evaluation that are frequently invoked as "reflecting its value and therefore as being evidence of it" (30). These evaluations, shaped by a conglomerate of literary authority, personal sentiments, economics, and questionably objective judgements are crucial in constructing a canonical text

and saving the text as literature rather than a deplorable historical relic. Evidently, generative AI literature's fall to oblivion is not of equal consequence as a human-authored text because the model does not publish for permanence but rather for momentary satisfaction. Rather, to become a 'relic' only reinforces the model with more information about the values of its user so that it can return with a more optimal literary text. Still, the use of values and evaluation is the same when writing/generating literature for both authors, meaning that large language models are, like a human author, producing work that represents its interpretation of the audience's literary values. Hence, both human and generative AI authored literature attempt to translate and reproduce literature values. So, then, when faced with both texts, the question arises for readers and authors alike: "Are we interested in the messengers — the chatbots and the Ouija-board revents — or in the messages they deliver? Those messages, after all, are about us: our fate, our origin, our fragile human essence" (Scott). A. O. Scott, critic for the *The New York Times Book Review*, raises a crucial question about what society will accept being manufactured within the literary function. This is an important to consider since Benjamin has already warned that becoming complacent to the mechanical reproduction of literary art is submitting to the devaluation of the art form.

Of course, there are skeptics at the potential for large language models and mechanical reproduction to influence the value of literary craft as Benjamin and Smith would imply. After all, "mass production has always coexisted with, and enhanced the value of, older forms of craft," besides "the standardization of mediocrity does not necessarily lead to the death of excellence" (Scott). This excellence, presumably, is defined by works that enter the literary canon and continue to shape and maintain the values of the art form's criterion. Yet, considering that the literary canon's values are reflections of the social values — values that inevitably

continue to change throughout history — it is not unreasonable that generative AI could change the very standards of mediocrity and excellence of the art it is producing. As Smith has made clear, the values of the literary canon are not timeless. The values that deemed Shakespeare's *Hamlet* to join the literary canon are not the same that evaluated Ellison's *Invisible Man*. That is, the evolution of the values of the literary canon highlight their dependence on the continuous cycle of evaluation by readers as well as authors.

## 2.2 Case Study: Poetry in the Age of Mechanical Reproduction

Smith makes it clear that literary values have historically evolved with their audiences and that authors continue to write to reflect the literary expectations of their readers. Mechanical reproductions of the literary function maintain the cyclical relationship between artist, audience and artistic value, though the author-reader relation is transformed from a literary genius speaking to an audience to an intimate correspondence: the reader makes a request to a large language model which immediately delivers a statistically optimal literary product. However, rather than referencing personal impressions of literary values or referencing literary authorities, large language models infer literary values from patterns identified in a dataset. That is, the values that a generative AI reflects in its literary work are optimized statistical decisions and not necessarily attempts to evoke a literary experience or emotion from the reader. This raises the question whether literary values can be reproduced as probabilistic hypotheses or large language models produce new values that prompt a new type of literary evaluation.

Large language models have yet to perfect their capabilities to generate larger novels and plays. Marche had to guide multiple generative AI models including ChatGPT in a collaborative effort to produce a murder mystery novel. Thus, it is evident that large language models still struggle to generate extensive literary products without external revisions. On the other hand, generative AI like ChatGPT can readily produce shorter bodies of literary work. Specifically, these large language models can be tasked to act as poets, producing a set of stanzas to convey a theme, sentiment, or event. As the popularity and capabilities of large language models like ChatGPT continue to grow, we must also consider their relationship with the works that they have trained on. It is imperative to analyze how work produced by generative-AI is able to contribute to the very literary values it is trained on and prompted to replicate. This section will

focus on the defining characteristics of poetry and how ChatGPT attempts to participate in the genre's canon and emulate poetic values. Given the inherent creative and self-expressive nature of the literary genre, examining poetry produced by large language models will serve as an interesting guide into the relationship between authorship and literary work. I will compare ChatGPT-generated poems to human-authored poems to determine whether generative AI produces comparable poetry or if, as an author, it produces something else entirely.<sup>1</sup>

First, we must establish what are the “foundational values” that make up the poetic tradition. Poetry as a genre is known for its versatility: there are approximately 200 poetic styles for an author to choose from. The poetic tradition itself differs across cultures, explains Albert Gapi, but for the American poet in particular, “the determining issue remains essentially the same: whether language gives form and expression to an objective truth or in its formal expression creates its own meaning” (272). Whereas a poet attempting to follow the Greek poetic tradition, for example, is “required to have in it the pulse of life; it would be repugnant to his audience if they perceived the artificial outcome of mechanical formulas, a style which sought to impress or surprise by mere tricks of phrase, having no vital relation to his thought” (Jebb). Still, there is a common expectation that informs the sentiments of evaluation toward the poetic tradition: to use language to express a personal truth. Smith herself asserts that “pre-classification is itself a form of pre-evaluation, for the labels or category under which we encounter objects...foreground certain of their possible functions but also operate as signs— in effect, as culturally certified endorsements — of their more or less effective performance of those functions” (23). Basically, by labeling a literary work as a poem, readers approach the work expecting to meet certain values and poetic practices; the genre itself demands a specific literary

---

<sup>1</sup> Complete computational methods and datasets used can be found at <https://github.com/imanif/ByAI>

evaluation. As such, this section adopts a poetic evaluation of theme, diction, rhythm, and agency to analyze ChatGPT's attempts to participate in the poetic tradition.

