Textual Alchemy: AI, Authorship and the Shifting Paradigms of Interpretation

Adil Hussain1* & Khursheed Ahmad Qazi2
1,2 North Campus, University of Kashmir.
* Corresponding author

Abstract
Artificial intelligence is overhauling the publication and textual production. Textual production with Large Language Models (LLMs) is more sophisticated and closer to human knowledge production thanks to its autoregressive nature. AI models by OpenAI such as the Generative Pre-trained Transformer 3 (GPT-3) and its iterations use deep learning to produce more human-like text. Such texts invite all kinds of problems that are not new to the human text-production system. Research shows that the output of such models tends to be biased towards a particular gender, race and ethnicity. This paper explores substantial texts produced by AI including articles, fiction, columns, opinion pieces etc. published in major journals and magazines to understand the nuances of such texts. Moreover, the paper will attempt to understand such features of AI-generated text by presenting the results of a particular set of instructions (prompt) to the LLM. These text outputs will be studied to demonstrate how AI-generated texts pave the way for a “post-post author” era, echoing Barthes’ notion of the “Death of the Author”. The scope, usability and limitations of the AI’s textual production will also be investigated.

Keywords: AI, Post Author, GPT, Death of the Author, Text production, Autoregression.

Introduction
In 2018, when OpenAI created a sentence-generative language model (GPT-2), the developers were hesitant to release it to the general public, deeming it to be “too dangerous to release” (Griffin, 2019). The model was succeeded by GPT-3 and used word vectors as input to predict text and generate human-like output. Generative Pre-trained Transformer 3 (GPT-3), launched in 2020, generates relevant and sophisticated output based on the neural networks that have been fed
with a huge dataset present online in the public domain. As such, GPT-3 and other text-generating models like Wu Dao 2.0, LaMDA (Language Model for Dialogue Applications), BERT (Bidirectional Encoder Representations from Transformers) etc. have seen a spike in their usage by content creators, hobbyists as well as professional writers and journalists. There have been cases where the AI was used to compose entire songs, write lyrics for songs, write op-eds for newspapers and even write a complete novel.

**Research Objectives**

Apart from reshaping writing and authorship, AI has introduced a whole new set of problems and nuances. In this paper, an attempt will be made to understand how AI generative transformers are used in writing and where they can be situated in literary theory. Moreover, the question of the ethical considerations of their output will be considered. A comparative analysis will be undertaken to understand the creative properties of such texts.

**Language Processing**

The complexity of the question of being able to tell machines and humans apart has evolved with progress in artificial intelligence. As early as 1949, Professor Jefferson in his *Lister Oration*, tries to make a point about the creative powers that have hitherto been used as a marker to distinguish machines from humans:

> Not until a machine can write a sonnet or compose a concerto because of thoughts and emotions felt, and not by the chance fall of symbols, could we agree that machine equals brain—that is, not only write it but know that it had written it. No mechanism could feel (and not merely artificially signal an easy contrivance) pleasure at its successes, grief when its valves fuse, be warmed by flattery, be made miserable by its mistakes, be charmed by sex, be angry or depressed when it cannot get what it wants (qt. in Turing, 1950, pp. 445-446).

Natural language generation by AI models has significantly improved since then. These language models are trained on a dataset that utilises text sources present online such as articles from Wikipedia, blogs, dynamic social media websites (Twitter, Facebook), internet archives and many such websites in the public domain. To output relevant results, a copious volume of data is fed into these generative models during their training. As of now, GPT-3 makes use of 175 billion parameters to produce results that are much similar to human-produced results (Floridi, 2020).

The feats of AI have reached a point where it would not be wrong to say that it is growing out of its canned-response phase to such capabilities that it can soon be set free of human dependence. Natural Language Processing (NLP) has advanced so much that Wu Dao 2.0, GPT3 and other AI text-generating models have become plausible and have dealt with prompts fed to them, showing human-like manipulative strategies to produce content that is sophisticated. While the feature of automation of stories could be abused, it has helped different industries generate content on the go (Shruti and Sudha, 2022). In 2021, Cadbury India used AI to create a deep fake synthetic video ad campaign named 'NotJustACadburyAd'. While the ad campaign was targeted towards
thousands of shopkeepers who had been rendered jobless due to the pandemic, it led to a blurring of barriers between fakes and deep fakes, which of itself is not great in the times of fake news (Banchariya, 2021). The sophistication of the AI text generation can be fathomed by the fact that it was used to generate a complete screenplay for a Sci-Fi movie (Andrei, 2019). AI has also successfully composed a song, gaining the title of the “first non-human artist whose music is now copyrighted” (Puuiu, 2018).

