# Copyright in Generative AI Models

# Group AI

CAP 5602

Maria Paula Sanchez
5914168
Florida International University
Miami, Florida
msanc371@fiu.edu

Sean Dollinger
6446079
Florida International University
Miami, Florida
sdoll004@fiu.edu

Jayavardhan Reddy Samidi
6396256
Florida International University
Miami, Florida
jsami003@fiu.edu

## ABSTRACT

This project investigates the capabilities of artificial intelligence, specifically utilizing the GPT-3.5 model, in generating children's stories, and differentiating them from human-written texts using our model using GPT-2. Through the development of an AI model integrated with a Streamlit application interface, we conducted a comparative analysis using metrics of perplexity, burstiness, lexical diversity, sentence complexity, and cosine similarity. The project aimed to address the creative potentials of AI in literature and its implications on copyright law, focusing on the generation of unique content and the identification of potentially infringing material. By creating and analyzing a balanced dataset of human-written and AI-generated stories, the project establishes classification thresholds that accurately distinguish between the two, offering insights into AI's role in content creation and its challenges in the digital copyright landscape. This project not only contributes to the understanding of AI-generated literature's originality but also proposes a Copyright Alerting System bolstered by human review to ensure copyright compliance effectively.

## CCS CONCEPTS

• Tokenization • TF-IDF • Cosine Similarity

## KEYWORDS

Copyright, Perplexity, Burstiness, Sentence Complexity, Lexical Diversity, GPT-2, Artificial Intelligence (AI), GPT-3, ChatGPT.

## 1 Related Works

An explosion in research has been initiated, particularly on the success of AI detection systems, due to the popularity of AI-generated content and its consequences for academic integrity. Chaka (2024) performed a recent evaluation that offers a thorough analysis of 17 publications that were published between January 2023 and November 2023. The focus of the review is on the effectiveness of these technologies in the higher education sector. This assessment of the literature is timely and relevant to our project's goals of differentiating AI-generated children's stories from human-written ones.

Chaka's research analyzes the various detection methods described in these investigations, pointing out the weakness and difficulties they have in accurately identifying texts generated by AI. The studies include a wide range of disciplines, including English language studies, the hard sciences, and medicine, highlighting the widespread concern in a variety of academic subjects. Additionally, this research emphasizes how important artificial intelligence (AI) is in educational environments, especially when it comes to promoting or compromising the integrity of academic and scientific writing.

The review clarifies the kinds of texts that were assessed—from fully composed works by AI models like ChatGPT to slightly modified or paraphrased versions—as well as the range of detection techniques that were employed, which ranged from two to sixteen. Though not without significant drawbacks, Crossplaig and Copyl eaks proved to be some of the more effective solutions among the detection tools in specific situations.

The main conclusions drawn from the evaluated publications emphasize the necessity of multifaceted strategies that combine human review with a variety of AI and anti-plagiarism techniques to successfully identify and distinguish AI-written content. These studies also highlight the urgent need for educators to receive digital literacy training, for detection tool creation and improvement, and for evaluation standards across academic publications to be updated in light of increasing AI capabilities.

In accordance with these results, our effort aims to overcome some of the issues found while also adding to the body of knowledge in the new area of AI-generated children's story detection. Our approach, which combines stylistic and linguistic metrics, is consistent with studies that suggest using many tools together to increase detection process reliability. Therefore, the review by Chaka (2024) provides valuable insights and establishes a basis for future research like ours, which aims to innovate in the areas of AI content detection and copyright alert systems. It also serves as a foundation for the ongoing conversation about AI in literary production and its implications on copyright law.

## 2 Introduction

Artificial intelligence presents new potential and challenges for copyright law, especially in literary works. The launch of ChatGPT, a generative artificial intelligence (AI) chatbot owned by OpenAI (OpenAI, 2022), on 30 November 2022, had a domino effect in cyberspace and in the real-life world. It not only rattled the AI world in which generative AI chatbots, which before ChatGPT were relatively unknown, suddenly emerged or announced their presence [1-4], but it also led to the emergence of AI content detection tools intended to detect and to differentiate between AI-generated and human-written texts [5], Such advancements underscore the necessity to evaluate the originality and copyright status of literary materials with unprecedented precision. In one revealing study, blinded reviewers identified 68% of abstracts generated by ChatGPT correctly, yet 14% of original abstracts were mistakenly flagged as AI-produced [6].

The goal of this project is to determine how much AI can generate original children's stories and establish a methodology for differentiating them from human-written texts. In the context of this study, copyright is defined as a measure of how alike AI-generated text is to works generated by humans. We conduct a thorough analysis utilizing a variety of linguistic variables. Recognizing the challenge that arises when generative models like ChatGPT produce content like human-created works, our project focuses on using a variety of linguistic features to evaluate the level of similarity between these children's stories. By using a variety of linguistic and stylistic metrics, we aim to highlight the creative potential of AI and its impact on publishers, writers, and the legal system around copyright. The use of these measurements simplifies the process of finding commonalities in children's stories while also highlighting AI's growing capacity for creativity. Serving as a bridge from detailed technical analysis to practical real-world usage, this project contributes to the academic discourse on artificial intelligence and copyright and has practical implications through the development of an AI detector and a Copyright Alerting System. In doing so, the project addresses the limitations of current copyright enforcement mechanisms in digital spaces and suggests a balanced approach that combines automated detection with human oversight. To achieve these goals, our project explores the complex interplay between creativity, technology, and law, offering detailed insights into the role of generative AI in modern literary production and copyright law.