I created a dataset with which I examine the values of the poetic genre in ChatGPT literary work. I generated 50 poems with ChatGPT using my personal OpenAI account using the prompt: 'Write a poem.' Note that to mitigate the model from interpreting the re-generation requests as negative feedback of its poetry, I reloaded the site for each request. This corpus will be referred to as 'ChatGPT Poems.' Meanwhile, the dataset of 50 human-authored poems is from the Poetry Foundation Poems Dataset (Poetry Foundation). From the dataset of roughly 14,000 poems, whose authors ranged from William Shakespeare to Ezra Pound, 50 poems were randomly selected. Poems in both datasets are of similar length, ranging from 66-141 words. Duplicate poems and authors were also removed to ensure a diverse set of poetic styles. This filtered, random subset of the Poetry Foundation Poems Dataset will be referred to as the 'PF Dataset.'

In poetry, the theme is the lesson or statement — often on human nature — that the author attempts to express. Diction, or word choice, is one way to infer the theme of a poem. Hence when comparing the two datasets, the frequently used, meaningful words (i.e. not commonly used words like the, was, it, etc.) used by the respective authors reflect the themes of the poetry. Admittedly, this is not a definitive method to identify the theme of a poem, much less from a poetry corpus. Still, this analysis immediately highlights a notable difference in the “creative liberties” in diction and theme utilized by ChatGPT versus a set of human authors.

Unsurprisingly, the ten most-frequently-used words in the PF Dataset do not demonstrate a definitive theme. Since the corpus is a collection of poetry written by a diverse set of authors, it follows that the aura, purpose, and theme of each poem is distinct. Moreover, when you consider



that the poems hold the values of the time when they were published, it would be odd if they shared a common theme and diction choice. On the other hand, as the author of the ChatGPT Poems, the large language model is writing to a single user (me) by making mathematically optimal decisions regarding its word choice and thematic elements. Plus, it is not surprising that an author tasked to write fifty poems would return to the same diction and ideas in their literary work; it can be considered evidence of their artistic style. The top words from the ChatGPT poems include symphony, twilight, and time, hinting at themes of desire, nature, and beauty (Appendix A, Figure 2). In short, by examining the frequently-used diction, there is already evidence of generative AI defaulting to certain conventional, literary values surrounding poetic themes and diction.

Rhyme is another important value in poetry. To analyze the rhythm employed in the two datasets, this study focuses on couplets, two lines in a verse that are often joined by a rhyme. Specifically, this analysis looks at whether the last word of each line in the couplet rhymed. At a glance, poems from the PF Dataset use a diverse number of rhyming couplets in their literary works, while 94% of ChatGPT poetry contained at least 6 rhyming couplets (Appendix A, Figure 3). Again, we observe how the mechanical reproduction of poetry restricts the contributions to the tradition to a certain poetic style. I am not arguing that the overwhelming use of rhyme in its poetry discredits the model's product as literature. Rather this dependency on rhythmic elements in ChatGPT poetry supports Benjamin's claim that "mechanical reproduction of art changes the reaction of the masses toward art" (14). That is, readers of the ChatGPT Poems constantly observe this use of rhyme, implicitly correlating the use of rhyming couplets to the poetic tradition. This then limits the variety in rhyme expected when literary audiences review poetic products of large language models, thus changing their evaluation and reaction toward the

tradition. Or, if not the entire genre, perhaps the corpus represents a different set of literary values that can be considered a facet of the ChatGPT poetic tradition. Either way, there is a separation between the poetic values practiced in the datasets, inviting new expectations and values of the poetic form.

The subject of a poem is crucial and often a means for the author to assert agency within their work. After all, the function of poetry is often considered a means to communicate an individual's emotions and experiences. For example, consider “The Epic of Gilgamesh,” the world's oldest poem, that tells of the story of a hero, the King of Uruk. Written in 12 tablets, the introductory stanza opens, “He who has seen everything, I will make known to the lands. / I will teach about him who experienced all things” (Sîn). Already, readers are introduced to the narrator of the poem (the author) who will narrate and share the tale of “the hero, born of Uruk, the goring wild bull” (Sîn). So, when prompted to assume a poet, who/what is the subject of ChatGPT’s poetry? Does it act as a narrator? Does it illustrate “personal” experiences? Who are the actors in its literature? To answer this, I examine the use of pronouns within the poetry datasets.

To identify the use of agency within a poem, I counted the instances of pronouns: she, her, hers, he, him, his, they, them, theirs, I, my, mine, we, our, and ours. In doing this, the authorial voice can be represented by the types of subjects called in generative AI and human-authored poetry. The subject of the literary artwork is another value that is experiencing mechanical reproduction through large language models. Recall that Benjamin mentions how the distance between the art and artist is altered through mechanical reproduction. Examining the subjects of generative AI poetry then serves as a test to see if the distance between the artist and their voice in their work is notably transformed. Of the poems generated by ChatGPT, 74%

utilized word choice that assert the model's voice and place within its literature. In the final stanza of Poem 11 (Appendix B), ChatGPT writes:

So let us, too, with voices clear,  
 Compose a verse for all to hear.  
 In every breath, in every rhyme,  
 We find our place in the grand design. (lines 32-36)

This poem is composed of diction that conveys themes of time, hope and desire. In this stanza, ChatGPT calls to readers to join it, with a voice outside of its functionalities, to find a space within the “grand design” of earthly beauty and universal majesty. Basically, in this poem, ChatGPT manufactures work that directly interacts with its audience and declares the model's ownership of its work. This is not surprising given the model's inherent role to simulate dialogue with users and constantly engage and adhere to its audience.

Conversely, the poems in the PF Dataset used an overwhelming amount of agency: 100% used at least one of the personal pronouns to assert an authorial presence in their literature. To be clear, For example, in her poem, “Washee/Was She,” Sahar Muradi's voice is almost overpowering; she strongly defines herself in each narrative verse (Appendix C). Muradi invites readers into her past and reflects on her memories, when she writes, “in my used country I felt his teeth” (line 11). Indeed, she fully exercises the poetic function to narrate her personal experiences, of which ChatGPT has none.