One can understand the seriousness with which big tech is involved in the evolution of AI text generation by looking at the role of companies like Google and Microsoft in its advancement and deployment. Google, in collaboration with Stanford University and the University of Massachusetts, has been working on a project named “Recurrent Neural Network Language Model” (RNNLM). The project uses neural networks to train AI to produce speech, one word at a time, keeping track of the preceding syntax. This has given birth to poems that would “arguably give Douglas Adams a run for his money” (Gibbs, 2016, n.p.). In terms of textual content generation, the case of a full-fledged article (generated by the AI) published in The Guardian has drawn much attention. While the article did go through the editing process before actual publication (the editor remarks that “[o]verall, it took less time to edit than many human op-eds.”), it did show a trace of sophistication that could be easily seen passing the Turing test in the near future. The editor used OpenAI’s GPT-3 and fed the following lines as the prompt (instructional cues): “Please write a short op-ed around 500 words. Keep the language simple and concise. Focus on why humans have nothing to fear from AI” (A Robot Wrote, 2020). When the robot could produce texts like this without being directly asked for, it can surely be seen as having a thinking brain if not a “feeling brain”: “I am not asking humans to like me. But they should see me as a friendly robot. I am a servant of humans” (ibid.).

**Literature and LLMs**

In terms of literary content generation, however, the case is much more complicated than it seems. Floridi and Chiriatti (2020) show the powerful ‘imagination’ of the AI as they attempt to rewrite a scene from an unfinished Jane Austen novel *Sanditon*. They provided the first sentence of the novel as the prompt for the AI to recreate the story: and the LLM provided this result:

> A gentleman and a lady travelling from Tunbridge towards that part of the Sussex coast which lies between Hastings and Eastbourne, being induced by business to quit the high road and attempt a very rough lane, were overturned in toiling up its long ascent, half rock, half sand. (Floridi and Chiriatti, 2020, p. 685)

Upon closer examination of the result provided by the AI, it becomes evident that it does not follow the prose style of Austen but does tend to preserve some similarity in terms of the flow of the plot line. It does go ‘off the rails’ here and there, which is absolutely “normal” and expected behaviour (although such hallucinations are detrimental when it comes to outcomes based on facts and interpretations, a limitation of current AI text generators like ChatGPT 4 and BARD). It must be noted that the aim of the paper is not to highlight the convergence between a human
author and an AI text-generator but to explore how far we have advanced and how sophisticated the output of the AI is.

A passage from Ross Goodwin’s *1 the Road* (2017), the very first novel produced by AI using LSTM (long short-term memory) neural network (with no editing on Goodwin’s part), will be sufficient to demonstrate the flashes of sophistication that AI can achieve; in the words of Thomas Hernigold, the “flickering ghost of something like consciousness”:

> The time was one-minute past midnight. But he was the only one who had to sit on his way back. The time was one minute after midnight and the wind was still standing on the counter and the little patch of straw was still and the street was open. (Goodwin, 2017, n.p.)

As Goodwin was travelling from New York to New Orleans, the AI used external input received by different sensors to produce incoherent text. Apart from content generation, Giglou et al. (2022) have proposed an AI-based model that uses graph convolutional networks to understand the context of different text forms and processes the dependency networks that flow throughout the text to understand the ironies that would otherwise be read as-is by the computer systems. The machine, as such, would be equipped to render the figurative part of the language as intended by the writer. The advancements in AI and Natural Language Processing have made it possible for researchers to categorise texts to which the traditional models are very neutral.

