GenAI ConceptLab: A Generative AI tool for Assisting Artists in Crafting Anime Characters and Narratives

Matheus Lobo[1], **Pamela T. L. Bezerra**[1], **Gabriella A. B. Barros**[2]

[1]Faculdade do Centro de Estudos e Sistemas Avançados do Recife – CESAR School
Avenida Cais do Apolo, n° 77 – 50.030-220 – Recife – PE – Brasil

{mlva,ptlb}@cesar.org.br

[2]Profissional Independente
Alagoas – Brasil

gabbbarros@gmail.com

Abstract. Introduction*: Generative AI has impacted how artists, either professional or amateur, create art. However, many generative tools require extensive training to develop good prompts and select the appropriate model features to produce high-quality results.* ***Objective:*** *To reduce the necessary learning curve and technical understanding of generative AI, this paper proposes a web application that integrates generative image and text models to create anime character concept art and minibiographies. The main goal is to assist independent creators and anime enthusiasts by providing an accessible and scalable tool that minimizes prompt engineering through an intuitive interface.* ***Methodology or Steps****: To develop the proposed solution, an initial study of similar solutions and generative models focused on anime style was conducted. This was followed by the development of the web application and the exploration of different prompts and interfaces. Finally, system and user evaluations were performed to analyze the solution's capabilities in generating images and text and to collect users' opinions.* ***Results:*** *User tests showed positive feedback on the interface and the quality of content generated. Meanwhile, the system's tests indicated that the solution is capable of generating art for a wide range of anime genres, with a positive correlation between diversity and the amount of details provided.*

Keywords *Generative AI Models, Concept Art, Anime Style, Game Art.*

1. Introduction

Recently, the anime industry has become a global phenomenon, valued at $24.80 billion in 2021 and expected to grow by 14% by 2030 [Grand View Research 2022]. From 2020 to 2022, the demand for anime media increased by 118% [Hollywood Reporter 2022], with successes like *Demon Slayer: Mugen Train* and cross-cultural projects such as *Cyberpunk: Edgerunners*. This growth has inspired many creators to produce anime-style media. Online platforms like Jump+[1], Wattpad[2], and Kickstarter[3] create spaces to support emerging talent, but few tools focus on helping beginners

[1]https://mangaplus.shueisha.co.jp/updates
[2]https://www.wattpad.com/
[3]https://www.kickstarter.com/

conceptualize characters. In games, the anime aesthetic shines in various titles, such as *Fire Emblem* [Intelligent Systems 1990], *Final Fantasy* [Square Enix 1987], and the *Xenoblade* series [Monolith Soft 2010], and this art style continues to rise in popularity in the gaming industry [Spotlight Report 2024].

Meanwhile, recent advances in generative AI, such as ChatGPT [OpenAI 2023] and Stable Diffusion [Esser et al. 2024], have transformed art creation, but these tools often require well-crafted prompts, making them less accessible to casual users. In early 2025, a viral trend emerged where AI was used to mimic Studio Ghibli's style, sparking both enthusiasm and controversy[4]. Ethical concerns about the replication of living artists' styles led to restrictions, reflecting ongoing debates about the intersection of AI and creativity.

This paper presents **GenAI ConceptLab**, a web-based system that generates anime-style concept art and character descriptions. It simplifies creation by requiring only key traits, automating the prompt engineering process, and delivering high-quality results. The system was evaluated through user testing and output analysis, showing strong usability and creative potential for artists, writers, and developers.

This paper is structured as follows: Section 2 describe similar solutions and research; Section 3 details the web system developed; Section 4 describes the experiments performed on the system and with users; Section 5 describes and discusses obtained results; Section 6 discuss the main conclusions and potential future steps.

