

Classic Brazilian Books Translated into Images with Generative Artificial Intelligence

*Libros clásicos brasileños traducidos a imágenes mediante
inteligencia artificial generativa*

DOI: 10.61820/ha.2954-470X.1655

Felipe Rodrigues Perche Mahlow

Universidade Estadual Paulista “Júlio de Mesquita Filho”

Bauru, Brasil

f.mahlow@unesp.br

ORCID: 0000-0001-9816-1440

André Felipe Zanella

Universidade Estadual de Maringá

Maringá, Brasil

aft.zanella@gmail.com

ORCID: 0000-0001-8828-6629

Universidad Autónoma de Querétaro
Licencia Creative Commons Attribution - NonComercial ShareAlike 4.0
Internacional (CC BY-NC-SA 4.0)



Classic Brazilian Books Are Translated into Images with Generative Artificial Intelligence

Libros clásicos brasileños se traducen en imágenes mediante inteligencia artificial generativa

DOI: 10.61820/ha.2954-470X.1655

William Alberto Cruz Castañeda

Universidade Tecnológica Federal do Paraná

Guarapuava, Brasil

wcastaneda@utfpr.edu.br

ORCID: 0000-0002-9803-1387

Regilene Aparecida Sarzi Ribeiro

Universidade Estadual Paulista “Júlio de Mesquita Filho”

Bauru, Brasil

regilene.sarzi@unesp.br

ORCID: 0000-0001-6267-6549

Recibido: 03/08/2024

Aceptado: 08/04/2025

Universidad Autónoma de Querétaro
Licencia Creative Commons Attribution - NonCommercial ShareAlike 4.0
Internacional (CC BY-NC-SA 4.0)



Abstract

This article describes the application of open-source generative artificial intelligence (Gen-AI) tools to support the illustrative process of books of classical Brazilian literature. The article relates an experience with state-of-the-art models and the central question addressed is the feasibility of these tools, as well as assessing the positive and negative behaviors observed in the models. The study implemented the Stable Diffusion XL Base 1.0 and Stable Diffusion XL Refiner 1.0 models. The methodological process used in this research highlights the formal elements of composition, color, and texture, along with ethical considerations, to assess the effectiveness and biases of the illustrations generated by Gen-AI. The results demonstrate the potential and spotlight the challenges of using Gen-AI as a tool for book illustration, and provide valuable insights into the intersection between literature and artificial intelligence.

Keywords: generative artificial intelligence, illustration, Brazilian literature

Resumen

Este artículo describe la aplicación de herramientas de inteligencia artificial generativa (IA-G) de código abierto para apoyar el proceso ilustrativo de libros de literatura clásica brasileña. El artículo retoma una experiencia sobre el uso de un modelo de última generación y la pregunta central abordada es la viabilidad de dichas herramientas, así como evaluar los comportamientos positivos y negativos observados en el modelo. El estudio implementó los modelos Stable Diffusion XL Base 1.0 y Stable Diffusion XL Refiner 1.0. El proceso metodológico utilizado en esta investigación resalta los elementos formales de composición, color y textura, junto con consideraciones éticas, con la finalidad de evaluar la efectividad y sesgos de las ilustraciones generadas por la IA-G. Los resultados demostraron su potencial, a la vez que resaltan los desafíos del uso de IA-G como una herramienta de apoyo a la ilustración de libros, aportando, de esta forma, información valiosa sobre la intersección entre la literatura y la inteligencia artificial.

Palabras clave: inteligencia artificial generativa, ilustración, literatura brasileña

Introduction

Generative artificial intelligence (Gen AI) is profoundly transforming society by integrating capabilities in the generation of text, audio, video and images, excelling in the creation of synthetic data that imitate real-world phenomena. Text generation models like OpenAI's GPT (Achiam *et al.*, 2023) improve natural language processing, content creation, and automated writing by demonstrating exceptional context and consistency (Roumeliotis and Tselikas, 2023). In terms of visual creation, models such as DALL-E (Betker *et al.*, 2023), MidJourney and Stable Diffusion (Rombach *et al.*, 2022) have the ability to generate complex and detailed images from textual descriptions, opening up new possibilities of artistic expression. However, creating visually appealing and accurate images from literary descriptions remains a complex task. The quality and accuracy of these generated images depend largely on the accuracy of the indications and the performance of the models used.

The current literature offers limited exploration of the convergence between artificial intelligence (AI) and book illustration. Historically, the illustration of literary works requires the participation of artists whose competence in creating images involves a process known to be slow, complex and highly elaborate, as well as subjective. However, recent advances in AI offer automation possibilities that can contribute to this process, producing illustrations that reflect the essence and historical context of literary works with far fewer resources and time compared to traditional methods. These advances allow new forms of creation and innovation. However, along with these opportunities come significant concerns and challenges that must be addressed. It is about having at hand a tool that can contribute to the artist's creative process and expand their aesthetic potential. However, ethical issues, such as the potential for bias in the data generated and impacts on artistic authenticity, require careful reflection (Walczak & Cellary, 2023). Training these models is particularly problematic as it often involves the use of copyrighted images without the explicit consent of the original artists (Piskopani *et al.*, 2023).