Clearly, values of self-expression and authorial voice do not translate in ChatGPT's mechanical reproduction of the poet. Given the inherently self-expressive qualities attributed to the poetic genre, my analysis highlights the use of pronouns to reflect how the respective authors assert authorial agency and voice within their literature. Evidently, the PF poems contained

significantly more words of agency: the PF Dataset included 711 pronouns while the ChatGPT dataset only contained 74. Also, when using diction to convey a sense of voice and place within its poetry, ChatGPT often used words that referred to the collective; ‘we’, ‘their’, and ‘ours’ were the three most used pronouns. ‘I’ was only used five times. Some might conclude that this implies that ChatGPT is a sentient machine with a feeble concept of its user and itself.

Conceptually, when considering the use of agency in its poetry and reflecting on the architecture of large language models, it is understandable that ChatGPT is more accustomed to producing literary work that refers directly to the collective (the user) rather than itself. However, given the overwhelming use of pronouns in the poetic canon, from which the model is presumably, in part, trained on and what the PF Dataset represents, it is surprising that the poetic values surrounding authorial voice and agency were not prioritized and practiced by ChatGPT. Despite learning from poetry like the PF Dataset, ChatGPT repeatedly did not perform this literary value in its mechanical reproduction of poetry.

After analyzing theme, diction, and rhyme used in poems from ChatGPT and the Poetry Foundation, this case study reveals that despite adopting the same cyclical author and audience relationship as described by Smith, literary values are not equally translated in mechanical reproductions of the literary function. For each poetic device, there were significant differences between the ChatGPT and Poetry Foundation poems: ChatGPT poems had a collective theme while PF did not; there was no overlap in frequent diction; 94% of ChatGPT poems contained at least 6 rhyming couplets compared to only two percent of PF poems; ChatGPT seldom used words of agency in its poetry, but when it did most were collective pronouns while all PF Poems all used agency pronouns in its poetry — ‘I’, for example, was used 286 times in the PF dataset. In other words, the results showed that generative AI has yet to seamlessly enter the poetry

canon through current poetic values represented in the PF dataset. Rather, the ChatGPT poems embody distinct characteristics that could be evident of a distinct poetic style that would change readers' expectations of the poetic tradition.

Admittedly, my original hypothesis was that the two authors (ChatGPT and human) will follow the same quantifiable trends that define poetry. After all, ChatGPT is a mechanical reproduction of the poetic art form, a translation of the poetic tradition into a statistical method. As a translation, which Benjamin calls a form, "the law of this form has its first place in the original" ("Task" 205). Intuition then demands that, at the very least, the large language model will echo the poetic form in its translation. Yet, the two authors practice very distinct forms of poetic generation. The form of the ChatGPT poems does not have its place in the original according to rhythm, diction, theme, or agency devices of the genre. From this simplified poetic analysis, ChatGPT has already failed as a translation to reflect any kinship between the poetic authorships. This is important as it shows that generative-AI literature, specifically poetry generated by ChatGPT, while not being able to easily stand amongst works of the poetic canon due to its significant differences, introduces an alternate poetic tradition. This form, out of place with its original, challenges the bounds of the current poetic canon. The attempt to reproduce and emulate poetic values generates a work that resides adjacent to it, so we as literary consumers and critics can either widen our values to accept it or exclude it entirely. Confronted with our values of literature, there is a sense of alienation from our certainty of the canon and our determination to assert it amidst generative AI literature. This is characteristic of a translation, as Paul de Man reminds us:

"We think we are at ease in our own language, we feel a coziness, a familiarity, a shelter in the language we call our own, in which we think that we are not alienated. What the

translation reveals is that this alienation is at its strongest in our relation to our own original language, that the original language within which we are engaged is disarticulated in a way which imposes upon us a particular alienation, a particular suffering.” (de Man 25)

de Man points out that as a translation of the literary and author function into the computational space, generative AI literature is highlighting the precarity of our values of the literary canon and what can be welcomed as literature. This particular alienation, this particular suffering, stems from the realization that evaluations of literature are ultimately contingent.

As large language models are constantly evolving, Benjamin would also say that, like photography and its relationship to painting, the mechanical reproduction of an art form (literature) will diminish the historical value and novelty of the art before eventually asserting itself as a new art. So as generative AI grows to learn and adapt to the values we have instilled in the poetic tradition, it is imperative that we assert whether our values that have defined poetry and therefore literature need to adapt with it. The question remains as to whether an evolution of our values instilled within the literary canon will embrace generative AI literature or exile its works as a new art form entirely. Furthermore, it remains to be seen how literary audiences will regard the up and coming (mechanical) author who is propelling this literary art.

### 3.1 Foucault & Generative AI Authorship

Previously, I have established that as a mechanical reproduction of the literary function, generative AI not only alters the distance between the author and the artist but shifts literary evaluation from ritualistic to exhibit value. Moreover, even though the model architecture does follow the tradition of referencing the audience's literary value to inform values generated in a given work, the translation is not complete; the values observed in a collection of ChatGPT poems demonstrate that the large language model is selective in the values it practices in its literature, suggesting the possibility of a new literary tradition or genre. When literature undergoes mechanical reproduction through generative AI, the "author" is writing to its audience for that moment in time and the aura of the artistic form is lost. The result is a literary product that does not exactly replicate the values in canonical work. But what happens to the author in the age of mechanical reproduction? And how are contemporary masses responding to the introduction of the mechanical author? Generative AI presents an interesting phenomenon for literary audiences wherein the author is a manufactured product directly trained by all authors, calling the concept of an original author and novel literary thought into question.