2. Related Work

Midjourney [Midjourney 2022], *DALL-E* [OpenAI 2021], and *Stability AI's Stable Diffusion* [Rombach et al. 2022] are currently the most popular models for image generation. All three are general-purpose systems trained on broad and diverse image datasets, and are commonly used to generate visuals ranging from realistic photography to abstract art. However, their generality also means that they are not specifically optimized for generating images in highly stylized aesthetics such as anime.

Platforms like *Artbreeder*[5] emerged to fill this specific gap, offering tools tailored for character creation, including anime-style portraits. Its "splicer" feature allows users to combine different base images and manipulate their visual "genes" to customize the appearance of generated characters. Still, the results are heavily constrained by the available base images and latent representations, often limiting the uniqueness and creativity of the output.

A more targeted approach is seen in "*This Waifu Does Not Exist*" and its successor, *Waifu Labs* [Sizigi Studios 2019], which use StyleGAN for image generation and GPT-3 for producing corresponding narrative descriptions. The first version focuses on fully automated generation, producing new faces and plot synopses every few seconds with no user control. Waifu Labs, on the other hand, introduces some interactivity by allowing users to guide the creation process by choosing base portraits, poses, and color palettes. Nevertheless, the generated outcomes tend to remain close to the original face, limiting visual diversity and stylistic range.

[4]https://variety.com/2025/digital/news/openai-ceo-chatgpt-studio-ghibli-ai-images-1236349141/
[5]https://www.artbreeder.com/

More recent tools like *Niji Journey* [Spellbrush 2022] and *Holara* [Holara 2023] offer a more advanced and style-specific solution for anime art generation. Niji Journey produces four image variations per prompt and allows users to refine the final image, while Holara specializes in anime aesthetics through a suite of eight proprietary models based on Stable Diffusion, each targeting different stylistic attributes (e.g., *Aika* for traditional anime, *Vibrance* for lighting effects, *Lucid* for realism). The main distinction between the two lies in the user interaction model: Niji Journey emphasizes prompt-based variation, while Holara focuses on curated visual styles.

Despite their innovations, most of these systems either prioritize visual fidelity or aesthetic variation, but fall short of offering true creative control or coherence between visual and narrative elements. Furthermore, they often lack mechanisms for integrating personalized storytelling into the image generation process or enabling deeper customization beyond superficial traits.

In light of these limitations, the present work aims to bridge this gap by proposing a generative system that integrates visual and textual generation with user-driven control over both design elements and narrative development. Unlike existing approaches that focus either on random generation or on shallow aesthetic tuning, our system introduces a collaborative pipeline that balances visual coherence, stylistic diversity, and narrative depth, thus enhancing the creative potential of generative AI in anime character design.

3. Proposed solution

The *AI-Based Character Concept Generator* (*GenAI Concept Lab*) is a web system that leverages generative AI models to create detailed text descriptions and conceptual images of anime characters based on user-provided input. It allows users to customize various aspects of their characters, such as species, gender, age, physical traits, artistic style, and narrative universe. The following subsections detail how the system was built, how users interact with it, and how it manipulates prompts and input.

3.1. Specifications

The system's **backend** is built in Python. It processes user input and generates prompts for AI models. Communication with AI models occurs through API calls to text and image generation services, utilizing the OpenAI library[6] for textual prompt generation and the previously mentioned Holara API for image creation.

The OpenAI API was chosen for its efficiency in generating coherent and context-aware text from minimal input, enabling fast and high-quality character descriptions. Meanwhile, Holara's API complements this by producing diverse and original anime-style visuals without relying on existing character templates, ensuring creative freedom and alignment with user-defined traits. Together, these models provide a streamlined and flexible solution for generating unique character concepts.

The **front end** is implemented using Streamlit[7], a tool that enables the rapid development of interactive user interfaces. This makes it easy for users to input descriptions, adjust parameters, and view results in real-time.