The following image showcases a piece of one of the stories we use in our dataset, in this particular example our model predicted correctly if the children story was generated by a human or AI.

| Human | AI |
|---|---|
| "Once upon a time, three neighbours living in a village were having trouble with their crops. Each of the neighbours had one field, but the crops on their fields were infested with pests and were wilting. Every day, they would come up with different ideas to help their crops. The first one tried using a scarecrow in his field, the second used pesticides, and the third built a fence on his field, all to no avail." | "In a small village, three neighboring farmers were facing a crisis, as their respective fields were plagued by pests, causing the crops to perish. Each farmer attempted a different method to salvage their harvest: one erected a scarecrow, another applied pesticides, and the third constructed a fence, yet all efforts were futile." |

**Figure1: Story Example Correctly Predicted**

The second image shows another children's story used in the data which our model failed to detect the AI generated story, classifying it as a human written story.

| Human | AI |
|---|---|
| "One day, king Akbar asked a question in his court that left everyone in the courtroom puzzled. As they all tried to figure out the answer, Birbal walked in and asked what the matter was. They repeated the question to him.<br>The question was, "How many crows are there in the city?"<br>Birbal immediately smiled and went up to Akbar. He announced the answer; he said there were twenty-one thousand, five hundred and twenty-three crows in the city. When asked how he knew the answer, Birbal replied, "Ask your men to count the number of crows. If there are more, then the relatives of the crows must be visiting them from nearby cities. If there are fewer, then the crows from our city must be visiting their relatives who live outside the city."<br>Pleased with the answer, Akbar presented Birbal with a ruby and pearl chain." | "During a session in his royal court, Emperor Akbar posed a question that perplexed everyone present. As the courtiers struggled with the response, Birbal entered and inquired about the dilemma. The question presented to him was the total number of crows residing in the city.<br><br>With a knowing smile, Birbal approached Emperor Akbar and declared that the city housed exactly 21,523 crows. When questioned about the accuracy of his count, Birbal confidently responded, "Have your servants tally the crows. Should they find more, it implies the local crows have guests from neighboring areas. If fewer, our crows are likely visiting their kin elsewhere." Impressed by Birbal's clever reasoning, Emperor Akbar rewarded him with a necklace adorned with rubies and pearls." |

**Figure 2: Story Example Wrongly Predicted**

# 3 Method Design

## Overview of the System

We made a system to check text closely using a few important language things like how hard the sentences are, how many different words there are, and how the text flows. Our main aim is to really understand what the text is saying. We're using this cool tool called a Streamlit web app for it. It helps us get feedback on what we find and shows the measurements in an easy way. This setup works well for lots of data projects, making it simple to understand and use text data better.

## Choice of Technologies

**Streamlit:** Streamlit was utilized for our project due its ease of use. Streamlit requires no web development experience to get started. Streamlit is great for setting up AI projects [12] because it's flexible and straightforward. The main reason we chose it is that it lets us see our project metrics update immediately.

**Transformers and PyTorch:** We used GPT-2 model from the Transformers library in PyTorch due to its understanding and processing of language [12]. GPT-2 is especially useful for making sense of how complex texts are, thanks to its ability to analyze text perplexity. We use PyTorch because it's fast at processing, even when we're working with a lot of data, and it handles it all through deep neural networks.

**NLTK:** This tool combines several text processing libraries that are great for natural language tasks. It helps with things like breaking text into words, tagging parts of speech, parsing sentences, classifying text, and identifying common filler words. We mainly use this tool in our projects to prepare text for analysis [12]. It does this by simplifying and standardizing words through processes like stemming and tokenization, making sure texts are consistent before we analyze them.

**Scikit-learn:** Scikit-learn was utilized due to its straightforwardness in turning text into numbers and figuring out how similar texts are. It uses TF-IDF to highlight the important words and cosine similarity to compare them [12]. This allows for quick analysis of data, allowing us to analyze large amounts of texts at once.

## 4 Text Preprocessing

**Regex Operations:** For preprocessing text, regex plays a pivotal role, used for cleaning and standardizing the text. It deletes unnecessary characters and any noise from the text. Unnecessary characters like quotations and non-alphanumeric characters are removed by them for the purpose of accurate results [9]. By performing these filtrations, we could lead to text purification by including only meaningful content. This step ensures the quality of content and helps us to analyze efficiently and easily when compared to others.

**Tokenization and Stopwords Removal:** Tokenization is the process of breaking down text into individual words or tokens. This helps us to break the clean text into words, which makes it easier for analysis. After performing tokenization, we had to remove common words or stopwords which did not contribute much information to the text analysis [9]. By removing stopwords, it reduced the dataset size and helped the system to analyze important words which carry meaningful information. This played a crucial role for us to perform analytical processes such as vectorization, etc.

**Normalization:** It is used for converting all text formats into a format that follows a normal and uniform distribution in analysis. For our system, the first thing we did was to convert all text into lowercase. This helps us to prevent the same words in different cases being counted as distinct [9]. For example, the words "Bike" and "bike" are normalized to be recognized as the same, facilitating consistent and accurate analysis. Overall, it reduces complexity and encourages simplicity by standardizing the text into an equalized format.

## Analytical Techniques

**TF-IDF and Cosine Similarity:** The abbreviation TF-IDF stands for Term Frequency-Inverse Document Frequency. It is used to evaluate the importance of a word in a document relative to a large volume of corpus. It measures how many times a word appears in a document and then normalizes this by the length of the whole document [8]. This overall metric helps to reduce the weight of each term which occurs many times and contributes less information to the document. Cosine similarity measures the degree of similarity between two text contents, excluding their size, by measuring the cosine of the angle between two vectors projected in multidimensional.

**Perplexity:** It measures how well a probability model predicts a sample. According to the GPT-2 model's capability, perplexity gives an idea about the complexity of the text in predicting the next consequent word. To know the perplexity score, you must calculate the exponential of the average negative log-likelihood of word probabilities provided by the GPT-2 model for our content. A lower perplexity indicates that the model is less complex and predicts the text more accurately [8].

**Burstiness:** It describes the variance in the frequency of words in the text. It also provides information about the distribution of words in the text. It calculates the measure by taking the ratio of the proportion of words which appears more than once to the total number of distinct words [10]. Higher burstiness indicates repetitive content and repetition of particular words in the text or individual words.