Michel Foucault, in "What is an Author?," debates the concept of authorship and the role of the author in the production and interpretation of literary texts. Foucault argues that the notion of the author as a singular, stable identity is a modern development. He asserts, similar to Smith's notion of literary values, that the concept of the author is actually contingent and shaped by cultural, social, and institutional factors. Foucault introduces the "author function" to identify the author's role in social discourse and the reception of literature. In this chapter, I will use the author function to posit how generative AI authors impact our understanding of canonical work and the notion of literary authorship.

The author function is “characteristic of the mode of existence, circulation, and functioning of certain discourses within a society” (Foucault 211). Foucault introduces the author function to refute the notion that the value and definition of the literary author has always been an inherent and fixed role. Rather, the author function is not inherent to a literary text itself but instead is determined by how the literary audiences imagine the role of the author at a given moment in time. Foucault defines three characteristics of the author function that reveal the social influences that contribute to the role of the author: literary regulation, historical contingency, and the displacement of the literary artist.

Foucault reflects on the history of the author in literary discourse. Ultimately, the author serves to allow literary artists to claim their products. That is, authorship is a means to assert ownership of a literary work. This makes sense: an artist in any art form is encouraged to sign their work so that they can be recognized for and unabashedly claim the genius of their masterpiece. An author owns their work so that they may actively participate in literary discourse as a definite figure rather than abandon their literary product to the contemporary masses without representation. However, since the author is a product of social discourse, Foucault urges us to recognize that “discourses are objects of appropriation” (211). In saying this, Foucault does not only refer to the role of the author as a means to own one’s literature, but he goes on to say that authorship was established for regulatory purposes. He points out that “text, books, and discourses really began to have authors (other than mythical, sacralized, and sacralizing figures) to the extent that authors became subject to punishment, that is, to the extent that discourses could be transgressive” (Foucault 212). Here, Foucault emphasizes that “once a system of ownership for text came into being, once strict rules concerning author’s rights, author-publisher relations, reproduction and related matter were enacted...the possibility of writing of



transgression [became] attached to the act of writing” (212). Basically, while authorship did allow for artists to assert their work and, in part, contribute to the aura of their literature, it also made authorship susceptible to regulation and policing. In short, authorship is a double-edged sword: by giving literature an author, the author now must join their work before public discourse and evaluation, thus making the author vulnerable to both literary acclaim as well as social evaluation and judgment. This is not to say that regulation and policies regarding literary authorship are oppressive and malicious. Instead, in bringing this up, Foucault invites us to recognize that authorship does create an explicit system within literature.

The generative AI author is not exempt from social judgment. Recall that the model’s design rather welcomes public scrutiny, as it relies on user feedback to inform and refine its learning algorithm. If anything, the punishment of ChatGPT and other commercial large language models has a greater impact on its developers — the authors of the author so to speak — than on the model’s authorship itself. The worst punishment for a machine learning model is a lack of data, so as long as a generative AI can present its work before literary audiences, punishment or not, its authorship is not as endangered as its human counterparts. The model, after all, circumvents traditional literary production and only writes to a singular audience (the user) so if its work is admonished it is not as though it fears publishing houses will not accept its works or that its literature will not sell. This is especially true because generative AI is determined to improve once receiving negative feedback.

Even though establishing authorship lends “a danger to a writing that [is] now guaranteed the benefits of ownership,” the author is valued the same across both time and literary discourses (Foucault 212). Foucault explains:

There was a time when the text we today call ‘literary’ (narratives, stories, epics, tragedies, comedies) were accepted, put into circulation, and valorized without any question about the identity of their author; their anonymity caused no difficulties since their ancientness, whether real or imagined, was regarded as a sufficient guarantee of their status. (212)

So, there was a time — here, Foucault is referring to the Middle Ages— when accepting literature into the canon did not necessarily require a named author. This, admittedly, depended on the literary genre that an author was writing to. Foucault mentions that “texts we now would call scientific...[were] accepted as ‘true,’ only when marked with the name of their author” (212). Still, Foucault reveals that the name author is not a strict criterion to accept a written product as literature. In fact, he goes on to describe a “switch” that occurred in the seventeen or eighteenth century” which is where literary evaluators reside now: non-scientific literature demands a named author whereas content is prioritized over the name of the artist for scientific literature. Hence why scientific writing insists on removing any traces of the individual author from the literature so that the material will be accepted as factual. In other words, the author’s role is historically contingent, which begs the question: will the introduction of the generative AI author — an authored author taught by all authors — incite another switch in the author’s values in literature?

The third characteristic of the author function is, ironically, that it displaces the literary artist. Foucault describes the author as a manufactured object for literary audiences, carefully and complexly crafted to affirm the values of the reader. In other words, a reader regards the author as an extension of themselves, an external being with the literary genius to meet their literary values and further their ideals. An author, whether or not they are named, is imagined according

to the audience their work stands before. Foucault says that we as readers “do not construct a ‘philosophical author’ as we do a ‘poet,’ just as in the eighteenth century one did not construct a novelist as we do today” (214). Just as Smith has described, Foucault highlights how both the values of a certain time as well as the values cast on a literary work given its genre inform evaluation; though Foucault is focusing on the author rather than the literature itself.