[6]https://github.com/openai/openai-python
[7]https://streamlit.io/

The primary model used for text prompt creation is the GPT-4o language model, which transforms user input into structured and optimized descriptions for image generation. The visual representation of characters is generated by an AI model specialized in text-to-image synthesis (Holara), attempting at providing an accurate match between the described concept and the generated image. Figure 1 summarizes the system's components and the flow of information. The user can select character traits via the web interface, which sends the information to the prompt engine. The prompt engine adapts prompts to fit received data and creates a character description, which it then feeds Holara API to generate images from. The prompt engine also records the generated descriptions for further modification, if necessary. Finally, when Holara API returns the images it created, they are shown to the user, alongside the generated description.

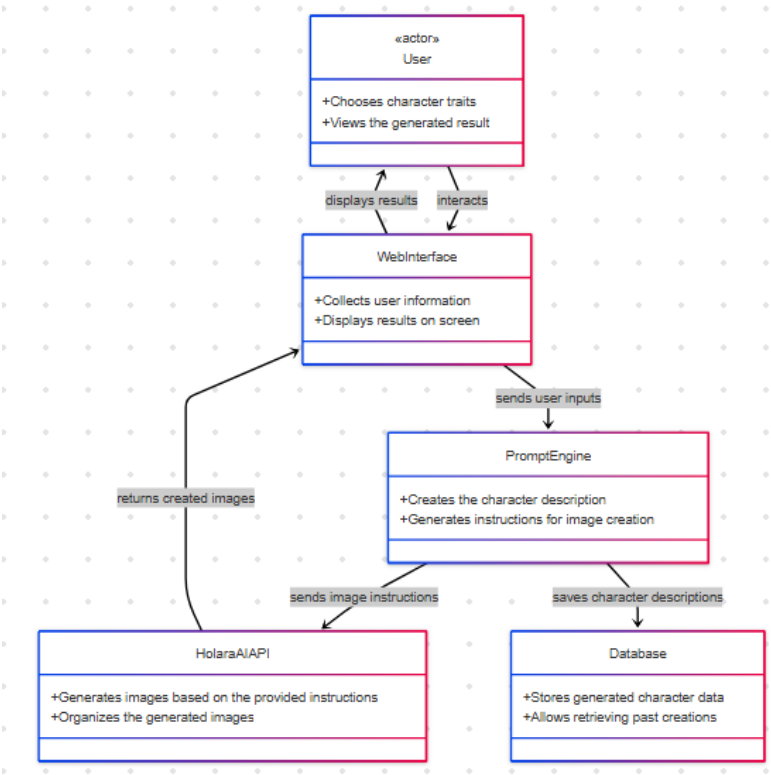

Figure 1. System Architecture. The front end was developed with Streamlit, while the backend, which is responsible for communicating with the generative models, is entirely in Python.

3.2. System overview

The system's primary functionality is to transform user-provided textual descriptions into structured prompts, which are then used to generate high-quality images via a specialized generative AI model. Additionally, users can adjust the *creativity level* (model temperature), influencing the degree of originality and adherence to the provided information. Higher *creativity level* indicate higher diversity of generated images.

3.2.1. User Interface and Input Process

The system provides an interactive UI where users can fill in multiple text fields, each corresponding to a specific character attribute. These attributes include *physical traits* (e.g., "Bright red eyes", "Futuristic outfit with metallic details") and *personality aspects* (e.g., "Extroverted and energetic personality"). The UI also allows users to select the desired artistic style and level of detail in the generated images. Upon submitting the information, the system processes these inputs and organizes them into **two structured image prompts** and a brief **character text description**, which serve as the foundation for AI-generated artwork of that character and its text-based backstory, respectively.[8]

3.2.2. Prompt Construction Process

The system modifies a **predefined template prompt**, which GPT API processes to generate both the structured character description and the image-generation prompt. This process consists of three main stages:

1. **Text-Based Character Description Generation -** The system sends a template prompt and the user input to the GPT API, which follows a predefined instruction prompt shown in Table 1. This prompt is trained to interpret the user's information and generate three elements:
    - Two structured prompts optimized for image generation
    - A narrative text describing the character's story, integrating the provided attributes into a coherent and detailed background.
2. **Image Generation with Holara -** The two structured prompts generated previously are sent to Holara, via an API key, in order to generate the images. This request include parameters like: The *model* used for image generation, the *prompt* defining the character's visual attributes, the *number of images* to be generated, and the *specific width and height* of the output images.
3. **Presentation of results -** Once both images and descriptions are generated, they are presented simultaneously to the user, who then can accept them or modify their input in order to continue the process.