**Lexical Diversity:** It is calculated by taking the ratio of distinct words to the total number of words in our text. It is used to measure the variety of words in text content. Indirectly, it indicates that higher lexical diversity maintains a broader range of vocabulary [10]. It is more valuable because it assists you with a range of different words and helps us to gain deeper insights into the writer's language usage.

**Sentence Complexity:** It is measured by the average length of sentences. It mainly analyzes the structure of sentences within a document or text. Not only that, but it also employs a variety of grammatical functions. It is calculated by taking the ratio of the number of words to sentences. This gives insights about the syntax of the text [10]. If there are longer sentences, it indicates more complex grammatical models. This analysis helps us to understand the style, details, and quality of the text easily.

## 5 Implementation

### AI Model Overview: From Idea to Application

Our focus of our algorithm to know the metrics of our text and how it is contributed and concluded to copyright issue. The central part of our project is to use GPT-2 model to get the analysis of text content. The selection of GPT-2 allowed us to use it in multiple functions, for example it is very useful in getting linguistic features and metrics. Behind this, each measure selected due to their own individual capability to meet our requirement of copyright detection with the help of GPT-2 model.

In our situation, perplexity serves as a gauge for the complexity and unpredictability of text sequences since it quantifies how well a probability model predicts a sample [8]. Burstiness examines the differences in word frequencies and sentence lengths, offering insights into the stylistic decisions that set human writing apart from texts produced by artificial intelligence.

By calculating the percentage of distinct terms used in the text, Lexical Diversity provides an indicator of the richness of the text's vocabulary.

The average sentence length and the syntactic structures employed are utilized to examine sentence complexity, which aids in determining the degree of sophistication of the generated content [10].

The Similarity Score helps identify possible copyright infringement by comparing the created content's textual parts to a database of original texts [10].

With additional assistance from libraries like NLTK for text processing and Scikit-learn for implementing the TF-IDF vectorization and cosine similarity measurements, these metrics are combined into a streamlined codebase that uses Python for scripting .

**Elements**

The model which we created has different features to adapt to its environment. Our model not only analyzes the standard format text, but it also can adapt to different types of text like different format styles of text. It accomplishes this through the use of our preprocessing text feature. The numeric values which we receive as an output after analyzing text can be checked by cross-checked by humans and can give reviewers insight as to whether the text is human generated or possibly AI. This can help with a hybrid analysis approach for copyright detection. With the help of streamlit interface, we created an interactive platform like website, which shows text boxes where you can enter or paste your text and can get immediate results simultaneously. This can create easy accessibility for the people who don't have any technical knowledge or experience programing as knowledge of the metrics and thresholds. According to the model which we have created and developed we believe it has a direct connection with the project objectives. The model is an effective system in analysis of text and helps reviewers get the results for copyright evaluations in AI content. Our study of models is based on in-depth analysis of text and how similar they are, which is a key component for copyright detection. The novel features like dynamic analysis environment and hybrid analysis confirm that our project system can be modified if there is advancement technology in future of Artificial Intelligence.

**Structure and Functions of the Project Code**

In this part of our project, we look at how our software analyzes text data. We used various methods that focus on language to help us figure out possible copyright infringement. What makes our code unique, how we put it all together, and the steps we took to reach our goals. This will be a review that shows how flexible our creative project can be.

## A Synopsis and structure of our code alignment

The code is divided into below modular parts:

**Module for Text Preprocessing:** This module takes care of the preliminary standardization and cleansing of text data. To guarantee consistency in the analysis, it eliminates unnecessary punctuation and symbols, normalizes whitespace, and changes all text to lowercase [9].

**Metric Calculation Module:** The following metrics are calculated independently: sentence complexity, lexical diversity, similarity score, burstiness, and perplexity. Although these functions are meant to be called separately, they can also be combined to offer a thorough examination of text data [11] .

**Analysis Engine:** This central component is responsible for coordinating the preparation module and metric computations when processing text data. To find out if the text demonstrates traits common to AI-generated content, it compiles the results and assesses them against preset thresholds [12].

**User Interface:** Streamlit-built, the interface lets users enter text and see the analysis findings instantly. It offers an intuitive interface for engaging with the system and shows interpretive insights in addition to basic metric statistics.

Principal Roles and Their Innovative Methods

**Preprocess Text:** To make sure that the input to the metric calculators is free of any artifacts that can distort the analysis, this function uses regular expressions to clean the text. Its adaptability allows it to take in numerous forms of textual input from different sources. It also removes words in quotes so that they will not be included in similarity analysis and give false positive results.

**Calculate Perplexity:** This function evaluates the predictability and complexity of text by computing its perplexity using the GPT-2 model. Here, a pretrained model is used in a novel way—it is explicitly customized to evaluate the subtleties of potentially AI-generated content, rather than just general language models [8].

**Determine Burstiness:** This function assesses the variation in the text's word frequency and sentence length. This novel method goes beyond conventional frequency analysis to comprehend the subtle stylistic differences between writing by humans and artificial intelligence [10].

**Determine Lexical Diversity:** This function provides information on the richness of the text's vocabulary by calculating the ratio of unique words to total words. This solution is distinct in that it applies to texts written by AI as well as humans, offering a direct comparison that is essential for copyright analysis [10].

**Calculate Sentence Complexity:** This function examines the length and structure of sentences to provide a syntactic complexity indicator, which is frequently indicative of the authorship level of the content, whether it be AI or human [10].

**Compute Similarity Score:** This function checks the input text against a database of original texts to find any overlaps using cosine similarity and TF-IDF vectorization. The process is adjusted to particularly draw attention to resemblances. We've put together our system by combining strong backend analytics with a straightforward frontend interface, using Python for the backend and Streamlit for the frontend. We chose Streamlit because it's user-friendly and gives instant feedback, making it easier for people who aren't tech experts to use our system. In our setup, we've added several smart features into a cohesive system that does more than just detect copyright issues it sets a new standard for how text analysis can be used to protect intellectual property today. This approach shows careful thinking we put into our work. It demonstrates we really understand both the latest tech and the tricky parts of copyright law, making our project stand out in the field of text analysis.