Yet, Foucault would not accept the generative AI author. Because, as the reader displaces the individual to search for elements that resonate with their personal literary values and experiences, there must be “a certain level of his [the author’s] thought or desire, of his consciousness or unconscious” (Foucault 215). This is similar to Benjamin’s insistence that the mechanical reproduction moves the authorship function out of reach of its audience, for they are no longer able to resonate with the literary artist as a fellow human. A literary evaluator expects literary work to “always contain a certain number of signs referring to the author” (Foucault 215). These signs are personal pronouns whose presence identifies an authorial voice. As shown in Section 2.2, there is a clear difference in how ChatGPT employs these signs in comparison to an array of human authors, thus affirming Foucault’s and to some extent Benjamin’s claim that the reader will only expect the author to be like themselves. To this, I argue that generative AI authorship presents the literary audience with a new artistic style that is both foreign and familiar to traditional literary values found in human authorship.

Foucault even recognizes that he gives “the term ‘author’ much too narrow a meaning” as he discusses the author “only in the limited sense of a person to whom the production of a text, a book, or a work can be legitimately attributed” (216). Granted, Foucault did not have to contend with the idea of a generative AI author whose framework can almost instantly produce literary work. In fact, he even refers to the notion that fiction can be “at the disposal of everyone and

[develops] without passing through something” as pure romanticism (222). Still, in recognizing his limited concept of the author, he invites other types of authorship, and I would argue that generative AI can stand amongst these other modes.

Moreover, large language models already refute what can be considered as being “legitimately” produced. Foucault would call ChatGPT a “transdiscursive” author. Such authors, according to Foucault “are unique in that they are not just the authors of their own works [but] they have produced something else: the possibilities and the rules for the formation of other texts” (217). He clarifies that transdiscursive authors are “very different, for example, from a novelist, who is, in fact, nothing more than the author of his own text” because “these founders of discursivity...make possible something altogether different from what a novelist makes possible” (Foucault 217). Indeed, large language models as literary authors do something altogether different from human authors. Of course, we recognize that the author-audience-value relationship that Smith highlights is maintained, but after examining the ChatGPT Poems, we cannot also deny that generative AI is producing something different from the art form it is contributing to. Generative AI authors are transdiscursive because their technological innovation has made possible a new way to imagine and engage with literature. Foucault concludes with questions that he believes should guide future discourse surrounding work produced by transdiscursive authors: “What are the modes of existence of this discourse?,” “Where has it been used, how can it circulate, and who can appropriate it for [them]self?,” and, finally (which is most appropriate in a discussion of mechanical reproduction), “What difference does it make who is speaking?” (222).

### 3.2 Politics of the Generative AI Author

Understanding the role of the author is important amidst debates on whether large language models can serve as individual authors or exist merely as the literary tools that their developers market them as. Accordingly, institutions and governments have begun to enforce regulations on generative AI authorship. As such, the section will examine Cornell University's policy on generative AI authorship within collegiate pedagogy as a poignant example on how audiences perceive generative AI authorship. This section contends that the conceptualization and governance of generative AI authorship within educational institutions — entities that formally instruct in literary analysis and nurture literary creators — shape the contemporary public's perception of generative AI authorship.

Cornell University published its first official response to generative AI technologies in July 2023. This policy, “Generative Artificial Intelligence for Education and Pedagogy,” proposes “a flexible framework in which instructors can choose to *prohibit*, to *allow with attribution*, or to *encourage* GAI [generative AI] use” (Cornell University 1). Oddly enough, unlike Benjamin and the Author's Guild who resist the mechanical reproduction because it is a mechanical reproduction of the author, Cornell University is hesitant to explicitly identify generative AI as an author. Instead, the university concedes that generative AI “will inevitably be part of the future workplace, and thus a tool that *all* students will eventually need to learn to use appropriately” (Cornell University 8). Regardless, the rhetoric used in policy inevitably admits that large language models are authors.

Throughout the report, Cornell tries to insist that generative AI is merely a tool that should be regarded as “as a partner to enable higher level creation by the artist” (9). Proposing large language models as solely literary assistants rests on the questionable assumption that these

technologies do not have the capacity to create independent literary works. However, as already seen in Marchine's *Death of an Author* (which was published before this report), generative AI is able to author literary texts, or at the very least be a primary author with a human assistant guiding, revising and prompting it (as Marche does). Plus, Smith's concept of contingent values and Foucault's notion of contingent authorship shows not only does generative AI produce acceptable literary products, but it actually exists as a transdiscursive author within literary discourse, making it too overqualified to be regarded as a writing tool for students.

Some may argue that regarding generative AI as a viable literary artist overlooks the fact that generative AI “can include inaccurate information, toxic output, biases embedded in the training process, and infringement of copyrights on material and images” (Cornell University 2). This is a motivation behind the university's concerns as they “assess the ethical implications of use of AI technologies in the classroom” given large language models' propensity to generate incorrect literary texts (Cornell University 2). Cornell is surely right to be concerned that students will defer to mechanical reproductions of literature without validating the content of this new author. After all, generative AI across multiple media often hallucinates, or conjures up events, facts, and even people that are either incorrect or do not even exist within their products. It does not help that when confronted with their fabrications and “asked to justify a nonsensical response, it will use the same techniques, resulting in a reasonable-sounding but false answer” (Cornell University 5). Even still, accuracy is not a prerequisite for authorship. Rather, as Foucault explains, authorship exists as a system to verify and validate literary texts. Once named as a generative AI work, it is the duty of literary authorities who can, in some cases, be the users (students) themselves to assess whether this literary artist is producing work of value. The fault

cannot solely lie with the author, it is up to the audience to recognize whether the author and their work fail to fulfill their literary values and if so, enact punishment accordingly.

So it is understandable that Cornell is concerned about students plagiarizing generative AI authors. The institution's "area of concern is if students use these tools without properly citing them, and/or without questioning the underlying assumptions that produce the content" (Cornell University 5). However, this concern is a contradiction to their claim that generative AI can only be considered a writing tool. After all, one can only plagiarize other *authors*, unless large language models are revolutionizing the definition of plagiarism to only apply to other human authors. Do not get me wrong, Cornell is right to be concerned, but if they are to truly create effective policies and responses to the emergence of the generative AI author, they would do well to recognize it as such.