3.2.3. System Output

The final output consists of:

- Two AI-generated images reflecting the provided character attributes.
- The textual prompts used to guide the image-generation process.
- A character description that integrates all input details into a cohesive profile.

Figure 2 and Table 2 below describe the final results presented in the interface. In Figure 2 we can see the two images generated disposed side-by-side, while Table 2 details the description prompts used for a different character.

[8] `https://ai-concept-generator-we6zzybsyfxeffyu9pgpsg.streamlit.app`

Table 1. Prompt used in the first step of generation.

'I want you to act as a prompt engineer. You will help me write prompts for an ai art generator called Holara.I will provide you with short content ideas and your job is to elaborate these into full, explicit, coherent prompts. Prompts involve describing the content and style of images in concise accurate language. It is useful to be explicit and use references to popular culture, artists and mediums.Your focus needs to be on nouns and adjectives. I will give you some example prompts for your reference. Please define the exact camera that should be usedHere is a formula for you to use(content insert nouns here)(medium: insert artistic medium here)(style: insert references to genres, artists and popular culture here)(lighting, reference the lighting here)(colours reference color styles and palettes here)(composition: reference cameras, specific lenses, shot types and positional elements here) the most important part is when giving a prompt remove the brackets, speak in natural language and be more specific, use precise, articulate language. always output me two full prompt options that are different Example prompt: Portrait of a Celtic Jedi Sentinel with wet Shamrock Armor, green lightsaber, by Aleksi Briclot, shiny wet dramatic lighting.Always write the prompts following this example. '
'after generating the prompt, gather all the informations that i gave you and the informations from the prompt, use as inspiration to generate a detailed backstory and description for the character. '
'give me the answer like this: '
' prompt 1: '
' prompt 2: '
' Description: text (Do not break a line here, write everything on the same line!)'
'do not give any more words or phrases that arent from the prompt and description, also dont write prompt1, prompt2 at the start of the sentence!.

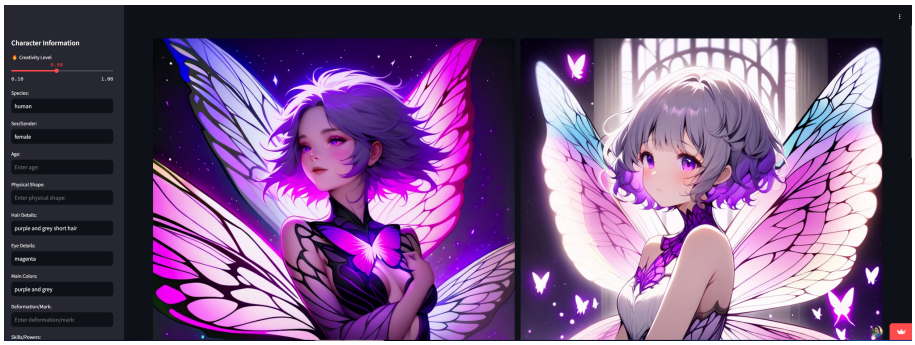

Figure 2. System interface showing the images generated.

Table 2. Description of a red-haired, blue-eyed, Viking character, and the prompts used to generate the images.