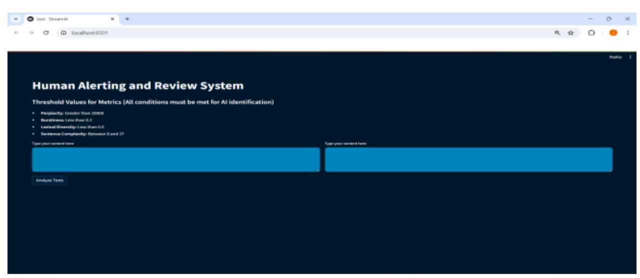


**Figure3: Initial setup without any user input all thresholds listed as well as an analyze text button for ease of use**

## The AI Content Generation Prompt's Design

We decided to go with an easily readable result when we came up with the prompt for our AI to generate content, making sure we stay original and on the right side of copyright laws. The prompt we set up is super important as it shapes the quality and features of what the AI comes up with.

**The Prompt Design's Foundations and Research of our prompts.**

We came up with a prompt that strikes a nice balance. On the other hand, we stick close to the original themes and ideas, which helps us avoid any copyright problems. Basically, we're making sure that what the AI churns out is not only unique, but also respects the vibe of the original material according to creating our prompts.

**Innovative Standards for the Prompt Thematic Coherence**: The prompt which we created is going to ask AI to maintain the original content and theme of story, so it is going to be checked with copyright issue. In this way, the AI can detect the moral theme and underlying background of story which is very important for our project. It also maintains the richness of text while generating AI content.

**Linguistic Transformation:** Instead of just tweaking a few words here and there, the prompt really mixes things up. It pushes for new word choices and different ways of putting sentences together. This kind of nudge gets the AI to think outside the box and explore new ways to put things. What you end up with is a text that stands out from the original but still keeps the main ideas. It's like giving the text a whole new personality while keeping its originality.

**Assurance of Originality:** So there's this rule that says whatever you tweak or make should be different enough to count as something unique. This sets a clear standard for being original. It pushes the LLM to get creative and come up with new stuff instead of just copying what's already out there. This way, we make sure that what the AI comes up with follows the rules of copyright.

**Directive Clarity:** The prompt which we used is clear and direct. The prompt which we give to AI tried to generate almost the same results every time without giving any wrong information by maintaining the quality of content. It generated straight forward answers when we ask AI to generate content using prompts.

**Flexibility for Creative Expression:** This prompt can make AI show what its capabilities are even when certain rules are provided in prompt. It also gives rules about how to make sentences too. By mentioning these rules promptly, the AI can adapt to dynamic changes, and it shows what can it do. In this way, the AI can give information or text content according to original story by maintain protocols.

**Feedback Mechanism:** The prompt was made very clear and concise. The way it is made is that it uses feedback loop like it tries to make AI content from original story by following rules of prompt. When we try next time, it tries to get better and so on. This way AI helps itself to give better results compared to before result by maintaining the theme and moral of the original story.

**A Novel Approach in Prompt Design:** We wanted AI to generate creatively, kind of like how a human artist might rethink a concept, instead of just giving commands. We shaped our prompt to balance innovation and creation, with a careful emphasis on copyright rules. Unlike usual prompts that just spit out content, our method was totally made by us, based on our unique ideas about how to use AI creatively while still following the rules we set.

In short, the prompt we used was, "Can you rewrite this story with new words and sentence structures while keeping the same meaning and themes?" When we talk about making sure our new version stands out, we're really aiming for a mix of being original, creative, and following the rules. It's really important that the updated content isn't just a redo; it should feel like its own thing, totally unique.

## Techniques for Evaluating AI-Generated Texts and Original

Our study uses a unique method to compare stories made by AI with the original ones in a clear way. To see how well the AI follows the creative instructions from the prompt while making the text different enough to be seen as original, we need to do a comparison analysis. We talk mainly about the tools and methods we use to get useful insights about the quality and originality of the text.

**Comparison**

In our setup, we use a two-text box interface for comparing texts. One box shows the original story, and the other shows the text that the AI created. This side-by-side layout lets us see the differences right away, which helps us do a quick quality check before we dive into a more detailed analysis. You can type in texts, see them, and check them out easily in real time using this user-friendly interface, which we made with Streamlit for interaction for users.

**Metrcis for comparison**: We used a collection of measures that were particularly selected for their capacity to draw attention to various facets of text quality and uniqueness in order to quantitatively evaluate the texts, perplexity, and lexical diversity are observed when AI preserves or even improves the language richness of the original by contrasting the lexical diversity of the AI-generated stories with that of the original work [10]. This is a sign of creative creation on the rise.

Sentence Complexity a determining factor as to whether the AI-generated text matches or surpasses the original's syntactic depth requires analyzing the intricacy of the sentence structures in the two texts [10]. This measure is essential for assessing how well AI is able to copy or develop sentence construction that is similar to that of humans.

Similarity Score was another metric that evaluates how closely the content of the AI-generated text resembles the original while maintaining enough uniqueness to qualify as an original work. It does this by using cosine similarity [11]. The evaluation of originality and conformity with copyright depend on this balance.

**Analytics(Methodology)**

The special thing about our process is how we mix and check these metrics together. Instead of looking at each number on its own, our method brings them all together to give a full view of what the text is like. In this way, we have a chance to find new patterns and themes. It shows what it is real capability and simultaneously to see similarities. This setup helps us see similarities and differences in texts and understand better how the AI changes and handles story information. This method is important for our research because it helps us reach our big goal: making AI that can come up with new, while following copyright rules. By checking all metrics, we will come to know how the text is generated whether it is AI or Human written. If its AI is written with the four linguistic metrics, then we can check similarity score whether it is copied or not. By this we can detect copyright violation. In conclusion, our project's way of comparing writings from humans and AI really focuses on deep analysis. This method makes sure we look at every text closely and in detail, which helps us learn more about what AI can do in making content.