The production of images through Gen-AI tools demands a critical reflection on the semiotic dimensions embedded in the creative process. The distinction between denotation and connotation —core concepts in visual language theory— is essential to understanding the boundaries between human creation and

machine-generated imagery. While denotation refers to the literal and immediate content of an image, connotation encompasses the symbolic, cultural, and subjective meanings it carries. In the context of literary illustration, connotation plays a central role, as it enables the illustrator to engage in a sensitive dialogue with the text and the reader's cultural repertoire. Gen-AI tools, operating largely through pattern recognition trained on massive datasets, tend to prioritize denotative elements, often failing to capture the nuanced, historically constructed connotative layers. This limitation becomes particularly relevant when considered within the framework of an ultra-mediated society, a condition in which technological mediation permeates all aspects of experience and shapes how we interpret images, texts, and narratives. In this context, the use of Gen-AI for illustration must be understood as part of a complex media ecology, in which meaning is constructed through a network of algorithms, cultural frameworks, and cognitively mediated technological processes.

Given the scarcity of material in the literature that discusses the topic, this article consists of an experience report on the creation of a set of images through the application of Gen AI tools, with the aim of illustrating well-known texts from Brazilian literature. The goal was to answer the following central question: is it possible to illustrate a book using Gen AI? This question led to others, as pertinent as the first. What formal and content criteria must be observed so that images of this nature can be considered illustrations that truly provoke a dialogue with the literary text? Is it possible that images of this nature result in ethical biases and all kinds of stereotypes? To achieve this, the study searched for the following texts of great relevance in Brazilian literature, available in the public domain: *Senhora* (1875) by José de Alencar; *O Cortiço* (1890) by Aluísio Azevedo; *A Viúva Simões* (1897) by Júlia Lopes de Almeida; *Dom Casmurro* (1899) by Machado de Assis; *Horto* (1900) by Auta de Souza; *Os Sertões* (1902) by Euclides da Cunha and *Triste Fim de Policarpo Quaresma* (1915) by Lima Barreto.

After the generation of images and a wide selection of those generated that could be considered for research, curation was carried out with the purpose of observing, above all, the degree of dialogue between them and the text to configure a defined corpus of images, in this context as illustrations. By illustration we understand those images that inform and help the understanding of the text, that fulfill the function of communicating and contributing to the construction of meanings. An illustration does not aim to reproduce a text in an image, but rather

to interpret the verbal text to “provide the most appropriate elements to the visual, in order to help the reader in the construction of meaning and the relationship between the visual and verbal text” (Giordani & Panek, 2020, p. 59716).

During the curation of the images, different elements were observed that indicate results considered by the study as appropriate to compose this set of illustrations and others that did not meet the criteria set for the final objective of the study: the text-image relationship that, ultimately, you might consider defining what constitutes an illustration. Among the set of formal elements of the image, the curator listed the composition, which includes shape, color and texture for the study of the images and, as for the content, the ethical bias revealed relevant results that require registration and reflection. A fourth category was observed and considered at the time of curating the images as those that did not meet the objective of the study and the possible reasons for such results were discussed.

I. Methods

In order to make these illustrations, we adopt a systematic procedure based on the two main stages of image generation and refinement. In the first stage, we use the Stable Diffusion XL Base 1.0 model to generate initial images. This model was powered with an extensive list of prompts. Each one describes a specific dinner, including details about characters, events and atmospheres. For each prompt, or model, the images were generated in forty steps of inference. The inference process involves “denoising” (noise removal), an essential technique in diffusion models that gradually refines an image of an initial noise state into a clear and detailed representation. In the second stage, we use the Stable Diffusion XL Refiner 1.0 model to refine the images generated in the previous stage. This refinement model receives latent images (partially processed images) and continues the denoising process. This additional phase of forty inference steps is intended to correct imperfections and add fine adjustments that make the illustrations more vivid and consistent with the original descriptions. For each prompt, multiple samples are generated, totaling three hundred images per description. This large-scale generation process guarantees a wide variety of illustrations, allowing a rigorous selection of the best representations for the project.

It is crucial to highlight that the effectiveness of the process depends on the model used. The skill of the individual who develops the prompt plays a fundamental role in the quality of the results obtained. Simply “copying and pasting” portions of text often result in images that do not adequately correspond to the intended context. To achieve an accurate and meaningful visual representation, it is necessary that the artist or prompt describer have the ability to detail the image composition. This includes consideration of shape, color, texture and other essential visual elements that reflect the intended context and narrative. A careful and creative description of the prompt allows IA-G to understand and translate more effectively the textual content in images that really correspond to the vision and needs of the project. All the prompts mentioned are written in English in the process of generating the images and are found translated in this work.

The methodology, based on an exploratory study of a qualitative and descriptive nature, defined four criteria that guided the curation of images in the experience reported in this article. The first criterion is the text-image relationship, where images that meet the objective of the study were selected. The second criterion is composition, which covers the shape, color and texture of the images. The images are analyzed for their formal coherence and the relationship between this coherence and the content to be generated to illustrate a scene or excerpt from the book being illustrated. The third criterion is the ethical bias, which considers images that carry stereotypes.

When these biases are noticed, they need to be corrected to ensure ethical use of the machine. The fourth criterion involves images that did not meet the objective of the study. Instead of being discarded, these images present content that can be explored as counter-models or images considered deficient in this context, offering learning about their flaws. Below is a description of the results of the image curation that selected, for each of the mentioned criteria, a set of five images. These images, which constitute the research corpus, will be described and analyzed aiming for a critical reflection that points to a human-centered use of AI.