Prompt 1: Portrait of a young male human with fiery red hair and striking blue-silver eyes, wearing winter Viking clothes, set in cold colors and adorned with freckles, in a snowy landscape.
Prompt 2: A description of a 16-year-old human boy with red hair and unique blue-silver eyes, dressed in traditional Viking attire for winter, featuring a cold color palette and landscape.
Character Description: Born in a small village on the edge of the Arctic Circle, this young man named Erik has always been fascinated by the tales of the ancient Vikings who once roamed the frozen lands. With his fiery red hair and piercing blue-silver eyes, he stood out amongst his peers. Erik's adventurous spirit led him to embark on a solo journey across the North, seeking to uncover the mysteries of his ancestors and test his own limits. Donning traditional Viking clothes to brave the harsh winter, he set out to explore the icy tundras and snow-covered mountains, determined to forge his own path and discover the secrets hidden in the frozen wilderness.

4. Experiments

The current experiments aim to evaluate (1) *the website usability* and (2) *the generative models' capability* of generating good and coherent images and texts. To evaluate the models' capabilities, a list of scenarios exploring different numbers of features was considered. As for the usability tests, the website was shared among designers,

professionals, and students, and their opinions were collected via a Google Survey. The following sections detail each type of test. Discussion on the obtained results is presented in section 5

4.1. System's Evaluation

This test evaluated the system's creative capacity while identifying potential limitations. Initial tests involved generating characters inspired by existing pop culture characters. The goal was to assess how well the system incorporates inspirations and whether the generated character was consistent with the provided input.

Additional tests were conducted to evaluate how different anime art styles influence the generated results, both visually and textually. Each test explored a **complete** approach, filling in all 17 description fields, and a **specific** approach, where only the essential aspects were provided. It is expected that the complete test will generate longer prompts but more detailed images, while de specific test will have fewer details and more creative freedom.

For example, in one test, a young Viking warrior inspired by the popular manga "Vinland Saga" was requested, with a dark color palette and strong contrasts of light and shadow. The complete version generated two images with specified details such as age, eye and hair color, personality traits, and accessories like a dagger and a scar, while the specific version was limited to fundamental characteristics like red hair, a scar, and Viking attire. In another test, a "Mahou Shoujo" (magical girl fighting evil using special powers) character, inspired by "Mahou Shoujo Madoka Magica," was described with pink tones and an ethereal atmosphere. The complete version included detailed information about magical powers, a school setting, and facial expressions, whereas the specific version retained only essential elements such as hair color, outfit, and the magical girl concept.

These tests focused, therefore, on how the system responded to different styles and levels of detail, analyzing the diversity of genre and aesthetics in the generations, as well as the impact of textual and visual accuracy on the fidelity of the created characters.

4.2. User Tests

To test the system's usability and outputs, a survey was developed to collect user feedback. The objective was to gather data about users' experiences and perceptions of the system. The survey was active for 18 days, and it was distributed among friends of the author and students and teachers from our higher education institution, CESAR School. A total of 20 responses were obtained. It included questions designed to evaluate different aspects of the application's functionalities. Most questions contained a rating scale from 1 to 5 and covered the following topics:

- **Previous experience with AI-based character creation**.
- **User's experience in character creation**: No experience, Amateur (hobbyist), Experienced (frequently writes/draws), Professional (works in the field)
- **Coherence of the character description**: This refers to how well the generated text integrates the provided attributes into a consistent and believable persona.
- **Originality of the generated character**: Measures how unique and creative the AI-generated character appears.

- **Aesthetic appeal of the generated images**: Assesses the visual quality and attractiveness of the artwork.
- **Accuracy of the image representation compared to the description**: Evaluates how faithfully the generated images reflect the intended character attributes.
- **User satisfaction with the generation time**: Determines whether the system's processing speed meets user expectations.

4.2.1. Participants Stats

The following insights about the participants were gathered from the survey:

- **Experience with AI for character creation (optional response)**: 9 participants reported previous use of AI for character creation, while 9 had no prior experience.
- **Level of experience in character creation (optional response)**:
  - 2 participants identified themselves as professionals in the field.
  - 5 participants were experienced creators (drawers or writers).
  - 9 participants were hobbyists or had no experience in character creation.