## Technique to Modify Threshold Values

The four metric threshold values are a central part of our project. By checking the thresholds, we could attempt to detect copyright infringement by maintaining the originality of content.

**First(Threshold)**

The ideal threshold values we took based on research we did on AI and the way it is generating content:

**Perplexity:** Set higher than 20,000 to make sure the language is sufficiently complicated and unpredictable, proving the AI isn't just copying well-known sentences [8] or readily recognizable text patterns.

**Burstiness:** If set less than 0.2, the text's variance will be more uniform and suggest a constant quality rather than simplicity[10].

**Lexical Diversity:** Lower than 0.5 means that more complex language use is being encouraged and the text is not [10] just restating a small vocabulary and there is chance of complex.

**Sentence Complexity**: Aim for a number between 8 and 27 that strikes a balance between unduly complicated and overly simple sentence patterns [10], thereby exhibiting a fluid and interesting writing style.

**Process of Manual Adjustment:** We improved our thresholds by looking at the numbers from both the original and AI-written texts. Firstly, we collected a bunch of data in the form of stories. We used our model with two boxes to pull together a lot of pairs of texts, one original and one made by AI. This big pile of data helped us figure out a starting point for each of our measurements.

**Analysis of Metric Values:** After setting up the baseline, we checked cases where the quality and originality of the text from AI matched or even outdid the original. We used scatter plots and histograms based on the metrics data for this analysis, which helped show clearly how the texts differed across different criteria. With analysis of metric values, we find out the average value for perplexity to know how much difference between human generated and AI content. We can know basic difference with this initial threshold values**.**

**Threshold Refinement:** we adjusted the initial threshold values to keep the balance between creation and complexity. So, we can take into account what insights we got from initial experiments of original AI generated stories. Here, we can explain it with an example that if perplexity is raised, it could lead to more complexity. Indirectly, it could get more results in false positives domain.

**Iterative Feedback and Finalization:** The usage of feedback mechanism helps us to adjust threshold values more efficiently. To get accurate threshold values, we increase our number of threshold values to 50. With this we got to know unique qualities and abilities of our AI model and how it is going to give metric values according to AI and human generated content. This is due to our iterative and feedback loop.

**Innovative Justifications for Choices:** Our methods are new because we used a lot of data analysis and kept improving it as we went through iteratively. Unlike some systems that just stick to rules they set at the beginning; our way lets us change things based on what's actually happening with AI-generated content. This helps our detection system get better over time and stay up to date with the latest changes in how texts are created by AI.

Our way of adjusting the thresholds by hand is tailored and based on real data and our experience running our model. We made it to make sure that the AI texts we produce meet high standards for quality and uniqueness and are good enough for detailed copyright checks. It also led to discoveries into the way AI writes and its particular average thresholds. Which allows for better discernment when deciding if content is AI or human generated.

**Establishing Copyright Status Through Integrated Metrics Analysis**

In our project, we used language measurements like similarity scores, perplexity, burstiness, lexical variety, and how complex sentences are. We are going to combine the analysis of our metric values to know whether someone might copy the work. This multi-level approach is important because it helps us really understand how unique the text is and if there might be any copyright problems. This approach allows for more accurate detection of potential copyright infringements, and we can also know Indepth analysis of text. These texts can be further reviewed by professional copyright reviewers if we would like to know about more copyright rules and regulations. With that, we can also implement a future in which copyright laws are potentially violated.

## A Comprehensive Method for Metric Analysis

Our method analyzes text through a framework that looks at different:

**Evaluation by Metric:** we will start checking whether the text is AI generated or human generated with threshold values. This involves looking at all four metric or linguistic feature values. This tells how well the text is there, we will come to know what to change in threshold values of our metrics

**Evaluation of Similarity Score:** This similarity score helps us to how much similarity is there between AI content and our copyrighted text. This also tells us that AI generated content generated by prompt is similar to already written story.

**Threshold-Based Analysis:** We start by setting initial threshold values for measuring with similarity score also. If any text does not meet the requirements in burstiness and lexical diversity or crossing them in perplexity and similarity, the text needs more examination with some more experiments.

## Finalizing the Project's Threshold Values

Our journey to fine-tune the AI content identification system involved lots of detailed and repeated analysis of data from both original and AI-created texts to nail down the final threshold settings. We aimed to pick thresholds that could reliably tell apart AI-generated content from stuff written by people, based on their unique traits.

Setting the Initial Threshold and Collecting Data

At the start, we decided on some threshold values based on what we initially thought made AI-written stuff different from human-written text:

Perplexity: We set it over 20,000 because we figured that AI-written texts are usually more predictable.

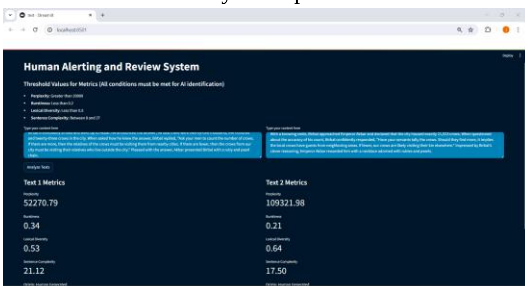


**Figure 4: Initial Result of basic threshold values**

Lexical Diversity: Set to under half, this was based on the idea that AI-created texts probably use a narrower range of words.

Sentence Complexity: We started at 8 and went up to 27, trying to cover a broad spectrum of sentence types found in both types of writing.

