Universiteit Leiden
The Netherlands

Informatica & Economie

What different business models can be identified in the market for AI for music creation?

Roan Maarten Otten

Supervisors:
Peter van der Putten & Max van Duijn

BACHELOR THESIS

Leiden Institute of Advanced Computer Science (LIACS)
www.liacs.leidenuniv.nl

13/08/2024

Abstract

With the rapid advancements in AI technologies, AI-based tools for music generation and production are becoming more common and influential in the music industry. Therefore, this thesis investigates the various business models emerging within the market of AI-created music.

In order to perform this research, a literature review has been conducted with qualitative market research to determine the state of the art. The research identifies key capabilities of these AI tools and their impact on the process of music creation and production.

To identify, compare and relate emerging business models, we developed the Unified Model canvas, a framework based on the Business Model Canvas and the Lean Model Canvas. The framework is validated by applying it to different categories of AI tools for music production. Two different business models appeared: The Lazy Mans Tool and Control Freaks. These business models represent two ends of the spectrum, with Lazy Mans Tool focusing on a large customer segment and quick music production, whereas the Control Freaks attract smaller customer segments of professionals and offer tools that are used to augment the artistry of the user, emphasising the process rather than the result.

The findings reveal distinct business models and provide insights into how AI is reshaping the music industry. The framework can be used to compare different brands in a meaningful and efficient way and to find similarities and dissimilarities. New entrants in the market can use this to structure their business model and to position themselves in the intended market. This research contributes to understanding the commercial potential and strategic implementations of AI in music creation.

Contents

1 Introduction

Music has been around for a long time. Humans make music using several types of instruments and sometimes in combination with their voices. In the 20th century, lots of breakthroughs happened in the way people experience music as well as the process of music creation due to digitalisation [Bar12].

Artificial intelligence (AI) has been around for the past decades, with the current breakthroughs in its generating purposes resulting in the recent spike of interest. This has led to AI becoming a vital part of our society with its implementation spreading in multiple industries. Computational creativity together with human interaction has been used in creative fields like songwriting [HKNR+20], music composition[STC21] and design [CRH+23].

Various institutions are researching different applications of AI on specific elements of music creation like Singing Voice Synthesis [Dai24] or algorithms like Generative Adversarial Networks[YZWY17] in order to understand rhyme and creative sequences. Since the underlying tools and knowledge to create music were available, it was a matter of time until this found its way to the commercial market. Examples of research projects that have been conducted where its generating purposes have been taken closer to commercial use are MusicLM[Fra23] and Jukebox [Sut20], but they lack the quality and long-term coherence of more such as Suno AI and Udio Beta. In April 2023, the world was taken by surprise when a song was taking over social media platform TikTok. What seemed to be a new song by the artist Drake [Cos23] in fact, turned out to be created by AI. This was the first time AI gained public interest at this scale by generating a song with the timbre of the voice of Drake or any other artist.

Since then, a lot of research has been done and improvements have been made to the tools. With the recent innovations, a lot of players entered the market. These companies differ from each other in how they use the co-creativity between humans and AI to aid in music creation. With these creative purposes of AI becoming more popular, this thesis will focus on the impact of generative AI in the music industry. Almost all companies with existing tools have implemented upgrades and additional features during this thesis research. Not only that, but new competitors with even better models and unique customisations joined the market.

This leads to the following research question: What different business models can be identified in the market for AI for music creation?

To answer the research question, the state of the art in this industry is identified with compelling arguments and examples of research projects and commercial tools. Then a unified framework is developed with aspects from the Business Model Canvas[OP10] and Lean Model Canvas[Mau12] that can be used to profile existing and future tools. The final part of this research validates the framework against a small sample of different types of tools.

The remainder of this thesis is organised as follows:. Section 2 discusses the related work to conduct this research. Section 3 describes the methods used to conduct this research. Section 4 provides a qualitative market research and displays an overview of existing research projects and commercial tools with their key capabilities. Section 5 applies the Unified Model Canvas to the AI created music industry with clear examples provided for each section. In section 6 two different kinds of tools are applied to the model in order to validate its usability. Section 7 describes the results, limitations and offers options for further research.