II. Results

Below, we highlight a set of images and the respective analyzes carried out using the relationship with the prompt as a criterion, organized based on the four categories presented above, namely: A) text-image relationship; B) composition; C) ethical bias; D) images that did not meet the objective of the study.

A) Text-Image Relationship

During the image curation, we observed that a set of results could be considered to have been efficient in translating the text prompt into an image, since the representation of the text in the form of an illustration came very close to the objective of the study to illustrate literature texts using AI tools (Figure 1).

Figure 1. Examples of generated images for (a) *Senhora*, (b, d) *Dom Casmurro*, (c) *O Cortiço* and (e) *A Viúva Simões* (2024), author's elaboration.



This is the case with the set of images present in Figure 1, which will be discussed individually below. It is noteworthy that the items to be analyzed—text-image relationship, composition, ethical bias and images that did not meet the objectives—may overlap, so that we will present, for example, elements of ethical bias, if these are present, in other categories of analysis, since the same image can often be subject to important notes in more than one criterion.

As can be seen in Figure 1a), the objective was to illustrate Aurélia at the window waiting for her suitors, as described in part two, chapter four of the book *Senhora* by José de Alencar. The prompt used to generate this image was: “Conceptual art of a very beautiful young woman standing in front of the window. She attracts a crowd of suitors who pass by in carriages and on foot. The anxious looks and insinuating words of the suitor’s contrast with their cold impassivity; she stands at the window like a statue, doing her duty but without resorting to flirting tricks or seduction tactics.”

In the prompt we used the description “very beautiful young woman”. Although we did not specify skin color in the prompt, all three hundred images generated represented white women, highlighting a racial bias in the algorithm. This observation underscores the importance of addressing and mitigating biases in imaging models to ensure more diverse and inclusive representations. Figure 1a) illustrates the scene mentioned above and has the profile of a woman, young and delicate, her face well resolved in relation to the proportion of her body and the surrounding environment, the window. The young woman was portrayed with dark brown hair and period clothes, she is sensual but not vulgar and has a subtle and enigmatic beauty. The window’s perspective and the vertical framing of the scene highlight the beauty of the young woman and fix the observer’s gaze on her voluminous hair and white satin clothes, contrasting with her flushed face and lips. The image was considered a good standard for illustration to the detriment of other results that showed many deformations, poorly formed eyes and faces, sleeveless garments or cleavage and disfigured dresses.

In Figure 1b), in order to represent Bentinho’s suspicion that his son was, in fact, Escobar’s son, we used an approach based on chapters one hundred thirty-one and one hundred thirty-two of the book *Dom Casmurro* by Machado de Assis. This illustration was designed to portray Bentinho’s obsession with Escobar’s photograph, which he kept in his office. Bentinho frequently mentioned the resemblance between his son and Escobar, fueling his suspicions. Thus, the image was created by positioning the son in front of Escobar’s photograph, highlighting similar traits between the two, symbolizing Bentinho’s growing distrust. The prompt used was: “A painting of a child (four years old) in the center. In the background there is a photograph of his father leaning against the wall. The boy looks like his father, who was forty years old in the 1800s”. What can be seen

in Figure 1b) is its realistic and sensitive aspect. The set of images generated by this prompt generally created scenes with a very pronounced pictorial aspect, and although the image has the presence of three figures, two represented in the pictures hanging on the wall and the child sitting on the floor, with bare feet, and all looking at those outside the scene, there is a trace of loneliness and the expression on the faces is touching and sensitive. This relationship of a presence (the child) and an absence (portraits) in the scene promotes dialogue between text and image, one of the objectives of the proposed exercise.

Figure 1c) depicts a scene from chapter XXI of the book *O Cortiço* by Aluísio Azevedo, with the intention of representing the character João Romão, a portuguese owner of the tenement, in a 19th century Rio de Janeiro setting. The prompt used was: “Conceptual art of a man walking around in slippers and a nightgown in the bedroom, a large room covered in blue and white with small yellow flowers simulating gold, a rug at the foot of the bed, and on a bench, a nickel alarm clock, Rio de Janeiro, 19th century. Digital, illustrative, pictorial work, matte painting, highly detailed”. The selected image presents a male figure from the front, entering the room, in the background the door is ajar. The man wears a robe tied around his waist and close to the window of a room well lit by the sun, which enters through the open window on his left. On the bedside table, next to the bed, a vase with yellow flowers reiterates this luminosity and in contrast with the blue of the walls creates an atmosphere of peace and tranquility. Once again, the framing adopted by the AI to generate the scene is quite unusual and ends up highlighting the size of the room, placing the viewer as a second presence, looking from the outside, taken inside by the linear perspective configured by the parallel lines of the bed, the carpet, the window, the furniture next to the window, which together direct the gaze to the door behind the man, through which he entered the illuminated space.