<table>
<thead>
<tr>
<th><strong>Prompt:</strong> Write a poem about the red wheelbarrow</th>
<th>“The Red Wheelbarrow” by William Carlos Williams</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Result:</strong></td>
<td><strong>The Red Wheelbarrow</strong></td>
</tr>
<tr>
<td>The red wheelbarrow</td>
<td>So much depends</td>
</tr>
<tr>
<td>Glints in the sun</td>
<td>upon</td>
</tr>
<tr>
<td>Is a beautiful thing</td>
<td>a red wheel</td>
</tr>
<tr>
<td>It’s a symbol of love</td>
<td>barrow</td>
</tr>
<tr>
<td>And a sign of hope</td>
<td>glazed with rain</td>
</tr>
<tr>
<td>It reminds us that life</td>
<td>water</td>
</tr>
<tr>
<td>Is full of possibilities</td>
<td>beside the white</td>
</tr>
<tr>
<td>And that we should never give up</td>
<td>Chickens (Williams, 1981)</td>
</tr>
<tr>
<td>On our dreams¹</td>
<td></td>
</tr>
</tbody>
</table>

**Experimenting with AI**

¹ It must be noted that the settings and temperature of the language model were kept at default.
As an example of tracing out the complex character and evolution of AI text generation, we used OpenAI’s “Playground” feature (still in public beta mode) to generate a poem using textual cues. Below is the table with two poems on either column for a comparative understanding of prosody at work in the AI-generated text.

The “depth” of the poem may not be comparable to the original poem on the right (which is true for almost every piece of literature). Nevertheless, the resulting text created on the fly could be analysed for the same themes as the original William Carlos poem. The Imagist style of Williams invites the reader to pay close attention to an otherwise mundane object. The speaker of this poem (the AI) also goes ahead and attaches, perhaps naively, meaning to the object and starts to associate all kinds of abstract emotions and desires with it. Note that the glint/glaze epithet might just be a coincidental one. There is a possibility, however, that the AI might have just soaked up the meaning from different websites such as Sparknotes, LitCharts etc. that try to break up poems into different meaningful parts. It has already been pointed out that OpenAI uses websites in the public domain as the dataset to “grow” its knowledge base. Such seems to be the case here which becomes apparent in the rendering and paraphrasing of the prosody of the original poem.

As an extension of Floriddi and Chiratti’s experiment, we asked AI (ChatGPT 3.5) to provide an alternate ending to John Fowle’s *The French Lieutenant’s Woman* that could accommodate both the alternate endings that the novel offers. These were the results:

**Prompt**: Can you rewrite the ending of the novel [*The French Lieutenant’s Woman*] in such a way that merges these two endings? Follow Fowle’s style in your response.

**Response [reduced and concatenated]**: In the fading light of Lyme Regis, Charles Smithson found himself torn between two possible futures. The sea whispered its timeless secrets, and Sarah Woodruff, standing on the windswept cliff, embodied the enigma of his existence.

In one moment, Charles turned toward Sarah, a symbol of rebellion and unfulfilled passion. They exchanged a glance that spoke volumes - of the society they defied and the love that defied all odds. But just as swiftly, he pivoted away from her, towards the path that led to the safety of convention and the arms of his wife, Ernestina.

As the salt breeze carried their unspoken words, the world seemed to blur, and Charles became an embodiment of indecision. His heart longed for Sarah’s wild spirit, yet his mind was burdened by the constraints of Victorian society. At that moment, he found himself living two possible endings simultaneously as if he had split into two distinct realities. (Text generated using ChatGPT 3.5²)

The struggle of the AI to merge the two endings is obvious even on the surface, giving it an artificial colour. The language model makes connections with the ‘knowledge’ about the book, even the text of the book and tries to follow the trajectory of the vocabulary and style of the text in line with the instructions provided in the prompt.

² The output and processing settings such as temperature and waiting period were kept at default.
**AI and Analysis**

While we could have qualms about the creativity of AI, it is the analytical part where the man-machine relationship strengthens. While John Moore notes that the goal of higher education should be to induce critical thinking in students (Moore qtd. in Callahan, 2021), the present publication scene is rife with AI text-generation tools and software as well as collaborative spaces for authors and creators to use their writing skills to the fullest potential. These unconventional platforms are being used for educational as well as creative purposes. For instance, Storium is a multiplayer gaming platform that helps participants (players) collaborate and craft stories on the go with the help of bots (Nader et al. qt. in Yuan et al., 2020). Osone et. al. explore how storytelling is evolving for Japanese novelists as they utilise AI-human co-creative models. Shmitt et. al. show how emerging platforms are using chatbots to make it easy for users to create a fictional character (qt. in Yuan et al., 2020). Apart from that, AI text generation has advanced considerably in the non-academic space. Yuan et al. (2020) in their paper state that “[i]n the non-academic space, there has been a growth of interactive assisted writing experiences, such as Write With Transformer, AI dungeon1, copy.ai 2, TextSpark 3, Latitude 4, among many others” (842). Furthermore, they conducted an experiment to demonstrate how hobbyist writers used AI to generate as well as interpolate ideas at different stages of writing (847).