2 Background and related work

With the recent interest in the intersection of creative industries such as music creation and AI as described in the introduction, this section will discuss several aspects of music creation as well as some background literature.

2.1 Impact of AI on music industry

The ongoing developments of AI and its capabilities have resulted in its influence reaching people's daily lives. People who use music-generating tools benefit from the easy accessibility. Users are able to generate studio-grade music without the knowledge of composing music or playing an instrument; this results in a lowered skills threshold that is needed to create and publish new music.

While this sounds great for enthusiasts, there are a lot of professionals concerned with the current trend. More than two hundred artists have signed an open letter stating to "stop devaluing music"[art]. In this open letter, they inform companies that even though these tools have massive potential if used correctly, they could become a threat to the craft as a music artist. Furthermore, the fast music generation of these tools results in the dilution of music quality that is published. A Dutch newspaper addresses this problem by saying that people can buy a SunoAI subscription for ten to thirty euros[Sun] and then compose thousands of songs for the average cost of two or three euro cent per song[Hij]. This while the training data of these models mostly happens on copyrighted music without the artist's approval or knowledge.

The United States has introduced a bill called the Generative AI Copyright Disclosure Act that forces brands to send the training data of AI products with generative properties to the Register of Copyrights. The objective is not to ban the use of copyright completely but rather to create a duty to report the data sets that are usually kept private [RE].

The impact of AI on the industry is therefore irreversible, and our society should focus on the correct usage of the tools. This should result in enthusiasts of music and these kinds of AI tools working alongside human artists rather than replacing them with these tools.

2.2 Key AI capabilities

This subsection will explain several terms and formats used in the process of music creation. These will be handled shortly and described in greater detail over the course of this thesis.

The process of music recording has seen a lot of change over the last decades. A major invention in the 1950s was multitrack recording, which was popularised by bands such as the Beatles and the Rolling Stones. This allowed for bands to record different instruments and vocals simultaneously. Later improvements of these analogue recorders and the development of digital recording changed the way music would be produced drastically. Recording music without analogue mediums allowed people to record digitally and, with the help of MIDI, communicate with other devices. An example of this is a Roland keyboard that is plugged into a Yamaha synthesizer. These MIDI files of instruments could later be edited in a Digital Audio Workstation (DAW).

These generative tools can record audio for a single instrument or create the whole song in one time. The first type of tool creates instrument tracks where the output file type is MIDI. The second type of tool generates raw music, which means that full audio is created. This allows for less flexibility if the user wants to edit something about the song.

The first format that will be handled is MIDI. MIDI, which stands for Musical Instrument Digital Interface, was developed to facilitate devices such as synthesisers to communicate with computers, for instance. This standardised format of MIDI messages was crucial in the development of music creation[Gib13]. DAWs are software programs that run on computers that allow users to manipulate recorded music. The use of special effects such as reverb and instrument plugins offers sound engineers and artists endless possibilities[Wika].

Now that these terms are explained, a short description of the four different types of workflows behind music creation is handled. These generative tools work differently because of the different input and output files. Several examples are text-to-music, audio-to-music, text-to-MIDI, and audio-to-MIDI. These types will be explained briefly.

First, text-to-MIDI systems generate MIDI files with text as input. The user gives a text prompt containing information on the desired output. The output is versatile in the way it can be manipulated easily in a DAW. A MIDI format gives the artist a lot of freedom in the process of music creation due to the capabilities of a DAW.

Audio-to-MIDI is the next example and differs slightly from text-to-MIDI. An audio file is used as an input, and the corresponding output is a MIDI file. A useful use case for this is when a user wants to split the stem of the piano from a full audio fragment, or the other way around, where a user hums a beat and the tool creates a MIDI file for an instrument. Just like the text-to-MIDI, this output can be changed in a DAW.

