AIDM7350

AI and Digital Media Workshop Group Project

**From Text to Tune: Unveiling the Power of**

**Suno AI in Song Generation**

Group members

CHEUNG Man Lok (22467653)

MAN Tsz Shan (22467661)

WU Chun Yiu (22457763)

**Introduction**

As the capabilities of AI continue to evolve, Suno AI exemplifies the potential of generative AI (GenAI) technologies by transforming text prompts into fully composed songs. From creating entertainment content to inspiring artists with new ideas, Suno AI has garnered attention and adoption across diverse sectors of society.

To explore the use of GenAI in songwriting with a focus on Suno AI, an educational website is launched, diving into Suno AI's development, applications, and community reception, alongside ethical considerations. This report aims to describe the emerging AI product and how the educational website uses different formats to present the rich content.

**Structure of the educational website**

The website consists of eight sections, including a homepage that provides a brief introduction to the four co-founders of Suno, a section introducing Suno and other GenAI music platforms, a section teaching how to use Suno to generate songs, a section showing recent and representative cases of songs generated by Suno (which is divided into three categories: entertainment, inspiration, and content creation), a section discussing the opportunities and risks brought by Suno, a section detailing how music industry practitioners perceive Suno, and a resources and contact page. On each page, there is a top menu bar with the eight tabs ‘Home’, ‘About’, ‘Tutorial’, ‘Cases’, ‘Controversy’, ‘Reception’ and ‘Resources’ and ‘Contact’, allowing visitors to navigate different webpages easily.

In addition to text content written in HTML format, CSS and JavaScript are employed to create visually engaging web pages (designing the color tone, layouts, fonts, and overall design styles) and increase interaction with users. The website uses plenty of online resources, particularly videos from YouTube to illustrate recent song examples created by Suno and other GenAI music platforms. It will arouse users' interest in browsing the internet.

**The emergence of Suno AI**

The landing page of the website shows a short tagline to introduce Suno AI, followed by a row of slider with cards that shows the profiles of the four co-founders of Suno AI. Each card displays the name, photo, short bio and optionally a quote by them.

The four co-founders worked together at Kensho Technologies before initiating Suno AI. They are Mikey Shulman, Martin Camacho, Georg Kucsko, and Keenen Freyberg, who worked as Head of Machine Learning, Chief Architect, Head of Machine Learning R&D and Head of Strategic Initiatives at Kensho respectively. Kensho was a tech company specializing in machine learning and analytics for the banking industry and the intelligence community, and was later acquired by S&P Global. Shulman and Camacho are both musicians who used to jam together in their Kensho days.

In an interview by *Rolling Stone*, Shulman stressed that by developing a music generating AI, they are not trying to replace artists. They in fact were trying to get more people to engage with music, targeting a billion more people. He believed that people would be more focused on creating, developing much more distinct tastes, which is obviously good for artists. The vision that we have of the future of music is one where it’s artist-friendly.[[1]](#footnote-0)

More background of Suno AI is detailed in the What is Suno? subtab under the About tab. In the past years, generative AI has been widely used and popular in various industries, such as ChatGPT and Midjourney, to create images for commercial use and personal exploration. However, there have only been a few discussions about generative AI in making music and audio. It could be a challenging topic for Artificial Intelligence (AI) in dealing with music as the amount of information found in the structure of music is higher than traditional text-based elements.

Suno AI is a newly generative AI platform for music, with the power to enable users to create with imagination with text by prompting based on its own trained model behind, just like the “ChatGPT for Music.”[[2]](#footnote-1)

Although Suno AI does not reveal the principle of its processing behind it, machine learning and deep learning are commonly used in the AI generation of music production. For instance, Recurrent Neural Networks (RNN) can be used for the interpretation of input sequences to generate the outcomes, while Long Short Term Memory (LSTM) networks can capture the thematic development of the music piece.[[3]](#footnote-2)

With the additional help of Convolutional Neural Networks (CNNs), feature extraction can be executed in the music training datasets. The model is trained with numerous datasets to learn the patterns of the musical structures, including the compositions, genres, and styles in different types of music.