**Analytical Procedure using Gathered Information**

We collected and checked loads of texts written by people and AI using a system with two boxes for text. To spot any patterns or odd bits, we lined each text up against the standards we set at the start, and used dot plots and bar graphs to show what we found. Look at figure 4 above for a story we marked wrong by just a little bit. It was just a tiny bit over our lowest burstiness number by 0.01. We think our numbers are mostly right for spotting AI stuff, but sometimes they mess up because the AI can rewrite really well.

Here's what we noticed from checking all the info:

- A lot of AI writings had really high perplexity—way over 25,000 at times. This led us to think maybe our original number was too low to catch all the complex text that AIs can come up with.

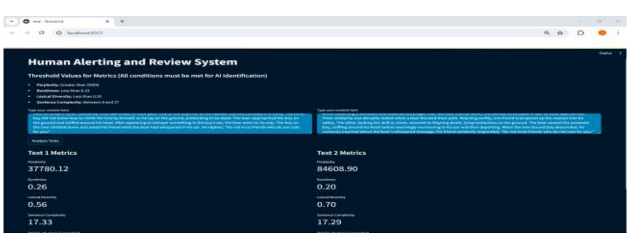


**Figure 5: Results of Intermediate Threshold values**

- The burstiness numbers for AI writings were all over the place more than we thought; a lot of them were around 0.23. We need to tweak this number to better catch the small differences.
- AIs were also more varied in their word use than we expected. Sometimes, their lexical diversity hit 0.55, making us think again.
- The way AIs build sentences often sat right between the highest and lowest numbers we first thought of. This has made us think about narrowing this range to spot AI traits better.

**Modification of Final Thresholds:** Adjusting the final thresholds: We've seen the need to tweak our thresholds over time.

-We've set the complexity threshold to more than 30,000 to handle the more complex stuff AI writes. This final threshold value helps us very much to predict the difference between AI and human generated

-Burstiness got adjusted to less than 0.25. This better shows the ups and downs in how the text flows and organizes. This value lets know how much variation of text is there in both content. For

-Lexical Diversity, we brought it down to less than 0.71. This change considers the bigger vocabulary that newer AI models use. This value indicates that broader range of vocabulary intensities. -

-We narrowed down Sentence Complexity to between 15 and 25. This gives us a clearer line for checking how complex sentences are and how the formation of sentences is there. With each of these changes, we did more data collecting and analysis. We wanted to make sure these new thresholds really show the difference between what humans and AI write.

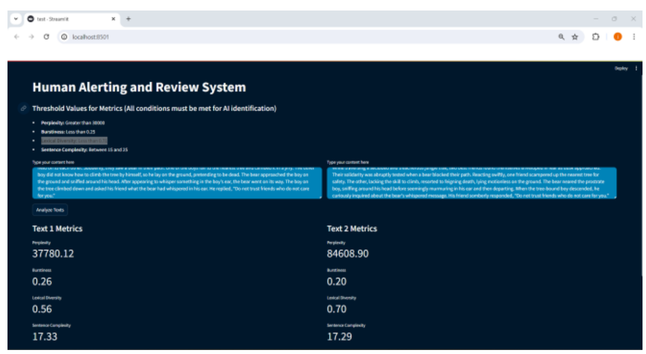


**Figure 6: Final Threshold value Results**

**Justifications for Choosing Particular Measures**

These metrics were picked because they help us understand the text's nature well: Perplexity shows how complex and unpredictable the text from the model can be. Burstiness helps us get how the text flows, which might tell us if a computer generated it. Phonological Variability and Sentence Complexity measure how well-structured and rich the text is things that AI often tweaks to make it seem like a human wrote it. In the end, choosing threshold values was key to making our AI detection system more reliable. This process shows our commitment to building a dependable and flexible copyright assessment system, since it was based on real data and deep analysis of text features. By using a careful step-by-step method, we've set cutoff points that work well in practice and are scientifically backed to tell apart content made by AI and Human generated text.

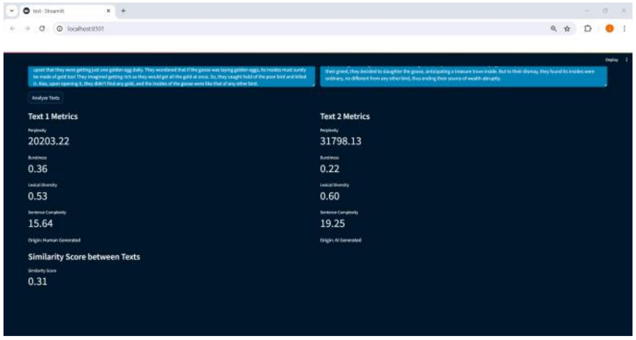


**Figure 7: Similarity Score Results**

## Entire Evaluation of Metric Values in Texts OF AI

In contrast to original human-written content, our team has really looked into and checked different language measurements from AI-written texts. We've mainly looked at how similar they are, burstiness, how varied the words are, how complex the sentences are, and perplexity scores. These measurements are important for figuring out how real and unique AI-created texts are, which is a big deal for the copyright issues with generative models.

**Perplexity Perspectives**

When we compare how AI and people write, one cool thing to look at is how predictable their writing is. We use a measure called perplexity for this, and it's pretty revealing. AI texts usually have a higher perplexity score—about 58,028 on average—while human texts are lower, around 34,860. What this means is that AI tends to be more complex and less predictable. Because AI's writing is more all over the place, it's great at cooking up something totally new. It doesn't stick to the usual ways we expect writing to go, which makes it less likely to sound like something you've read before.