The next examples describe the full music process and take a step further by generating complete audio tracks. Text-to-music is used in systems that take a text prompt and generate a full song based on the information. This often includes both instruments and vocals and offers fewer options to manipulate certain settings for instruments or voice separation in the music creation process without having to regenerate the output as a whole. Therefore, these raw audio systems are tailored for beginning artists or users that are more interested in the output rather than the process of creating music.

Finally, audio-to-music takes an audio file and generates full audio based on that file. This can be useful, for instance, if a person wants to generate a song based on a specific rhythm.

3 Methods

This thesis uses a literature research method with qualitative market research to gain a better understanding of the state of the art in research and in practice. Besides the market investigation, a research by design study has been done by developing a framework that uses elements of both the Business Model Canvas and the Lean Model Canvas. Due to the vast changes and developments in this industry, validation by market research is limited. The Business Model Canvas and the Lean Model Canvas have been studied, and the limitations of both models have been taken into account in the creation of the new unified model. To verify this model, it is applied to the music industry to differentiate several business models on their AI capabilities in regards to the co-creation of music alongside humans.

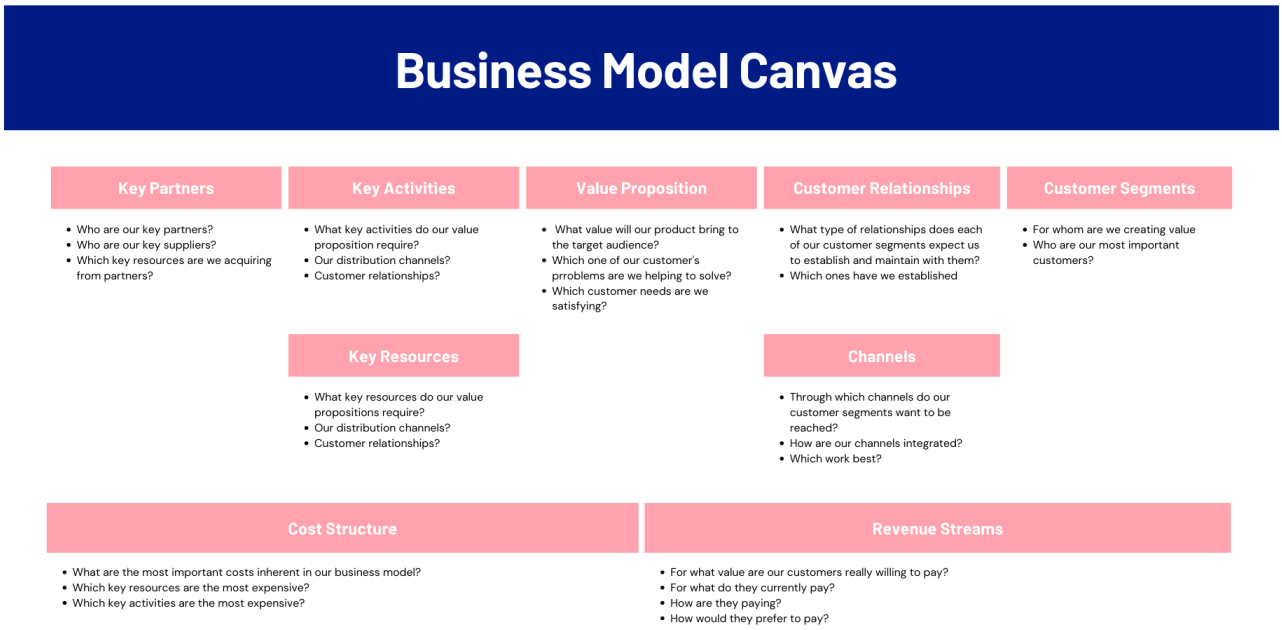

Figure 1: Business Model Canvas, source text (from: [Cas])

3.1 Business Model Canvas

A Business Model Canvas is a diagram that should represent the business model of a company. Designed by Osterwalder and Pigneur, it captures how a company creates, captures, and delivers value to its customers[OP10]. It is divided into 9 sections: key partners, key activities, key resources, value proposition, channels, customer segmentation, customer relationship, cost structure, and revenue streams. The following sections will discuss these aspects briefly with supportive examples to gain a better understanding.