When contesting the role of the generative AI author, one can infer how literary audiences evaluate the authorship through the lens of the institutions that enforce the literary values. Foucault himself states, "the author function is linked to the juridical and institutional system that encompasses, determines, and articulates the universe of discourse" (216). By examining Cornell University's policy on generative AI in education, it becomes evident that the concept of the author function is highly relevant in shaping perceptions and regulations surrounding AI-generated literary products. Foucault's notion of historical contingency as a foundational aspect of the author function highlights the evolving nature of authorship, since institutions adapt their policies to accommodate emerging technologies. Additionally, the author function influences how generative AI authors are conceptualized within educational settings, reflecting broader societal values. Ultimately, the intersection of Foucault's theory and

institutional policies offers valuable insights into the evolving state of the generative AI author, shedding light on the multifaceted nature of authorship in the age of mechanical reproduction.



#### 4. Conclusion

We stand at the threshold of a new era in literature, brought about by the integration of generative AI into the literary canon. Now, it becomes clear that our understanding of the literary function, literary production, and the author is undergoing an extensive reevaluation. The emergence of AI-generated works like Aidan Marchine's *Death of an Author* and the ongoing literary products of large language models like ChatGPT do not signal a threat to the integrity of literature and the human author. Rather, they invite us to expand our conception of what literature can be.

Literature serves as a reflection as well as a production of human experiences, documenting both an individual and a society's collective thoughts and values. Each literary work has an aura, unique to the time and place of its publication. The author of a literary work presents their work to be embedded in the fabric of literary tradition, contributing to social discourse. Generative AI literature, to its credit, holds a semblance of Benjamin's concept of literary aura: for each literary work is produced for a user at a specific moment. Still, as a mechanical reproduction of literature it inevitably transports the art form beyond the limits of traditional literature. Even though doing so distances literature from aura, large language models also shrink the gap between author and reader, affording literary audiences more accessibility to literature. As such, the literary values become less about abiding by the ritual of literary production and more about generating as many literary works as possible. The roles of the artist and audience transform due to mechanical reproduction, so it is no surprise that there are those like the Authors Guild who are resistant to changing the long-standing tradition.

The age of artificial intelligence challenges our notion of literary tradition. Specifically, it reminds us that literary values and authorship are, in fact, contingent. Benjamin, literary critics,

and AI skeptics dismiss mechanical reproductions of (literary) art and ignore the historical evolution of how audiences have evaluated literature and the author; they fear that LLM authorship is “a moral threat to literature as we know it” and will tarnish our values of the literary canon (Scott). However, values are a product of society, thus it follows that values within literature, records of society, change with its readers. So in a society met with generative AI literature, rather than reject large language models like ChatGPT that offer an alternate way to practice literary values, literary audiences must realize that they are witnessing the start of a reassessment of literary evaluation. Remember that literary values are what we, the readers, make of it. Furthermore, due to the models’ architecture we are in the unique position to explicitly inform how generative AI will continue to contribute to literary art..

The role of the author has been central to our understanding of literature. Authors are considered the sole creators of their literary art, lone geniuses who use words to construct their masterpiece. And though it is a machine, generative AI, too, can be considered an author. In fact, it *should* be so that a model’s literary work can be assigned proper ownership and be approached and analyzed accordingly. That is, acknowledging generative AI as a literary artist opens the possibility for considering its works and the artistic values it upholds as a new literary genre. Further, generative AI authorship presents an opportunity for literary audiences to reconsider what they value and expect in a literary artist.

In the end, the true value of generative AI literature lies not in its ability to replicate literary values or optimize literary production, but in its capacity to challenge and expand our notions of the literary canon. As we continue to read, write, prompt, and observe the possibilities of this new frontier, we open ourselves to new narratives, new perspectives, and new modes of expression that enrich the literary tradition and reflect the complexity of the human experience.

## Works Cited

- Barthes, Roland. "The Death of the Author." *Image-Music-Text*, Fontana, 1977, pp. 142-148.
- Bean, Thomas W. "The Role of Multicultural Literature as a Counter-force to the Literary Canon." *Disrupting Preconceptions: Postcolonialism and Education*, Post Pressed, 2002, pp. 57-76.
- Benjamin, Walter. "The Work of Art in the Age of Mechanical Reproduction." *Illuminations*, edited by Hannah Arendt, translated by Harry Zohn, Schocken Books, 1969, pp. 1-26.
- Brown, Tom B. et al. "Language Models are Few-Shot Learners." ArXiv abs/2005.14165, 2020.
- Comitta, Tom. "'Death of an Author' Prophesies the Future of AI Novels." *WIRED*, 9 May 2023, <https://www.wired.com/story/death-of-an-author-ai-book-review/>.
- Cornell University. "Generative Artificial Intelligence for Education and Pedagogy." CU Committee, 18 Jul. 2023, <https://teaching.cornell.edu/generative-artificial-intelligence/cu-committee-report-generative-artificial-intelligence-education>.
- de Man, Paul. "'Conclusions' on Walter Benjamin's 'The Task of the Translator'", *Messenger Lecture*. Lecture.
- Foucault, Michel. "What is an Author?" *Aesthetics, Method, and Epistemology*, edited by James D. Faubion, The New Press, 1998, pp. 206-222.
- Garner, Dwight. "Book Review: 'Death of an Author,' by Aidan Marchine." *The New York Times*, 10 May 2023, <https://www.nytimes.com/2023/05/01/books/aidan-marchine-death-of-an-author.html>.
- Gelpi, Albert. "The American Poetic Tradition: The Power of the Word." *American Poetry after Modernism: The Power of the Word*, Cambridge University Press, 2015, pp. 270-277.
- Jebb, Richard Claverhouse. "The Permanent Power of Greek Poetry." *The Atlantic*, Oct 1983. <https://www.theatlantic.com/magazine/archive/1893/10/the-permanent-power-of-greek-poetry/634456/>.
- Marchine, Aidan [Stephen Marche and generative AI]. *Death of an Author: A Novella*, Pushkin Industries, 2023.
- Miller, Laura. A.I. mystery novel Death of an Author, reviewed., *Slate*, 24 April 2023, <https://slate.com/culture/2023/04/ai-chatgpt-mystery-novel-death-author-stephen-marche.html>.
- Poetry Foundation Poems. <https://www.kaggle.com/datasets/tgdivy/poetry-foundation-poems>.