5. Results and Discussion

This section presents and discusses the results of the experiments previously described.

5.1. System's evaluation

These tests, performed on the generational aspects of the system, focused on how it created different styles and dealt with various levels of detail. Key observations include:

- **The system performed well when handling pop culture characters, generating images with good fidelity to the requested character, and incorporating recognizable features.** However, none of the generated images were 100% faithful to the original characters. This does not imply that the tool exclusively generates original content, but rather highlights its ability to do so.
- **Descriptions of female characters tended to emphasize physical appearance, with less focus on personality and abilities. Male character descriptions were more detailed, including both physical and personality traits.** For example, in the *Viking* test, male characters were described with greater depth, highlighting both their physical characteristics and distinct personality traits. On the other hand, the *Mahou Shoujo* test, which focused on female characters, leaned more towards emphasizing visual details such as hair, clothing, and body, with less attention to the inner qualities or abilities of the character. This bias may stem from the stylized representation of female characters in Japanese anime and reflects a broader pattern in anime, where female characters are often portrayed with a stronger focus on external appearance rather than internal depth or personality.
- **The number of fields filled in the input directly impacted the quality of the generated art and text.** When more fields were completed, the results were more aligned with the input specifications, and the generated text was more comprehensive. In contrast, when fewer fields were completed, the tool often added extra details not requested.

- **As the creativity level of the prompt increased, the generated results became more detailed, often incorporating additional elements that were not specifically requested.** This variability in creativity led to a wider range of results, with some details added to enrich the generated character.

The tool's ability to handle specific prompts was also noted: if a user emphasized certain aspects (such as the character being a student), the generated art reflected that focus (e.g., school-related outfits), but still adhered to other requirements. However, when conflicting instructions were given (e.g., requesting a character with dark skin while also asking for a "Shonen Protagonist"), the generated image may reflect one instruction over the other, leading to discrepancies.

A common issue is the tool's tendency to generate art from the Japanese cultural viewpoint, which may not always align with other cultural or aesthetic preferences. For instance, requests for a "witch costume" were interpreted through the lens of the Japanese "Mahou Shoujo" style, rather than a European-inspired version, unless explicitly stated.

In conclusion, the experiments demonstrated the system's flexibility and creativity but also highlighted some limitations, especially when dealing with contradictory prompts and cultural biases. It works well for users who provide detailed and consistent input but may produce less accurate results when faced with ambiguous or contradictory instructions.

5.2. User experience

The results indicate that users generally had a positive experience with the character creation system. Key results include:

- **Coherence of character description**: Most users (13 out of 20) rated the coherence between 4 and 5, indicating that the generated descriptions were aligned with the users' expectations.
- **Originality of the generated character**: The originality ratings were diverse, with users expressing varying opinions based on their level of experience. While 9 participants rated it between 4 and 5, 4 rated it lower, between 2 and 3, which may indicate room for improvement in creativity.
- **Aesthetic appeal of the generated images**: 15 out of 20 users found the images visually appealing, rating them between 4 and 5. A few users, however, gave lower ratings, indicating a need for potential refinement in image quality.
- **Image representation accuracy**: Most responses (16 out of 20) rated the accuracy highly, with ratings between 4 and 5, confirming that the images largely reflected the users' descriptions.
- **Generation time satisfaction**: The majority of users were satisfied with the time taken for the character generation, with most rating between 4 and 5, showing that the system is efficient in producing results within an acceptable time frame.

5.2.1. Participants Comments

Five participants provided suggestions and critiques in the open-ended text section. Their feedback (shown in Table 3) highlighted opportunities for improvement, such as the

inclusion of additional character trait options, refinement of the prompt phrasing within the interface, and the ability to modify the character's backstory provided text.

Table 3. User feedback received. Each row is feedback from a diferent user.