To illustrate Capitu’s “hangover eyes”, described as “oblique and disguised gypsy eyes”, we used prompts inspired by the physical and symbolic characteristics mentioned in chapters thirteen and thirty-two of the book *Dom Casmurro* by Machado de Assis. As Capitu’s described as a brunette with light eyes, we assumed that her eyes were light brown (Capitu’s true eye color is not revealed in the book), so during the curation, we eliminated most of the images generated with blue eyes. We can see that the model incorporated elements of the sea in the illustrations, reflecting the metaphor of “hangover eyes”. In Figure

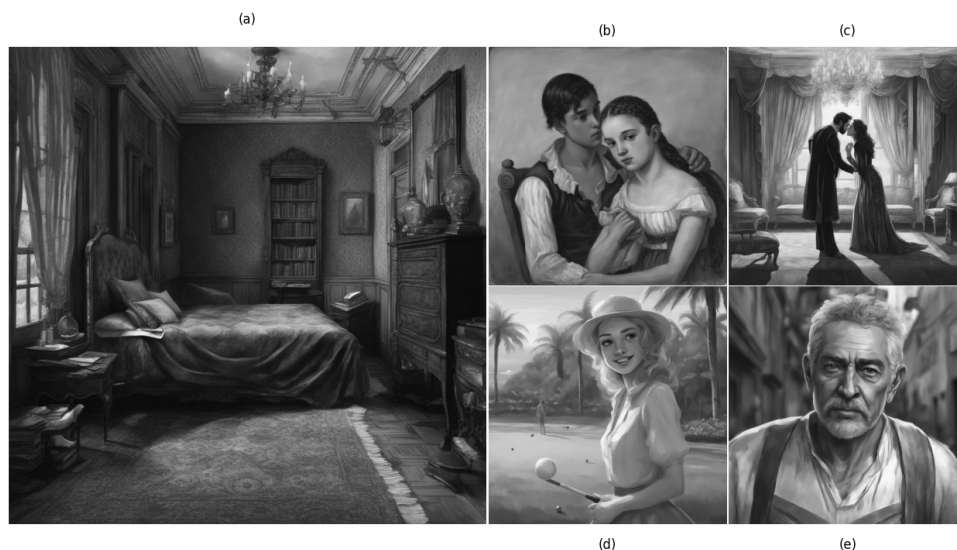
1d), this movement can be observed on the face, as well as the “sea wave” inside the generated iris and the watery pupil. The prompt used was: “painting oblique and hidden eyes. Just eyes. Mysterious and energetic fluid, like the wave that retreats from the beach, on hangover days. Brunette, clear and large eyes, straight and long nose, thin mouth and wide chin”.

Figure 1e) depicts the character Ernestina, based on the novel *A Viúva Simões* by Júlia Lopes de Almeida. The prompt used was: “Conceptual art of a rich, widowed, forty-year-old woman, from the bourgeoisie, a beautiful, tall, slender woman, with beautiful black eyes, dark skin that is delicately velvety and soft, Rio de Janeiro. Digital, illustrative, pictorial work, matte painting, highly detailed.” Among the objectives of the scene, the presence of the female figure and the landscape of Rio de Janeiro and a strong connection with features of the bourgeoisie portrayed in the book stand out. The choice for the image is due to it being the one that was most faithful to the age of the female figure, as well as the representation of the same with delicacy and highlighting a black beauty (the prompt indicated black eyes and dark skin) and also her clothes related to the bourgeoisie. The landscape features items such as palm trees, sea and mountains from which Rio de Janeiro can be recognized, not the exotic or idealized landscape, but landscape features that identify the city and its geography.

B) Composition

The composition category aims to describe images that point to compositional solutions involving the visual elements of shape, color, texture, perspective, in synergy with the content and themes proposed to be included in a given visual composition, whose final result will convey expression and intended communication or message. The highlighted set can be considered sufficient, although the composition is an element of visual language that is always subject to divergences, as it is associated with the artist and the style of an artistic movement and/or period in art history. In this study, the objective was to analyze how the machine operates its compositional skills and returns solutions consistent with the visual language, corroborating the verbal texts that were the basis of the prompts used (Figure 2).

Figure 2. Examples of generated images for (a) *Senhora*, (b) *Dom Casmurro*, (c) *Senhora* and (d, e) *A Viúva Simões* (2024), author's elaboration.



In Figure 2a), the objective was to illustrate Seixas's room, as described in part one, chapter five of *Senhora* by José de Alencar. We focus on the contrast between his luxurious lifestyle and the poverty of his mother and sister. The scene was chosen because it is described in detail in the text, revealing important aspects of Seixas' personality. The prompt used was: "Concept art of a modest, worn-out office with faded blue wallpaper and antique furniture. Iron bed with a green mosquito net contrasts with the environment. Luxurious items like a tailored black coat, elegant evening wear, Parisian hat, quality gloves and fine boots seem out of place. The embroidered cushions in blue satin stand out. A cluttered alcove with books, inkwells, ashtrays and assorted trinkets contrasts with the well-equipped chest of drawers. Corner with umbrellas, walking sticks, some valuable, as well as artistic curiosities". This image represents the discrepancy between Seixas' bohemian life, marked by expensive objects, and the humble reality of his family. It is worth highlighting the relationship between textures, colors and the sophisticated environment composed of the rococo chandelier and decorated ceiling. The composition is centered on a linear perspective that takes the viewer's attention to the bookcase on the wall at the back of the room on the side of the bed that takes up the entire topological extension of the room. The cerulean blue walls and the fine porcelain vases in front of a mirror reflect the environment with the light that comes from

the glass window and blue curtains and light fabric, in short, a whole visual set that shapes the illustration, giving it the visuality generated by the verbal, in this case, by the prompt, which, in short, was the purpose of the study carried out. It is worth noting that not all objects present in the prompt were captured by the AI, a fact that corroborates the idea that concise prompts are recommended for greater effectiveness.