3.1.1 Customer segmentation

A firm can't survive without its customers; thus, it's essential that it understand who its target market is and who its potential customers are. Groups of people or entities that share the same attributes and behaviours belong in the same segment. This results in one or multiple user segments all requiring something different from the product or service the company supplies. A brand can cater the product or service to a segment's specific needs if the company decides to take a profitable and desired segment on as a client. The trade-off between a niche market or mass market can help with certain brand positioning compared to competitors in the same branch[Mau12].

3.1.2 Value proposition

A value proposition is the way a company delivers its value for a customer. It combines products or services to satisfy the customer's needs. The value propositions should be catered to a specific

customer segment. It provides value to customers by being innovative or able to outperform the competition. With novel solutions to a problem, new opportunities and new needs arise, which results in the continuous development of value propositions. Another way to have an edge over the competition is simply by outperforming the competition. Simple examples can be taken from the tech branch with solutions providing more powerful computing power or a more efficient operating system. Customisation is the last example to consider a certain company. The provided product offers a personalised product or service, resulting in a unique fit for the clients needs [OP10].

3.1.3 Key activities

Key activities are the most important activities a company has to do to successfully apply its business model. Like key partners and key resources, key activities differ depending on the value proposition and, to an extent, the business model of the firm. For software companies, implementing and developing software are important activities, while consulting firms focus on problem solving. Several kinds of activities are problem solving, production, administrative, and management.

The category problem solving is based on a service that a company offers. When a patient comes to the hospital with an injury or problem, the business model of the hospital is to provide the service of operating the patient. These business models require knowledge of the employees and often need continuous development in order to solve a problem.

Another example of a key activity is production. This category is straightforward in the way a business model requires a division to create a product. Once the required knowledge is applied to a specific field, manufacturing the product is the key activity. BMW makes vehicles, but after the framework and parts of the car are developed, producing and assembling the car are the only next steps in order to create value for the customer[OP10].

3.1.4 Key resources

Key resources are necessary to perform the key activities and create the value proposition. Depending on the focus area of the company, some types of resources are more important than others. The four main types of resources are: human resources, intangible resources, physical resources, and financial resources.

All companies require human resources, but the requirements differ per company. A consultancy requires extensive knowledge from its employees to solve the customers problems, while a local supermarket requires a grocery clerk to help customers with product shopping. The next type of resource is intellectual resources such as patents, copyrights, knowledge, and brand image. Intangible trade secrets are a crucial asset for a company such as ASML, whereas Zalando personalises shopping based on specific knowledge of the customers that is stored in a database. Physical resources are all types of physical assets, such as storage buildings, manufacturing facilities, and equipment. A supermarket has to have a lot of physical assets, like storage areas, while a consulting agency solves problems and requires human resources for their knowledge. Finally, financial resources are necessary to meet the financial needs of a company, such as investments, certain bonds to expand a business, or licenses[OP10].

3.1.5 Key partners

Partnerships are essential for a company to make its business model work. There are several kinds of partnerships, and their function varies from risk reduction to acquiring resources.

The first example relates to optimisation and economies of scale. It is illogical and suboptimal for a company to perform all activities by itself. A reason to outsource aspects of the production process is to reduce costs. The company that is hired by the outsourcing brand could benefit from economies of scale and operate on a larger scale.

Another reason to forge alliances is to reduce risk. Especially in a competitive environment where uncertainty plays a big role. Companies that compete in a specific field could be partners in another one. As an example, several companies contributed to the development of the 5th generation of cellular networks (5G). While these companies have worked together to make this available, they compete in the market of electronic devices that use 5G, smartphones, for example[OP10].

3.1.6 Customer relationships

The different ways a company can have a relationship with its customers are described in this section. The relationship can vary per customer. Some require personal assistance, while others benefit more from an automated service or through an online community. When acquiring new hardware for an office, personal assistance from a real person can be beneficial in the sales process and result in an increased user experience. On the other hand, with a simple reminder mail to plan a new dentist appointment, an automated assistant that is available at any moment can be a more fitting solution.