**What makes Suno AI so powerful?**

The Semantic Recognition Power - In contrast to previous generative AI platforms primarily relying on pattern recognition and statistical modeling, Suno integrates natural language processing (NLP) methods to understand textual prompts.[[4]](#footnote-3) This empowers Suno to compose music that maintains structural coherence and resonates thematically with the user's input.

Extensive Customization - Suno provides high customization and interactivity, allowing users to input specific text prompts and adjust various parameters to tailor the generated music to their preferences. Users can define genre preferences, mood, instrumentation, and other musical aspects, resulting in a deeply personalized and immersive journey.

Seamless Integration of Text and Music - Suno seamlessly integrates text and music generation processes, ensuring the generated compositions reflect the intended themes and emotions conveyed in the input text. By dynamically adjusting musical parameters based on the semantic content of the text, Suno produces compositions that feel cohesive and expressive.

Scalability and Performance - It efficiently handles enormous volumes of data and computational tasks. Suno maintains a high-performance level to deliver fast and convincing results and ensure the growing demand for users in the music production industry.

**Steps in generating music with Suno AI**

Under the tutorial section, we demonstrate the process of using Suno to make a song step by step with real examples. Regarding the creation, Suno offers two modes for the users: the Default mode and the Custom mode.

In the Default mode, users only have to write the description of the song, in a similar way to prompting in ChatGPT, MidJourney, etc, but in the form of making songs. User can choose the specific version of their model to generate the results. The song could be created easily through Suno’s model behind, and it names the song title for you.

On the other hand, Custom Mode offers more adjustable parameters for the creator with more ideas. It enables the users to enter their lyrics in the song. By choosing the style by inputting keywords, users can set the vibe, mood, and voice according to their keywords to generate their tailor-made genre. Then, after naming the title and the mentioned details, Suno integrated all the parameters to create the song in a few minutes.

For both modes, there is an option “Instrumental” for choosing whether the output is composed only with lyrics and vocals or only music.

After your creation, you can view your own pieces in the Library on the Suno platform or explore other users' works in the community in Suno.

**Applications of Suno AI by online users**

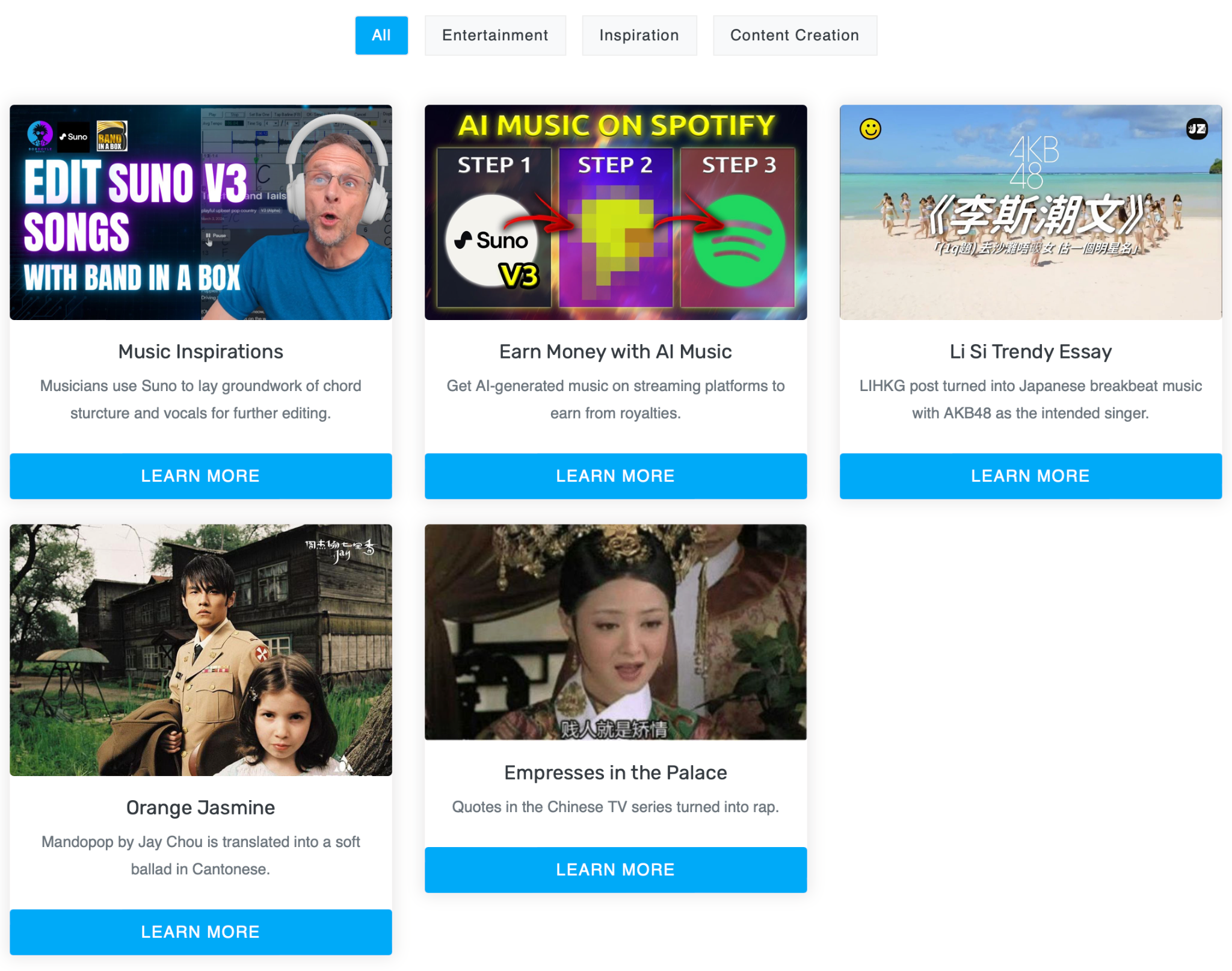
Suno AI has garnered significant attention as soon as its launch of V3 in March 2024. It then generated widespread adoption among online users globally, serving as a versatile tool for various applications in music creation, content production, and entertainment.