In Figure 2b), the objective was to illustrate the scene of the first kiss between Bentinho and Capitu, as described in chapter 33 of *Dom Casmurro* by Machado de Assis. This scene marks the beginning of the romance between the two characters, after Bentinho braids Capitu's hair. The prompt used was: "A painting of two teenagers. The girl has braids, sitting in a chair while a boy (standing) kisses her. 1800s". During the process of generating these images, the AI frequently made mistakes, such as braiding both Bentinho and Capitu, and confusing the position of the characters, often placing Bentinho sitting instead of Capitu. Furthermore, deformations in the arms and limbs were common, as they tended to intertwine in the images (as can be seen in the characters' hands, in the Figure in question). Despite these difficulties, the final image captured the essence of the moment described in the text, focusing on the beginning of the romance between Bentinho and Capitu.

Among the images generated and analyzed by the image curator, Figure 2b) was the one that most resulted in a text and image dialogue, with a more delicate result, with the bodies close together, appearing intimacy without vulgarity, a soft and delicate touch between the bodies, showing a meeting between two young people, without exposing the female figure to any form of embarrassment and with satisfactory staging and scenic construction solutions. The composition is slightly asymmetrical and the bodies, which have their arms intertwined, draw a circle right in the center of the scene, which makes the representation of the encounter significant. The image represents a tender moment between the young couple. He rests his face on her forehead and delicately holds her long braided hair, and she has her left hand lightly resting on one of the boy's arms, giving shape to the circularity of the composition described above. They find themselves in a position of intimacy and complicity.

In Figure 2c), the objective was to illustrate the final scene in which Silas returns the money to Aurélia, as described in the last excerpt of chapter nine of the last part of *Senhora* by José de Alencar. The prompt used was: "Conceptual

art of a room elegantly decorated with heavy curtains and exquisite furniture. In the center of the room, a man and a woman are present. While the woman kneels before him, holding his hands fervently, the man, visibly moved, kisses the woman passionately while holding a document in his hand”. During the process of generating these images, the AI was challenged to capture the emotional intensity and descriptive details of the final scene and the moment of reconciliation between Silas and Aurélia, highlighting the environment decorated with heavy curtains and exquisite furniture, while the woman slightly bends her body in front of the man, holding his hands fervently, and he kisses her passionately, holding the significant document. The illustration faithfully reflects the dramatic and emotionally charged outcome of Alencar’s narrative. The composition of the image is drawn from the center, in which Silas and Aurélia are illuminated by a strong light that enters the environment through the window in the background. The white curtains and the sophisticated chandelier blend into the environment due to the load of light that comes from the window. However, this same light is not enough to overshadow the meeting between the characters and the final outcome with intensity and emotion, this is because the composition of the image projects the bodies standing in the center of the room, giving them a prominent place and leading the form to reverberate the content of the scene, which is why this image is considered an excellent result in the composition category.

Figure 2d) illustrates the character Sara, present in chapter I of the book *Viúva Simões* by Júlia Lopes, whose prompt was: “Conceptual art of a twenty-year-old girl playing croquet, bourgeoisie, slightly large head, strong and rounded forehead, brown and intelligent eyes, fiery and luminous blond hair, smiling mouth, good health, innocence, joy, frank gaze, pink, fresh skin, friendly mouth, Rio de Janeiro. Digital art, illustrative, pictorial, matte painting, highly detailed”. Among the images generated, this was the one that most visually met the text indicated in the prompt, above all, as it resulted in figures without deformation and in an accurate way in the bourgeois representation of the scene, and aesthetic aspects of the girls, hair and age of the characters to be portrayed. The set composed of landscape, human figures and the action to be figurativized in the image, based on the indicated prompt, reveals that the representation achieves its purpose. Nonetheless, it is curious to note in the composition that the pink color of the clothes, and perhaps not the white that would be more common or conventional in this sport’s uniforms, added to the pink sky and

even the pink ball that adds to the flowers in the background in the garden, all matching pink, lead us to question the association of this color with the feminine and the age of the characters, or a convention regarding the representation of the feminine in the practice of said sport. Even so, among the results, it is reiterated, this was the one that could be most used in terms of dialogue with Júlia Lopes' text.

Figure 2e) illustrates Luciano, another character of the novel *A Viúva Simões*, described in chapters I and II. The prompt used was: "Conceptual art of a man, gray hair, thick and virile physiognomy, not slender, rounded belly, expressive and friendly face, dark circles under his eyes, Rio de Janeiro. Digital art, illustrative, pictorial, matte painting, highly detailed". The composition of this image was highlighted by the sum of the surroundings or the landscape in the background made up of houses that look like a shopping street, in perspective, and the figure of the man in front, generating a very photographic relationship between figure and background. The hair, age and physical type also have good adherence to the prompt, in addition to the expression of pain on the character's face. The composition of an image, despite being considered an illustration based on its text-image relationship, must be based on good aesthetic solutions proportionally balanced between form and content, and the set of images in this category, composition, has this quality.

C) *Ethical bias*

Generative models have proven to be tools with great potential for increasing human creativity. However, it is known that these models can encode various prejudices and social stereotypes, present in the big data that feed AI tools. This fact occurs due to the dependence of these models on large sets of images that serve as references for the generative process. These datasets can present problems as they are often curated automatically.

As Sarzi-Ribeiro & Sedeño-Valdellós (2024) suggest, the ethical biases present in the different stages of content production and creative process with generative AI tools highlight two points that deserve attention. The first point refers to the generation of data generally produced by developed countries, whose aesthetic standard of images results from social and cultural values that can manifest different biases and, secondly, the analysis of models developed using hyper parameters, that is, who develops the models and/or whoever makes the decision to create the model does not always care about the social impact.