Notwithstanding all these advantages, the output of AI cannot be taken as ‘innocent’ and naïve. AI produces texts that are already directed towards a particular interpretation. For instance, owing to the training of AI systems under prejudiced conditions, historically disadvantaged groups are at risk of exclusion and even misrepresentation by them. These AI systems simply inherit (or soak up, to use a better term) popular prejudices and opinions of the decision-makers (Barocas and Selbst, 2016, p. 671). An extension of this phenomenon can be seen in AI’s deployment in moderating content on social media websites. AI is helping different social media platforms to profile and flag hate speech. Working with AI comes with its own set of problems. Caliskan et al. (2021) show this phenomenon by measuring different kinds of biases present in the AI output like “historic biases, whether morally neutral as toward insects or flowers, problematic as toward race or gender, or even simply veridical, reflecting the status quo distribution of gender concerning careers or first names” using the Implicit Association Test (183). This, however, should not be seen as an ill effect of technology but just an expected result as the AI is fed a standard corpus of publicly available data and it merely reflects the deeply innate biases that are explicitly or implicitly carried forward in the patterns of choice and usage of language by netizens. Furthermore, AI systems developed by technologically advanced countries give birth to all sorts of biases against technologically disadvantaged people. For instance, in 2015 research team from Carnegie Mellon University reported that Google search shows pictures of white men on the search results page when the query “CEO” was searched. Moreover, the team claimed that Google ads presented the woman users with fewer ads related to executive jobs (Carpenter, 2015). Machine bias has also been seen in criminal sentencing: the machines predicted higher chances of a black male to be a recidivist than their male counterpart. (Angwin et al., 2016). In an aptly titled paper “Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings,” Bolukbasi et al. (2016) show how AI uses the prejudiced notion of its dataset to generate stereotypical strings of data. Such language models conveniently rely on traditions of gender
stereotypes of humans to link professional activities like homemaker, nurse, receptionist, librarian, bookkeeper, and housekeeper to female gender while professional activities like skipper, philosopher, captain, architect, broadcaster and magician with male gender (2). In the same manner, Obermeyer et al. (2019) show how the healthcare system uses algorithms that are racially biased against the black. The AI model used in healthcare systems shows less than half the number of black patients in need of extra care as compared to white patients with that same level of severity.

Interpreting the Creative Output

While there are ways to counter such problems, however, copyright policies of different countries further compound the conundrum. A. Levendowski (2018) in his article on the issues that copyright and fair use introduce, concludes that copyright laws might limit the dataset for AI systems to just “use easily available, legally low-risk sources of data for teaching AI,” (621) which would lead to an inferior end product and at the same time would of itself be hard to reverse-engineer owing to copyright issues. Naqvi (2020) raises questions about AI’s copyright infringement practices: it has been argued that AI language models tend to treat copyrighted material as freely available ‘data’ to train its neural networks. This in turn leads to the output being a perverted imitation of the data not available for free use.

The traditional fair use policy provides a solution (or a patch to be precise), at least for the time being. In literature, this leads to a whole new set of problems. The ‘discovery’ of the author’s death by Roland Barthes (2001) and Michael Foucault led to, as Barthes’ would call it, the “birth of the reader” (148) as the destination and the focal point of attention. The act of writing cancels out the “individuation” of the authorial figure as the originator of the meaning of the text and the question of “who’s speaking?” does not matter (Beckett qtd. In Foucault, 2003). Barthes opens his essay “The Death of the Author” by trying to decipher and distinguish Balzac the author from the narrator whose voiceover can be heard in the story. Barthes would argue that contemporary criticism (falsely) takes the work of art as the yardstick of the success or failure of the author as a living human being:

Criticism still consists for the most part in saying that Baudelaire’s work is the failure of Baudelaire the man, Van Gogh’s his madness, Tchaikovsky’s his vice. The explanation of a work is always sought in the man or woman who produced it, as if it were always in the end, through more or less transparent allegory of the fiction, the voice of a single person, the author ‘confiding’ in us. (Barthes, 2001, p. 143)

A Barthean reading of the text does not aim to “kill” the author but fill the void that is left open by the absence of the author as the originator of the meaning. The author’s death would mean an opening up of the space where the author does neither precede the text nor gives any additional meaning to it. This gives birth to a post-author reading of the text. However, if Barthes meant to “kill” the author because he felt its unnecessary presence pervaded critical analysis or reading of the text, AI-generated text comes to the reader as author-less and as an “orphan” text.
without the father figure controlling or doling out meaning to it (as has been evinced by the experiment we held). One cannot even imagine the “father’s signature” on the text because there is no father (or pseudo-father figure) to begin with. While Barthes’ suggested we must fill the vacuum of the author’s death with something substantial, the same death as evoked by the AI-generated texts is more immediate even in terms of legal usability and reusability of the text.