Musicians found Suno mesmerizing to draw inspirations for further editing and creation. The YouTube channel Bob Doyle Media showed generating a "playful upbeat pop country" before converting it to editable music to be imported in other music editing programmes.[[5]](#footnote-4) The YouTuber commented that Suno could be used for tune inspiration. "If you lay a groundwork in Suno a chord structure you like, even some vocals that you like, you can export that, bring it into Band in a Box (music editing tool) and start using instruments to make whatever tweaks you want."

Even for people who are not a music professional, they started making money leveraging the use of Suno AI. AI Controversy, a YouTube channel, uploaded a video showing how to publish music generated by Suno on online streaming platforms, such as Spotify, in the hope of generating profits from royalties when a number of users listen to the song.[[6]](#footnote-5) He also used Suno to develop instrumental music as background music for his videos.

Suno AI is also used for entertainment. In Hong Kong, LIHKG, a popular online forum, discussed the use of Suno AI in a series of posts consisting of more than 1000 replies.[[7]](#footnote-6) The focus of the AI product soon shifted to experimenting music with "trendy essays" in the past as the lyrics. One of the outputs is "Li Si Trendy Essay", of which the content came from a LIHKG post in 2019, referring the TVB Li Si actor Chan Kwok Pong as the answer to an IQ question.The creator of the music translated the post content into a set of lyrics with intro, verse, pre-chorus, chorus and outro. With the use of AKB48 music video clippings as visuals, the text became a lively Japanese breakbeat.

On our website, the Cases tab is linked to a gallery of outputs created by online users with Suno. Each example of application is explained in detail by clicking through the corresponding case. They are categorized into three groups, which are Entertainment, Inspiration and Content Creation and can be filtered by using the selection bar. Apart from the above mentioned examples, two other cases that generated a Cantonese version of Orange Jasmine and with quotes from the TV series Empresses in the Palace are also included.