Excellent tool, I was very satisfied with the characters it generated. It was also super easy to use compared to other sites I've tested, so I didn't waste any time on my creations.
The main appeal of the site is its ease of use, combined with the good quality of the generated text and images. I wished this easiness-of-use was expanded further, allowing users to directly edit the story or the prompt, with the images adapting accordingly to the changes.
Very simple and quick to use, and the AI-generated art is very impressive.
Super intuitive platform. I was creating high-quality characters and stories within minutes. The experience is so seamless that it stands out from other tools I've tried.
The AI art and story generation are top-notch, it's impressive how effortlessly everything comes together, making the creative process both fun and efficient. It'd be great if the text generation features were even further developed. Perhaps allowing users to kickstart a full story or an interactive narrative based from the generated content.

The results suggest that the system performs well in terms of generating accurate and visually appealing characters based on user input, and was well-received by at least 25% of participants. However, there is room for improvement in the areas of **originality** and **creative flexibility** for certain user profiles. The system works effectively for both beginners and experienced users, with most users expressing satisfaction with the generated results and the overall process.

6. Conclusion and Future Work

In this paper, we introduced *GenAI ConceptLab*, a web-based tool for generating anime character concept art and textual descriptions using generative models. The system simplifies the early ideation process by allowing users to input a small set of character traits, producing both visual and narrative outputs automatically, without the need for manual prompt crafting.

Experimental evaluations showed that the system is capable of generating aesthetically pleasing and stylistically consistent content. However, limitations emerged when handling long or conflicting user-provided traits, which occasionally resulted in mismatches between the image and the description. Still, user feedback reflected high levels of satisfaction with the overall visual quality and character coherence.

Several key lessons emerged from our study. First, the system works best when users provide concise and well-aligned traits. Second, although the generated images are generally coherent, occasional inconsistencies between visual and textual content point to the need for improved prompt control mechanisms. Third, customization features were positively received but remain limited in scope, suggesting a demand for more granular control over design elements. Lastly, some stylistic repetition was observed, indicating potential model biases that should be addressed in future iterations.

To overcome these challenges and further improve the system, future work will focus on (1) enhancing the prompt generation process to improve text-image consistency, (2) expanding user customization options, (3) integrating additional generative models to support stylistic diversity, and (4) optimizing system performance and responsiveness. Additionally, new user studies, particularly involving professional artists, will be conducted to deepen the evaluation and ensure the tool's relevance in creative workflows.

References

Esser, P., Kulal, S., Blattmann, A., Entezari, R., Müller, J., Saini, H., Levi, Y., Lorenz, D., Sauer, A., Boesel, F., et al. (2024). Scaling rectified flow transformers for high-resolution image synthesis. In *Forty-first international conference on machine learning*.

Grand View Research (2022). Anime market analysis by product type (tv shows, movies, merchandising, others), by revenue (usd billion), and segment forecasts, 2025 - 2032. https://www.grandviewresearch.com/industry-analysis/anime-market. Accessed: 2025-02-28.

Holara (2023). Holara. https://holara.ai/.

Hollywood Reporter (2022). Japanese anime: The world's most bankable genre. https://www.hollywoodreporter.com/business/business-news/japanese-anime-worlds-most-bankable-genre-1235146810/ Accessed: 2025-02-28.

Intelligent Systems (1990). Fire Emblem. Digital.

Midjourney, I. (2022). Midjourney. https://docs.midjourney.com/.

Monolith Soft (2010). Xenoblade. Digital.

OpenAI (2021). Dall-e: Creating images from text. https://openai.com/index/dall-e/.

OpenAI (2023). ChatGPT: OpenAI Language Model. https://openai.com/chatgpt.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., e Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10684–10695.

Sizigi Studios (2019). Waifu Labs. https://waifulabs.com/.

Spellbrush (2022). niji journey. https://nijijourney.com/home.

Spotlight Report (2024). The rising popularity of anime-themed games.

Square Enix (1987). Fire emblem. Digital.