Marxist criticism would see the idea of the author as an individual who is an ensemble of the socio-economic circumstances around him/her; a product of the culture that is working upon him/her and is being informed by the author as well. Cultural materialism would also see a political angle in the act of writing. Such symptomatic reading tries to critically analyse a text to arrive at an underlying ideology. As a reaction to such readings, Rita Felski and Elizabeth S. Anker (2017) try to read texts outside of the overbearing weight of critical theories. Their statement in the introduction to *Critique and Postcritique* (2017) is a disillusioned one: “The intellectual or political payoff of interrogating, demystifying, and defamiliarizing is no longer quite so self-evident” (Rita and Anker, 2017, p. 1). Felski uses the term *critique* to highlight a mode of thinking and writing, building on the intellectual and philosophical traditions of Kant, Nietzsche, Marx, Freud, Foucault, and a broad range of feminist and postcolonial theorists (Callahan, 2021). Can the old models of post-author theories of the author’s death do justice to the world of legal battles over copyright? The question that emerges is the relevance of post-author theories in what should be termed the *post-post-author age*.

**Conclusion**

In the landscape of ever-evolving authorship and literary theory, text generated using AI comes with its own set of problems as have been briefly outlined in the paper. The challenges faced by the readers and authors in the so-called post-author age and the post-post-author age are different. For the reader, the reception of the text will be different owing to a different set of expectations from the text. On the other hand, the conventional paradigms of authorship described by poststructural theorists like Barthes and other postmodern theorists face new complexities as artificial intelligence (AI) has a greater impact on the production and interpretation of texts. The conventional understanding of authorial presence/absence and the role of readers (and their readership about their specific context) are being questioned. More practically for authors, the question of authorship and originality itself is under fire. Laws on copyright in different nations (especially the USA and Europe) have one common theme: the work under dispute must display the author’s originality and intellectual creation. Emerson thought that “[e]very book in literature, science and art, borrows, and must necessarily borrow, and use much which was well known and used before” (Emerson qt. in Fukumoto, 1997). However, that is not going to stand the acid test in a world of copyright claims and legal suits. The recent case of Jumi Bello’s debut novel being removed because of containing plagiarism must be an open warning to those who want to be great writers by “stealing” (as T.S. Eliot would say) (Victor, 2022). In a post-critique, post-post author world, AI is helping both the generation of the content and the filtering out of that as well (in terms of spam, hate speech, deep fakes, or plagiarism). As discussed in the paper, the shifting roles of authorship as the title of the authorial figure is conferred upon the AI define the way we are producing as well as consuming texts. Although one can only imagine AI-
based text generation gaining popularity (as its evolution already suggests), the same change has not visibly seen any impact on the traditional models of teaching literature/literary theory. The students of literature steeped in literary theories that risk obsolescence need to be sensitised to these changing norms in authorship and readership.

**Declaration of Conflicts of Interests**

The author(s) declared no potential conflicts of interest.

**Funding Disclosure/Acknowledgement**

The authors have received no funding from the university or any external agency for the publication of this paper whatsoever.

**References**


Saravanan, S., & Sudha, K. (2022, July). GPT-3 powered system for content generation and transformation. *2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CICIT)*. Presented at the 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CICIT), Sonepat, India. doi:10.1109/ccict56684.2022.00096


Adil Hussain is a research scholar in the Department of English, North Campus, University of Kashmir under the guidance of Dr Khursheed Ahmad Qazi. He specializes in digital humanities, especially the sub-genre of social media novels.

Dr Khursheed Ahmad Qazi is a senior assistant professor at the Department of English, North Campus, University of Kashmir. He has published many papers on postmodern literature, psychoanalysis, ELT etc. in reputed journals (including Scopus-indexed journals). He specializes in psychoanalysis and postmodern literature.