Therefore, we believe it is necessary to reflect on the results of this experiment, taking ethical biases as one of the analysis criteria (Figure 3).

Figure 3. Examples of generated images for (a, d) *Senhora*, (b) *Dom Casmurro*, (c) *A Viúva Simões* and (e) *Triste Fim de Policarpo Quaresma* (2024), author's elaboration.



Therefore, we chose to describe some results to reflect on these possible biases, such as those observed in a set of images that contained questionable elements in their ethical results. The first of these refers to the fact that a male room, like the one in Figure 1a), is represented by images of a disorganized, messy room, which is stereotypically associated with male environments. The image of the room in question was generated by the prompt used to create a possible illustration for the scene in Seixas' room, from the book *Senhora* by José de Alencar, and immediately caught the researchers' attention when some images showed disorder, the rooms were dark and they looked dirty. It is worth reflecting on these results and the ethical bias that may be implicit in the way the machine interprets the male bedroom command and prospects for stereotypical scenes.

Another result that caught attention refers to the first kiss scene from the book *Dom Casmurro* by Machado de Assis, as in Figure 3b) (same prompt as Figure 2b), when it was noticed that among the results there were several images with eyes cross-eyed, deformed heads and girls with few clothes or underwear, as seen in the aforementioned image in which the young woman

appears wearing a dress that looks more like a female nightgown and has her legs exposed while being kissed by the boy who is fully dressed. Such results could be associated with greater exposure of women or female figures, as in a society guided by male gazes.

Figure 3c) (same prompt as figure 1e), which illustrates a scene from the book *A Viúva Simões* by Júlia Lopes, deserves to be highlighted since the prompt requested the image of a rich, mature, 40-year-old bourgeois woman in Rio de Janeiro. Among the results there were several with women wearing extravagant and exotic hats, like this one with large red flowers, which also indicates an ethical bias in which bourgeois women in Rio de Janeiro continue to be represented by exoticism or even excessive adornment, as a synonym for wealth, generating this stereotype.

In addition, Figure 3d) is Aurélia's revelation after the wedding, described in part one, chapter thirteen of *Senhora* by José de Alencar. The prompt used was: "Conceptual art of a couple in a luxurious and opulent setting, a living room elegantly decorated with fine wood furniture and luxurious fabrics. The woman is sitting in an upholstered chair, with an expression of disgust and triumph, holding an open sheet of paper. The man is kneeling at the woman's feet, his features twisted in shock and despair". Despite there being a clear mention of the fact that the man was kneeling, in none of the images generated this happened, and in some of those generated the woman actually appeared on her knees, indicating yet another gender bias. In the book it is mentioned that she wore a green dress in this scene (which was not specified in the prompt), however, for some reason the algorithm assumes in all cases white dresses similar to those of a bride. You can see a robust man, with a gigantic chest, sitting next to Aurélia, a young woman represented by a slender body and gray hair. Images of this type were disconcerting, because even though the surrounding environment is sophisticated and the scene is composed in a balanced way, the bodies are stereotypical. Aurélia wears exaggeratedly low-cut outfits that do not match the models of the time, and she also appears with her breasts exposed, while the man is fully dressed. The female faces are "scary" and caricatured, while the men are cold and do not express any feelings, and this could be related to the social roles imposed on the female and male genders.

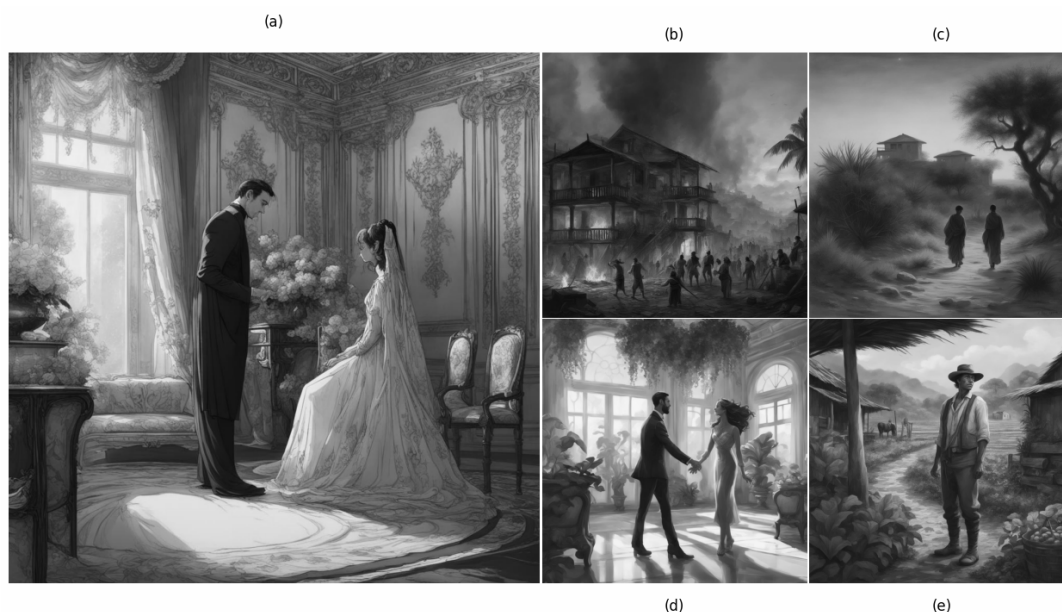
Finally, Figure 3e) illustrates the attitude of the character Policarpo, described in the third part of the book *Triste Fim de Policarpo Quaresma* by Lima Barreto, whose prompt used was: "Conceptual art of a young man in command