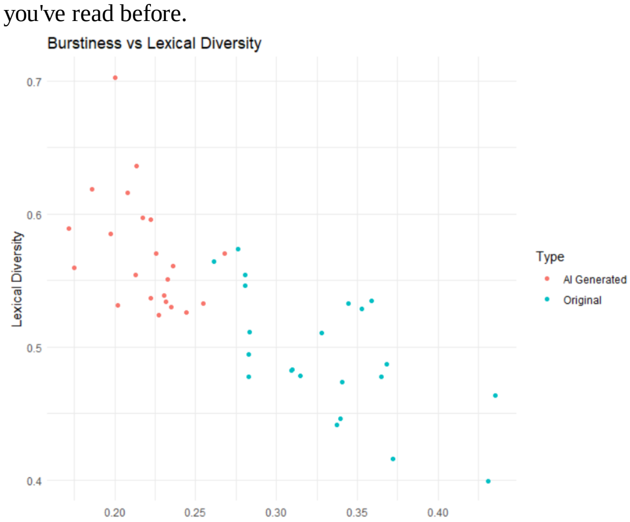


**Figure 8: Burstiness vs Lexical Diversity**

**Lexical Diversity vs. Burstiness**

When we look at traditional writing, there's a cool connection between Burstiness and lexical diveristy used. Generally, the more consistent the sentence lengths, the richer the vocabulary. It's like human writing tends to flow more smoothly when there's a good mix of words. AI writing doesn't quite follow this pattern. It throws in a mix of word choices and sentence lengths that can seem a bit out there. This shows how AI can get really creative, bending the usual rules we're used to in writing. It's like AI has its own unique style, making its creations stand out from what we typically see in human writing.

**Analysis of Complexity and Similarity**

The test we used to compare how complex sentences are in AI-written versus human-written text didn't find a big difference. Basically, the numbers showed that both AI and humans use sentences that are about equally complex. For AI, the average complexity score was 19.42, and for humans, it was 18.24. This means AI can mimic human grammar really well and can write with almost the same level of complexity. The model also gives similarity score on the interactive platform to show how similar both original text and AI text content. The similarity score which we got from our analysis is below 0.573, this tells us that AI generating content with the prompt is unique and less likely to copy from original. This lower similarity score indicates that there is low risk of copying from original source and it is more unique

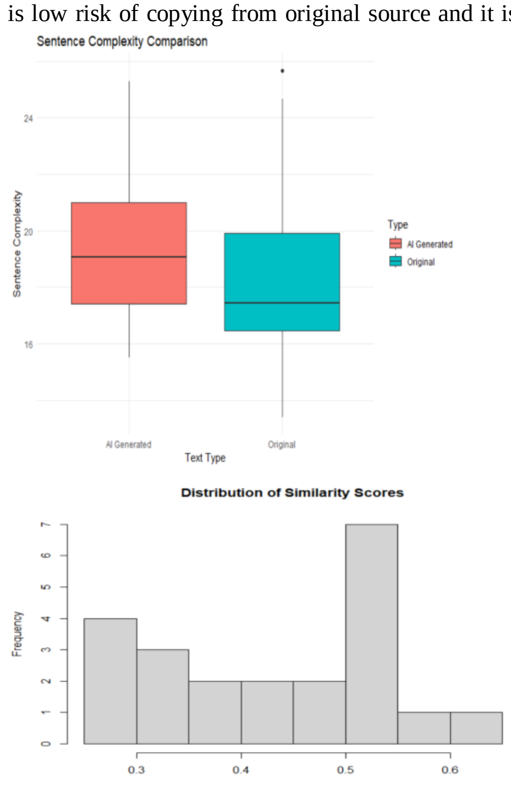


**Figure 9 and 10: Sentence complexity and Similarity Score Distribution**

**Correlation Analysis to Determine Copyright**
After looking into correlations, it turns out that just checking how complex a sentence is might not be enough to tell if it's original. Researchers found that the link between using lots of different words (lexical diversity) and being unique isn't very strong. They also noticed that longer, fancier sentences usually use a wider range of words, showing how these things are all connected. What does this mean? Well, it suggests that while complexity helps to measure, we can't depend on it alone to decide if something's original enough for copyright stuff. It's like saying there's more to a cake than just the icing. To really figure out how unique and copyright-friendly AI-generated content is, we need to look at a bunch of different signs and not just one. It's like putting together puzzle pieces to see the bigger picture.

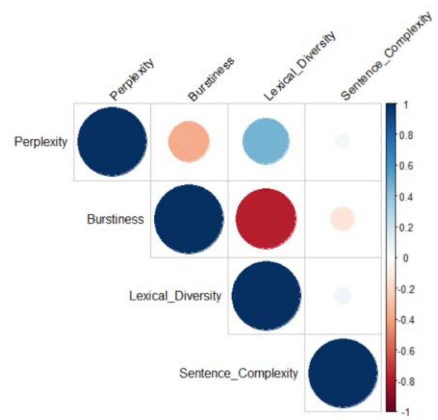


**Figure 11: Correlation of all Metrics**

**Final Thoughts**
So, our project found out that AI can write just like people, not only by being complex and correct in grammar but also by using unique ways of saying things. This is really important for updating how we check copyrights, especially since AI writing is becoming more common. By mixing number-based analysis with a good look at the quality of the text, we've come up with a solid method to see if AI-made texts are truly original. This helps us figure out better what counts as original content in our digital world today.

## 6 Technical Evaluation of our whole prject

Our model was tested using stories created by GPT-3.5 based on original human written input stories. We then tested the human written stories and stories created by GPT-3.5 to detect if they were AI created as well as their level of similarity. The model was tuned using turn-it in as a reference of accuracy. Our model was 77% accurate at discerning between the original stories and the AI generated content of GPT-3.5. The prompt we used asked for unique stories separate from the initial content. Although we asked for uniqueness the average cosine similarity was 0.573 which was higher than our threshold for originality of 0.47.