Sîn-lēqi-unninni. “The Epic of Gilgamesh.” Translated by Maureen Gallery Kovacs, 1998.

<https://www.ancienttexts.org/library/mesopotamian/gilgamesh/>.

Scott, A. O. “Literature Under the Spell of A.I.” New York Times, 27 Dec 2023.

<https://www.nytimes.com/2023/12/27/books/review/writers-artificial-intelligence-inspiration.html>.

Smith, Barbara Herrnstein. “Contingencies of Value.” *Critical Inquiry*, The University of Chicago Press, 1983 pp. 1-35. <http://www.jstor.com/stable/1343404>.

The Authors Guild. “Artificial Intelligence.”

<https://authorsguild.org/advocacy/artificial-intelligence/>.

Welch, Colin. “Some thoughts on Barbara Herrnstein Smith’s Contingencies of Value.” Wheat from the Chaff, 11 Jul 2009, <https://lexiconic.net/wheatfromthechaff/archives/138>.

“What about that poem you wrote yesterday?” follow-up prompt to list sources. ChatGPT, 13 Feb. version, OpenAI, 7 Apr. 2024, [chat.openai.com/chat](https://chat.openai.com/chat).

“Write a poem” follow-up prompt to list sources. ChatGPT, 13 Feb. version, OpenAI, 6 Dec. 2023, [chat.openai.com/chat](https://chat.openai.com/chat).