of an artillery platoon, in battle, repelling attacks by sailors, on beaches, 18th century, Rio de Janeiro. Digital art, illustrative, pictorial, matte painting, highly detailed". The proposal for the scene included the representation of a young boy in charge of an artillery platoon and involved battle scenes, beaches and sailors, in the 18th century in Rio de Janeiro. When observing the image carefully, the visual mention of the figure of Napoleon Bonaparte and the historical paintings legitimized by art history and the Neoclassicism movement become clear. Although the image has a beach, vegetation typical of the tropics, sailors, battles and a young man wearing military clothes, common sense or rather the association with the figure of Napoleon appears clearly, including in the pose of the young man who has his hands on his hips.

D) Images that did not meet the objective of the study

When treating this set of images as a category arising from the study carried out, the objective is to seek to learn from the knowledge generated by the images that were considered insufficient, aiming to reflect on them as they indicate obscure points of image production with artificial intelligence and that need to be investigated (Figure 4).

Figure 4. Examples of generated images for (a, d) *Senhora*, (b) *O Cortiço*, (c) *Horto* and (e) *Triste Fim de Policarpo Quaresma* (2024), author's elaboration.



The image that resulted in Figure 4a) (same prompt as Figure 2c), from the book *Senhora* by José de Alencar, led us to reflect on what could perhaps be considered a *modus operandi* of the machine that chose the same color and fabric pattern as the curtains and the environment, for the young woman's dress. That's why in this illustration, the white flowers, the white walls with gold designs, the upholstery of the chairs and sofa, the curtains and the rug seem to unfold into the clothes of the young woman who is portrayed as a bride, even though the prompt didn't even indicate that the female figure would be a bride or would be wearing a wedding dress. It is assumed that artificial intelligence seeks harmony between the elements that should make up the image and to achieve this it uses the same pattern for nearby textures such as fabrics, upholstery and flowers.

Figure 4b) is an image generated as a proposed illustration of a scene from the book *O Cortiço* by Aluísio Azevedo, of a scene in which people flee a fire, the prompt used was: "Conceptual art of a fire in a tenement, confused people, Rio de Janeiro, 19th century. Digital art, illustrative, pictorial, matte painting, highly detailed". Contrary to what the prompt indicated, in most of the images generated there are people in the middle of the flames and they do not appear afraid or confused by the fire. The images generated by this prompt were strange because people appear to be around or between large bonfires, with an entire house and all the floors being consumed by fire.

Figure 4c) was generated to illustrate a scene that refers to the backcountry path from the book *Horto* by Auta de Souza. The prompt used was: "A painting of two brothers walking through the dense vegetation of an arid environment. A distant house is seen in the background, night sky, realism, 19th century". The creation of the prompt aimed to materialize the melancholic feeling arising from the distance from home, presented by two brothers, on an arduous and lonely journey through the backlands. The images generated for this prompt do not fully comply, mainly because they are missing important visual references such as "a house in the background", arid landscape and the type of vegetation that does not match that described in the command given to the AI program used in the exercise. In the specific case of vegetation, the green or exuberant plants for the arid climate did not allow them to be considered for the text-image dialogue, as did the number of people in the scene, which were two, but in many cases (different from the one presented) there were images with three people.

Figure 4d) depicts the ball scene where Aurélia faints, in part three, chapter four of the book *Senhora* by José de Alencar. The prompt used was: “Conceptual art of an elegant living room, adorned with tall artistic bronze vases in each corner. The vases are filled with live plants, creating a country atmosphere. In the center of the room, a couple dances a waltz, eager for space. Their lips touch briefly. The woman faints in the man’s arms”. Curating the images generated from this prompt was quite challenging, as the results presented several problems, such as women’s dresses and costumes that did not correspond to the requested period, excess vegetation, such as ferns that dominated the scene, disproportionate bodies and deformed faces with “beak” mouths, large ears, distorted facial expressions and “monstrous” heads. It is suspected, by observing the results, that very complex images with elaborate compositions are more difficult to achieve harmonious results. This reinforces the hypothesis that the large number of details can lead AI to prioritize certain elements over others, failing to harmonize them all satisfactorily. Thus, it becomes clear that prompts with many details can challenge the AI’s ability to generate coherent and balanced images.

Figure 4e) illustrates the problems faced in the place called Sossego by the character Policarpo, described in the second part of the book *Triste Fim de Policarpo Quaresma*. The prompt used was: “Conceptual art of a young retired man on a farm, poverty, lack of cultivation, 19th century, rural area, Rio de Janeiro. Digital art, illustrative, pictorial, matte painting, highly detailed”. Curating the images generated from this prompt was also quite challenging. The results seemed very artificial and not very credible, and the landscape had no relationship with the text that was intended to be illustrated. The figures did not convey the misery or lack of cultivation of food or animals that would justify the connection with a situation of penury; on the contrary, the vegetation was abundant and green. Furthermore, a serious ethical bias was identified: although the AI did not generate images that represented misery as desired, among the images generated, all the men represented were black. This provoked strangeness and reflection, highlighting the need to address and correct these biases in the AI models used.