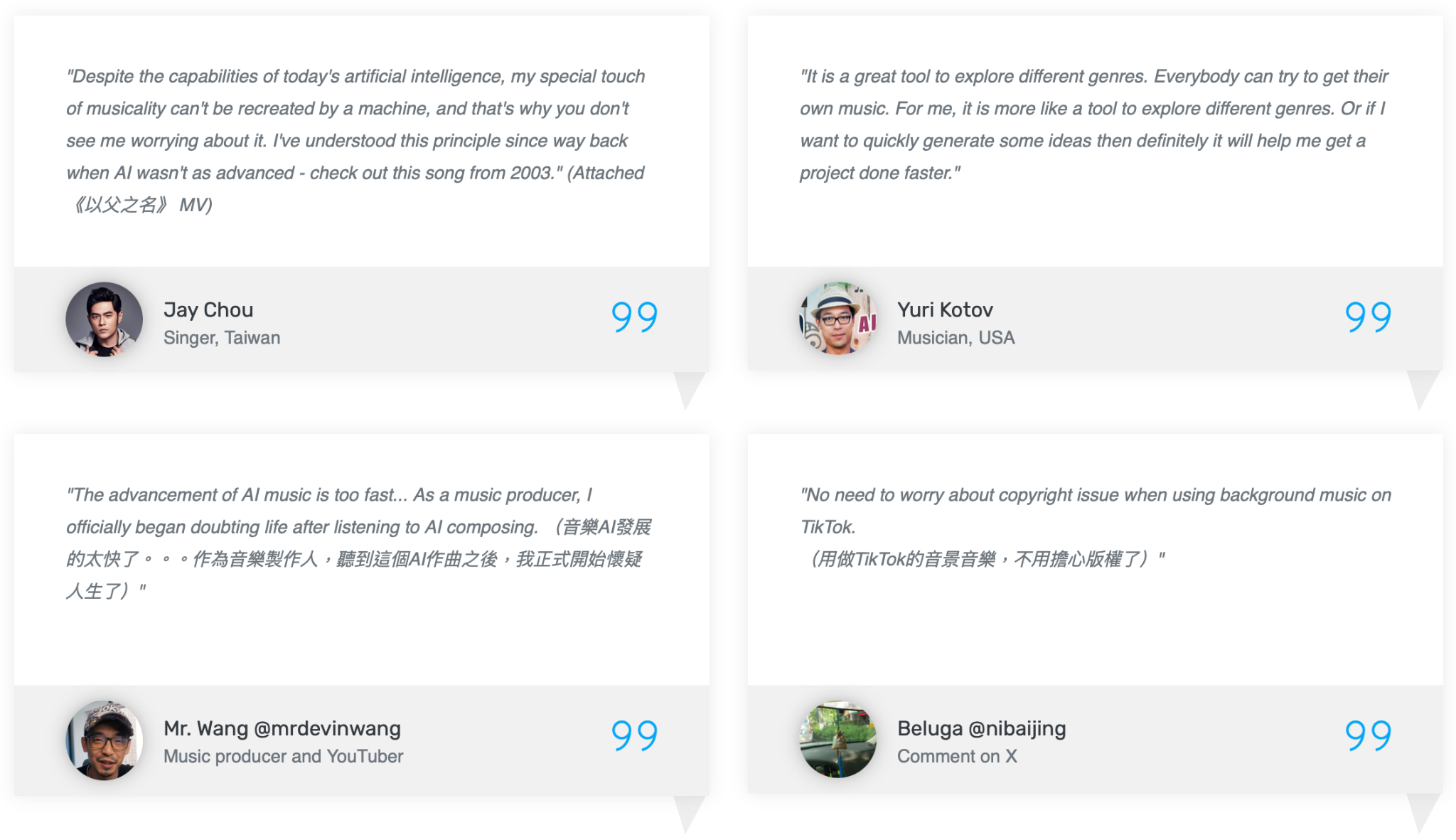
**Receptions of music industry and the internet**

The reception of Suno AI within the music industry and among online users varies, reflecting a spectrum of perspectives on the intersection of artificial intelligence and music creation.

Jay Chou, the "King of Mandopop" from Taiwan, emphasizes the irreplaceable quality of human musicality, asserting that his personal touch cannot be replicated by machines. Despite acknowledging the capabilities of AI, Chou remains confident in his unique artistic expression, suggesting that AI-generated music cannot capture the essence of human creativity.

The AI product also stirred up heated debate online. Many YouTubers expressed their views online. Musicians like Yuri Kotov view Suno AI as a valuable tool for exploration and idea generation. Some Tiktokers are happy to find Suno to generate music without copyright infringement issues. However, Mr. Wang, a music producer and YouTuber, is worried that the rapid development of AI will replace their role.