Due to the high similarity, we do not believe GPT-3.5 with limited prompting can create unique work that passes our threshold. We also tried to increase the areas in which some AI created stories failed to pass the threshold by prompting increased perplexity for some and decreased burstiness for others. Despite this additional prompting the stories did not pass the threshold.
Our model could be used in instances where a person believes their work was infringed on and could be used to determine whether AI may have been used to copy and its similarity to copy-written material. We also added a line of code to not include quoted material for instances where quotes are heavily used such as scientific abstracts. Due to the nature of some text needing to be confidential we chose cosine similarity as it has been stated that “Cosine with other similarity metrics are efficient for plagiarism detection in secured systems in which submissions are considered confidential, such as conferences.” (4)

## 7 Conclusion

We used GPT-3.5 to recreate children's stories and then analyzed using the provided code and thresholds we refined, it was observed that the detection model achieved impressive results. With no false positives and an accuracy rate of 77% in identifying AI-generated content, the system demonstrated robustness in distinguishing between human-written and AI-generated texts. Our results were surprising considering in other research papers “AI detectors performed well in accurately recognizing human-written essays. In contrast, they performed poorly in detecting ChatGPT-generated and enhanced essays” (8).
The average similarity score of 0.573 indicates a significant similarity in the AI-generated and human made input content, further corroborating the idea that GPT-3.5 with minimal prompting does not create what we define as unique content. These findings suggest that the approach implemented in our model can serve as a reliable tool for detecting generative AI content created by GPT-3.5, especially in scenarios involving copyright infringement or content authenticity verification.
It is crucial to acknowledge the potential misuse of generative AI in creating copywritten material. While our model shows promise in identifying AI-generated content produced by GPT-3.5, there remains a possibility of more sophisticated AI models even the newer GPT–4 could produce content that more closely resembles human writing, making it challenging to detect. This underscores the importance of continual refinement and enhancement of detection algorithms to stay ahead of evolving generative AI capabilities.
In conclusion, our design science study demonstrates the effectiveness of our model in identifying AI-generated content with high accuracy and no false positives. There must be ongoing vigilance and innovation to address the challenges posed by the potential misuse of generative AI in creating copywritten material in the future. If there are tools that can be made to combat this problem maybe a large language model can fix the problem, it is causing.

## 8 REFERENCES

[1] Chaka C. 2023a. Generative AI chatbots - ChatGPT versus YouChat versus Chatsonic: Use cases of selected areas of applied English language studies.

International Journal of Learning, Teaching and Educational Research, 22(6), 1-19. https://doi.org/10.26803/ijlter.22.6.1

[2] Eliaçık E. 2023a. Best ChatGPT alternatives: What awaits us beyond ChatGPT? DataConomy. Available at https://dataconomy.com/2023/02/beyond-chatgpt-alternatives-best-jasper/

[3] Hetler A. 2023. Bard vs. ChatGPT: What's the difference? TechTarget. Available at https://www.techtarget.com/whatis/feature/Bard-vs-ChatGPT-Whats-the-difference

[4] Kamran S. 2023. Microsoft Bing vs Google Bard: ChatGPT cloning battle has begun. Available at https://techcommunity.microsoft.com/t5/itops-talk/microsoft-bing-vs-google-bard-chatgpt-cloning-battle-has-begun/m-p/3737979

[5] Chaka, C. 2023. Editorial: Journal of Applied Learning & Teaching. Journal of Applied Learning & Teaching, 6, 2. University of South Africa, Pretoria, South Africa. ISSN: 2591-801X. Available at http://journals.sfu.ca/jalt/index.php/jalt/index.

[6] Gao C. A., Howard F. M., Markov N. S., Dyer E. C., Ramesh S., Luo Y., Pearson A. T. 2022. Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. bioRxiv. https://doi.org/10.1101/2022.12.23.521610

[7] Author Unknown. n.d. Top Short Moral Stories for Children. FirstCry Parenting. Retrieved from https://parenting.firstcry.com/articles/top-20-short-moral-stories-for-children/

[8] Gamallo P., Pichel J. R., Alegria I. 2017. A Perplexity-Based Method for Similar Languages Discrimination. In P. Nakov, M. Zampieri, N. Ljubešić, J. Tiedemann, S. Malmasi, & A. Ali (Eds.), Proceedings of the Fourth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial), Association for Computational Linguistics, 109–114. https://doi.org/10.18653/v1/W17-1213

[9] Van Leijenhorst D.C., van der Maas H.L.J. 2013. The Functional Segregation and Integration Model in Complex Cognitive Systems. Procedia Computer Science, 18, 449-458. Available at https://www.sciencedirect.com/science/article/pii/S1877050913001385

[10] Zhao M., Bryant P. 2023. Comparing Measures of Syntactic and Lexical Complexity in Artificial Intelligence and L2 Human-Generated Argumentative Essays. Available at https://www.researchgate.net/publication/377029323_Comparing_Measures_of_Syntactic_and_Lexical_Complexity_in_Artificial_Intelligence_and_L2_Human-Generated_Argumentative_Essays

[11] Alzahrani S. M., Salim N., Abraham A. 2012. Understanding Plagiarism Linguistic Patterns, Textual Features, and Detection Methods. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42(2), 133-149. https://doi.org/10.1109/TSMCC.2011.2134847

[12] ScienceDirect. (n.d.). The impact of text pre-processing on document clustering. *Knowledge-Based Systems*, 52, 178-189. https://www.sciencedirect.com/science/article/pii/S0927025623006535

[13] Vyas N., Kakade S. M., Barak B. 2023. On Provable Copyright Protection for Generative Models. Proceedings of the 40th International Conference on Machine Learning, in Proceedings of Machine Learning Research 202:35277-35299. Available from https://proceedings.mlr.press/v202/vyas23b.html.

[11] Lucchi N. 2023. ChatGPT: A Case Study on Copyright Challenges for Generative Artificial Intelligence Systems. European Journal of Risk Regulation, 1–23. doi:10.1017/err.2023.59

[15] Hayes C. M. 2023. Generative Artificial Intelligence and Copyright: Both Sides of the Black Box. SSRN. Available at https://ssrn.com/abstract=4517799 or http://dx.doi.org/10.2139/ssrn.4517799