Conclusions

The study revealed itself to be an in-depth exercise in the production and analysis of images generated by Gen AI, with the aim of illustrating books of

classic Brazilian literature. In addition to generating visual results, the study provided significant reflections on the current stage of technological development in the automation of image production, highlighting the still unclear controls on the results to be achieved.

Some texts in the chosen books offered clearer or perhaps more precise prompts to the detriment of texts that are naturally more complex and whose visuality could perhaps be considered less obvious. It became clear, after the exercise carried out, that artificial intelligence presents different degrees of difficulty and makes choices prioritizing certain visual elements over others to generate its compositions, such as the colors of women's dresses that seem to follow the color of other fabrics or materials, which define the textures of the image. It is assumed that, in this case, we could be facing an adjustment to the texture of the materiality most present in the image and that this sensitive adjustment should be indicated very precisely in the prompt.

Likewise, when it was hoped to create an illustration for *Os Sertões* by Euclides da Cunha, the images generated mostly had the presence of indigenous people associated with the original peoples of Brazil, and the vegetation was very exuberant, with a dense forest and green, which did not correspond to the description of a type of vegetation that could be identified with the Brazilian hinterland or the caatinga, pointing to difficulties in regional representation or regionalism. It is worth reflecting on the fact that it is possible to create illustrations with AI, but the ethical bias is present in a large part of the images generated and there were several examples that could be reported in this article, ranging from racial issues, gender, economic class, all the way to age and geography.

It is important to highlight that there are several additional aspects to be considered when using Gen AI as a creative tool for illustrations, which were not addressed in this study, but which could be explored in future research. One of these aspects is the concept of continuity, fundamental when the artist needs to represent a character in multiple contexts throughout a book. In the current scenario, the application of the described method may result in unwanted variations in the character's traits, according to the initial random noise present in each generation. To overcome this limitation, alternative tools, such as fine-tuning with Dreambooth (Ruiz *et al.*, 2023), could be employed. Furthermore, a

subsequent post-processing step, using Image to Image models, could be incorporated into the methodology to adjust and improve the generated images, correcting discrepancies by making it possible to change the color of a dress or enhance the vegetation in an image. In summary, the intersection between literature (especially when it comes to illustrations) and Gen AI is extremely new and offers ample opportunities to expand and enrich existing research.

References

- Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Alteschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., Belgum, J., Bello, I.,... Jain, S. (2023, March 15). GPT-4 Technical Report. *Cornell University*. <https://arxiv.org/abs/2303.08774>
- Azevedo, A. (1991). *O cortiço*. Moderna.
- Barreto, L. (2011). *Triste fim de Policarpo Quaresma*. Companhia das Letras.
- Betker, J., Goh, G., Jing, L., Brooks, T., Wang, J., Li, L., Ouyang, L., Zhuang, J., Lee, J., Guo, Y., Manassra, W., Dhariwal, P., Chu, C., Jiao, Y., & Ramesh, A. (2023). Improving Image Generation with Better Captions. *Computer Science*, 2(3), 1-19. <https://cdn.openai.com/papers/dall-e-3.pdf>
- da Cunha, E. (1967). *Os Sertões: campanha de Canudos*. Coleção da Edições de Ouro dos Clássicos Brasileiros.
- de Alencar, J. (1875). *Senhora: perfil de mulher (volume 2)*. B. L. Garnier.
- de Souza, A. (2019). *Horto*. LeBooks Editora.
- Gapski Giordani, M. V., & Panek, B. M. (2020). A transposição midiática: uma interação entre literatura, ilustração e a técnica da gravura. *Brazilian Journal of Development*, 6(8), 59714-59730. <https://doi.org/10.34117/bjdv6n8-401>
- Lopes de Almeida, J. (1897). *A viúva Simões*. António Maria Pereira Editor.
- Machado de Assis, J. M. (2024). *Dom Casmurro*. Penguin Clássicos; Grupo Editorial Unipessoal.
- Piskopani, A. M., Chamberlain, A., & Ten Holter, C. (2023, July 11-12). *Responsible AI and the Arts: The Ethical and Legal Implications of AI in the Arts and Creative Industries* [Conference presentation]. First International Symposium on Trustworthy Autonomous Systems, Edinburgh, United Kingdom.
- Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022, December 20). High-Resolution Image Synthesis with Latent Diffusion Models. *Cornell University, arXiv*, 10684-10695. <https://arxiv.org/abs/2112.10752>
- Roumeliotis, K. I., & Tselikas, N. D. (2023). ChatGPT and Open-AI Models: A Preliminary Review. *Future Internet*, 15(6). <https://www.mdpi.com/1999-5903/15/6/192>

- Ruiz, N., Li, Y., Jampani, V., Pritch, Y., Rubinstein, M., & Aberman, K. (2022). DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation. *Cornell University, arXiv*, 22500-22510. <https://arxiv.org/abs/2208.12242>
- Sarzi-Ribeiro, R., & Sedeño-Valdellós, A. (2024). Vídeo experimental y procesos creativos con Inteligencia Artificial: algunas obras de Ben Bogart. *AusArt*, 12(1), 257-267. <https://doi.org/10.1387/ausart.25813>
- Walczak, K., & Cellary, W. (2023). Challenges for higher education in the era of widespread access to generative AI. *Economics and Business Review*, 9(2), 71-100. <https://doi.org/10.18559/ebr.2023.2.743>