On the website, these comments are kept in their original quotes with the name and background in the "Receptions" section. They are shown as conversation boxes for each commenter to make them more lively and relatable to readers.



**Other GenAI platforms are innovating and changing the music industry**

The landscape of generative AI in the music industry is flourishing, with various tech companies rolling out various platforms in music and audio generations in recent months. In addition to Suno, it is worth to highlight other GenAI platforms which have a huge potential in revolting the future of music industry. In particular, MusicFX by Google, Stable Audio by Stability AI, and OpenAI's Voice Engineer, each of them has a distinct features and potential applications in music production.

MusicFX, powered by Google and launched in February 2024, is a generative AI text-to-music platform that allows users to create tunes up to 70 seconds in length and music loops. It is available through Google’s AI Test Kitchen. Users can add and adjust musical prompts (e.g. instruments, speed or style) through text prompting. However, MusicFX can only generate the output in melody without vocals for protecting original artists’ voices and styles, which is a distinctive difference compared with Suno.

Stable Audio 2.0, powered by Stability AI and published in April 2024, is a new model goes beyond text-to-audio to include audio-to-audio capabilities. Stable Audio 2.0 not only have text-to-audio function, but also allowing users to upload audio samples and transform into a wide array of sounds through text prompting. The model also expands sound effect generation and style transfer, providing artists and musicians more flexibility, control, and an elevated creative process. Differs from Suno’s text-to-audio generation, audio-to-audio allows provide a great convenience for music composers to convert the melodies in mind into a song in anytime and anywhere.

On the aspect of AI vocal generation, OpenAI announced a vocal model called Voice Engine in March 2024. By providing a text input with a single 15-second audio sample, the model can generate natural-sounding speech that closely resembles the original speaker. It also allows for the translation of a speaker's voice into multiple languages to reach a global audience. The translated audio preserves the native accent of the original speaker; for example, generating English with an audio sample from a French speaker would produce speech with a French accent. It is different from Suno that it can translate the music pieces into different languages. Although the platform is not currently designed for the music industry, there is high potential that it could revolutionize the field of music composition in the future. It is possible that singers can create a song that in any language with their vocal if the function of Voice Engine combined with other GenAI tools.

**Risks and opportunities of GenAI music platforms**

A major benefit of generative AI lies in its ability to speed up the creative process. Instead of starting with a blank page or from scratch, users can now engage AI to generate preliminary ideas and suggestions (e.g. text-to-audio or audio-to-audio) that propel creative endeavors forward. This makes generative AI particularly attractive to those in music and creative fields. Additionally, for both individuals and companies, generative AI contributes to significant time savings and reduces production costs.

The embracement of GenAI with reality also explore a new field in music creation. Riding on the AI hit, the YouTube content creators used AI to generate voice of Hong Kong cult icon Wan Kwong before integrating it with two Cantopop songs. The creative work has attracted a lot of attention on the platform and simulate Wan Kwong to create a song “Dear myself” by both his AI and himself, noting that the creativity is worthy of praise. It recorded a nearly 2 million views with around 48,000 likes on Youtube platform[1], and Wan Kwong became one of the top 5 most favorited male singer in 2023 Ultimate Song Chart Awards Presentation.

Besides new creations, public are also willing to accept work produced by generative AI for individuals who have passed away (e.g. AI Deng Li Jun, AI Wong Ka Kui) or no longer sing for a long time (e.g. AI Stefanie Sun) as a means of remembering them. Millions of views have been accumulated in Youtube.[1] In particular, Stefanie Sun has expressed her thoughts on the transformative power of AI across various industries, including music. She has also reflected on the importance of individuality and authenticity, suggesting that remaining true to oneself will continue to be significant in a world increasingly filled with AI-generated content.

While AI literacy is opening new opportunities and possibilities in the music and creative industries, the use of GenAI in these sectors has sparked significant concerns about originality and copyright issues. One pressing question is whether the use of publicly available data for training purposes without prior consent is justifiable or not. A numerous lawsuits and articles accused technology companies of misappropriating artists’ work to train their AI models in 2023 alone. For example, a high-profile case in the US saw the New York Times suing OpenAI and Microsoft for copyright infringement. To avoid the lawsuit, some GenAI music platforms, such as Suno, do not disclose the materials and training process of its AI model.