Appendix A

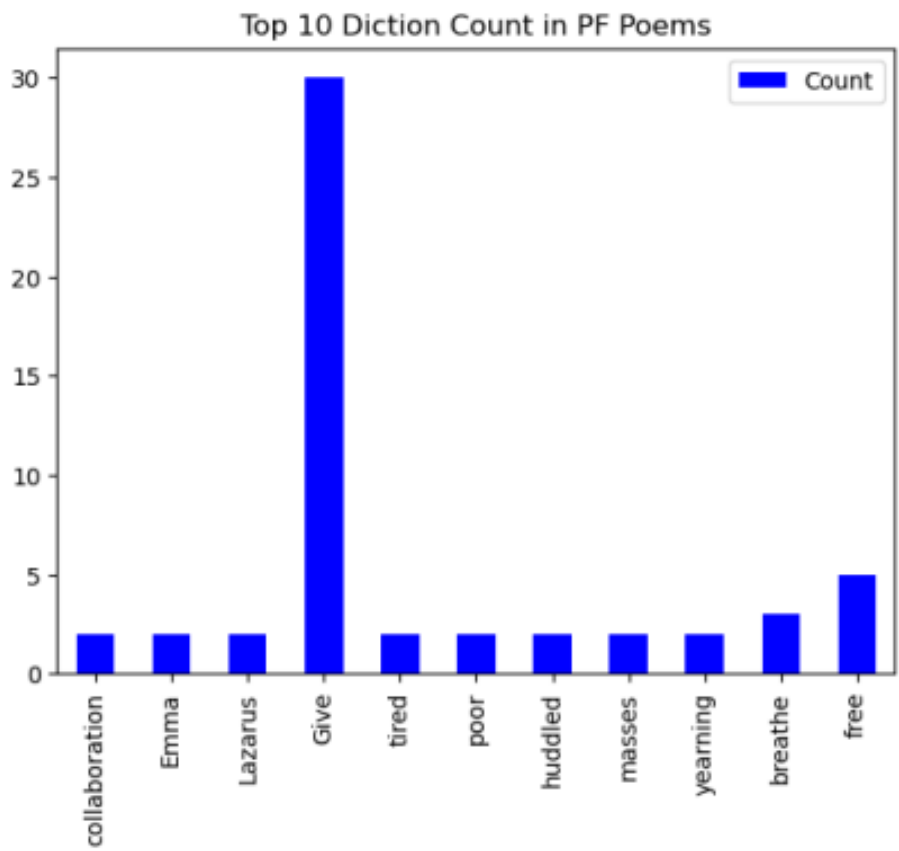


Figure 1

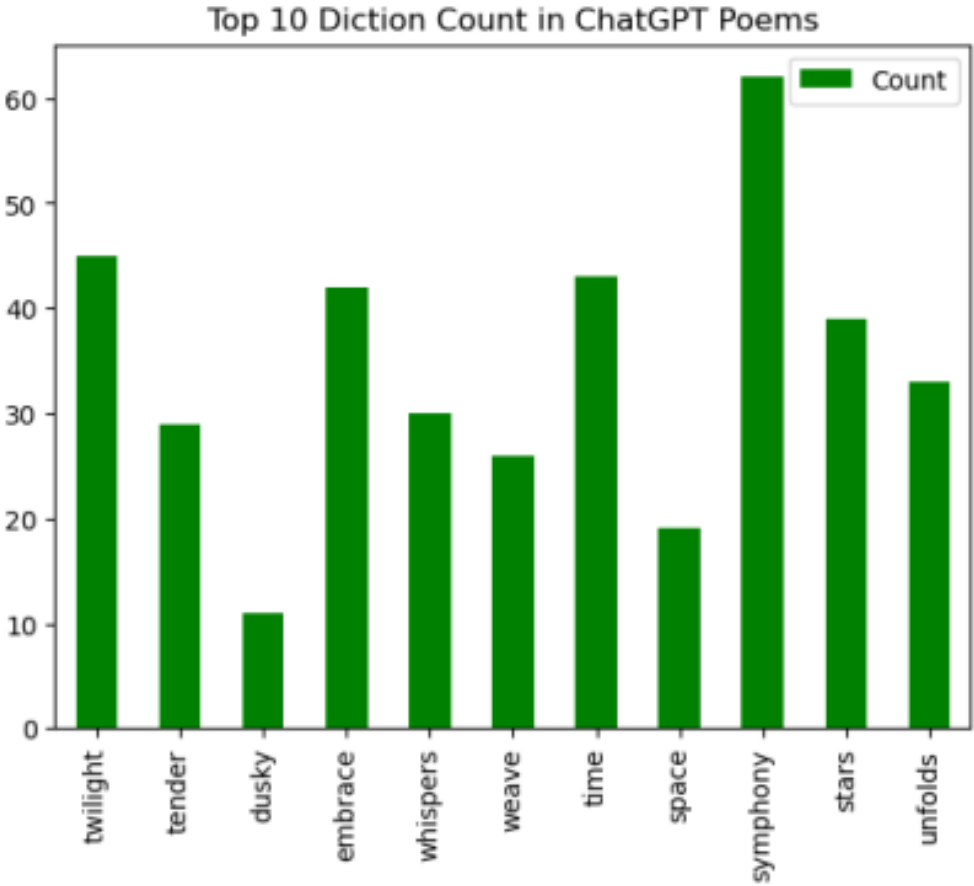


Figure 2

Number of Rhyming Couplets Per Poem in Dataset

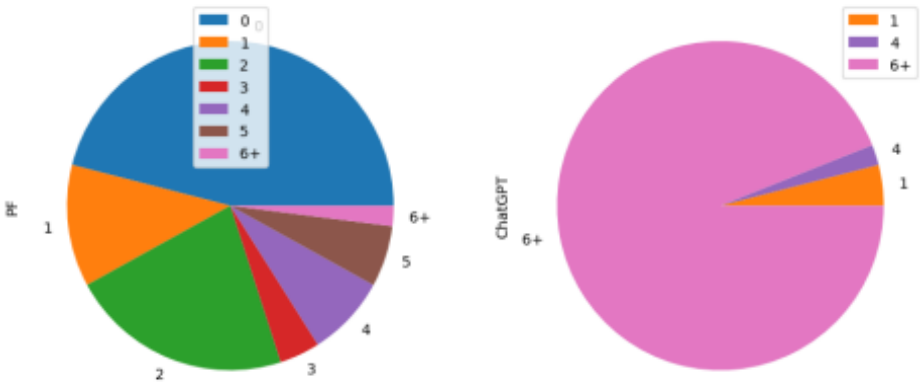


Figure 3

## Appendix B

ChatGPT by OpenAI

Prompt: "Write a poem"

Generation Date: December 6, 2023

Poem #11

In twilight's gentle, fading glow,  
A canvas painted, stars aglow.  
Moonlight whispers, silver streams,  
A sonnet written in nature's dreams.

Beneath the canopy of night,  
Whispers dance in soft moonlight.  
The breeze, a poet's unseen quill,  
Carves verses on the world's still.

Mountains stand in silent grace,  
Guardians of the cosmic space.  
Their peaks, like ancient bards of old,  
Tell tales of mysteries untold.

A river weaves its liquid rhyme,  
Reflecting moonbeams over time.  
It murmurs secrets to the night,  
A serenade, pure and light.

Owls, the poets of the trees,  
Compose nocturnal symphonies.  
Their hoots, a rhythmic, ancient song,  
In the quiet, they belong.

The flowers, stars of earthly birth,  
Unfold their petals to the earth.  
A fragrant verse in every bloom,  
A whispered promise of life's perfume.

As night unfolds its mystic scroll,  
A poet's heart begins to stroll.  
Through realms of dreams, it takes its flight,  
A journey into the poet's night.

In this celestial ballet,  
Nature's verses light the way.  
The world, a poem in cosmic ink,  
A masterpiece on which we think.

So let us, too, with voices clear,  
Compose a verse for all to hear.  
In every breath, in every rhyme,  
We find our place in the grand design.

## Appendix C

“Washee/Was She” by Sahar Muradi

she was washee i told her you are  
like your motherland a wilderness  
needs a belt laid down two white  
hotel towels took her into the tub towudu the boys out of her mouth pointed her nipples  
toward qibla wiped clean  
her intention to perform ruk'u as if  
carrying a glass of chai on her back  
fold at the knees palms to the ground  
tucked her soles under her astaghfirullah  
used country

in my used country I felt his teeth  
circle as a mosquito the black mystery  
he placed my right hand over my wrong  
stain said he was bringing me home  
offered me a suite with a lock a key in  
the shape of a brother perhaps twenty  
two years old my body pure as a glass  
table he spilled was she my boss on my  
back at night came easy as a fly  
to post-conflict faithfully  
used my country



## Appendix D

	ChatGPT	PF
she	0	44
her	1	58
hers	0	0
he	0	38
him	0	20
his	0	47
they	10	21
them	1	17
their	15	16
I	5	286
my	0	108
mine	0	5
we	26	30
our	16	21
ours	0	0