The products of GenAI technology also challenges existing copyright law. First, the law must determine whether to grant authorship and ownership rights to expressive works created with the assistance of GenAI models. Suppose we can identify a certain part of song is composed by GenAI models, is it possible for us to define the originality of a song by drawing a simple line that a song which is composed by GenAI model below a particular threshold? Secondly, the law must resolve whether works created with the help of GenAI models constitute copyright violations if the models were trained using copyrighted materials.

Another controversial question is whether the music piece created by generative AI is truly considered art. Art is not only defined by creativity but also by its emotional connectivity with humanity and the ability to learn from experience. “That sense of interplay, or the ability to react in the moment, is something that artificial intelligence can’t reproduce.” Artists, as well as general public, may need to rethink about a product from GenAI model, which is a mixture of training samples with some random factor, should be considered as an artwork.

The accuracy of the GenAI product remains a doubt. Some of the elements of minority cultures could be misrepresented from the GenAI products given a limited sample and insufficient model supervision. For example, Google Gemini AI generated an illustration of 1943 German solider with black and Asian faces[1], which fails to produce historically accurate results in an attempt at racial and gender diversity. Similar cases could be happened in music generation if users do not have enough knowledge to assess the accuracy and the appropriateness of the GenAI products.

Similar to other GenAI products, the displacement of artists and music creators by Suno and other music GenAI models is a key concern within the industry. In the UK, Equity, a trade union representing 50,000 performers and creative practitioners, launched its Stop AI Stealing the Show campaign to lobby the government to update the law and better protect artists’ livelihoods. Many professionals from the music, literary, and television industries are particularly worried about how AI platforms, capable of plagiarizing and reproducing their work without compensation, will affect their living and career.

Techniques used when presenting the knowledge on a website

Conclusion

Workload distribution

|  | CHEUNG Man Lok (22467653) | MAN Tsz Shan (22467661) | WU Chun Yiu (22457763) |
| --- | --- | --- | --- |
| Development of topic | ✔ | ✔ | ✔ |
| Research and background understanding | ✔ | ✔ | ✔ |
| Website development | ✔ | ✔ | ✔ |
|  |  |  |  |

1. Brian Hiatt (2024, March 17). A ChatGPT for Music Is Here. Inside Suno, the Startup Changing Everything. *Rolling Stone*. https://www.rollingstone.com/music/music-features/suno-ai-chatgpt-for-music-1234982307/ [↑](#footnote-ref-0)
2. Sabrina Ortiz (2024, March 19). What is Suno? The 'ChatGPT for music' generates songs in seconds. *Zdnet*. https://www.zdnet.com/article/what-is-suno-this-chatgpt-for-music-generates-songs-in-seconds/ [↑](#footnote-ref-1)
3. Alberto Martinez Jr (2024, January 4). The Rise of AI-Generated Music: What It Means for Artists. *Flourishprosper.* https://flourishprosper.net/music-resources/the-rise-of-ai-generated-music-what-it-means-for-artists/ [↑](#footnote-ref-2)
4. Henceforth Solutions Pvt Ltd (2024, Apr 9). Unleashing the Power of SUNO AI in Music Composition. https://www.linkedin.com/pulse/unleashing-power-suno-ai-music-composition-henceforthsolutions-8qz0c?trk=organization\_guest\_main-feed-card\_feed-article-content [↑](#footnote-ref-3)
5. Bob Doyle Media (2024, March 14). How a "Real" Musician can use Suno v3 AI Generated Music. *YouTube*. https://youtu.be/JCBc3-SNNUY?si=iwxYVIlTyUCvFN4b [↑](#footnote-ref-4)
6. AI Controversy (2024, February 27). Suno AI V3 Will Blow Your Mind! This Is How-To Publish AI Music On Spotify. *YouTube*. https://youtu.be/RIwRC6kQqaE?si=nScUTpIm0cfO179v [↑](#footnote-ref-5)
7. (2024, March 20). 乜原來宜家AI連廣東歌都寫到... *LIHKG*. https://lihkg.com/thread/3655261/ [↑](#footnote-ref-6)