The way a company handles these relationships influences the user experience and their loyalty. A satisfied customer remains loyal for a longer period of time, resulting in a higher value and profit for the company in the long run. Besides customer service and retention, acquisition is an important aspect. The acquisition of customers can be done through commercials or word-of-mouth in the case of a positive relationship with the brand and is important to grow as a company [OP10].

3.1.7 Channels

The section channels describes how brands communicate with customers in order to deliver the value proposition. Besides delivering the proposition, channels have the function of reaching new customers, increasing brand awareness, and offering aftercare post-purchase. There are five distinct stages. These are as follows: awareness, evaluation, purchase, delivery, and after sales, in chronological order. Each channel can cover one or more phases.

Several examples of these channels for the first stage could be a website where products and their descriptions are given as well as a chat bot where customers can ask questions if they want more advice on a product. Other forms of channels could be advertising on public transport or on other websites of partners. Social media could be a final example of this phase. A channel that is focused on the after-sales stage could be a website that offers support and where customers can make appointments for products that are malfunctioning.

The first phase focuses on raising awareness for the company and its product. This can be accomplished through direct channels such as a company website or an advertisement that is owned by the brand or in an indirect fashion that involves third parties, such as partner stores. The next step is to allow potential customers to evaluate the product and determine whether it

satisfies their needs. The third phase is the purchasing and covers how customers can obtain the product or service. After a product is bought, delivery is the next step. If a customer ordered new clothing online through the brand's web shop, delivery is possible by picking up the product at a store or having it delivered to their home address. The final step is after sales and focuses on the support for the customer post-sale. Customer care and feedback collection are two examples to maintain customer satisfaction. While channels can have multiple stages, these customers can also use multiple channels at the same time or separately depending on their needs.[Mau12].

3.1.8 Revenue streams

This section of the canvas represents different ways a company generates cash. A company can earn revenue through recurring transactions or one-time transactions. A clear understanding of what the product or service is worth to the different customer segments is important to maximise revenue. The price of the product or service could be either fixed or dynamic. An example of a factor that influences the price is segment-related; for example, a hotel room can be rented to a loyal customer for a lower fee than a new customer. On the other hand, an auction where the product is sold to the highest bidder is a way of dynamic pricing[OP10].

3.1.9 Cost structure

The cost structure includes all the incurred costs of a company. Companies can incur two different kinds of costs: fixed costs and variable costs. Fixed costs remain the same and are not influenced by the fluctuations in sales. Variable costs, on the other hand, proportionally vary depending on the quantity a company produces. Several examples are the costs to maintain a system, the costs of research & development, and service costs. Whilst inherently a company wants to minimise costs, each one has a different view. On one end, there are cost-driven companies that prioritise keeping the price low, which impacts the quality of the product; cheap flights by Ryanair is a prime example. On the other side of the spectrum are value-driven companies that deliver a high value proposition, like a luxurious restaurant with exclusive food and master chefs as employees[OP10].

3.2 Lean Model Canvas

The Lean Model Canvas is based on the Business Model Canvas, with figure 2 showing an example diagram. Similar to the Business Model Canvas, the Lean Model Canvas is used to gain an understanding of the structure at a glance. This canvas is designed for start-up companies and focuses on the added value of a product. It is concise and offers high-level information about the start-up. It comprises the core elements such as problem, solution, and unfair advantage. The key problems of the target segment are elicited and are written down. The solution element should explain what problems are targeted, and each solution should link directly to one of the core problems. To make sure the product is feasible, the amount of problems and therefore solutions are limited. When the desired solutions are found, unfair advantages should create a barrier for competition to replicate the created solution. Besides being difficult to replicate, it should also be something that cannot easily be bought. Examples of this could be a passionate community or a unique brand experience that a customer is exposed to. The section key metrics includes measurements like Key Performance Indicators (KPIs) that allow the brand to track progression

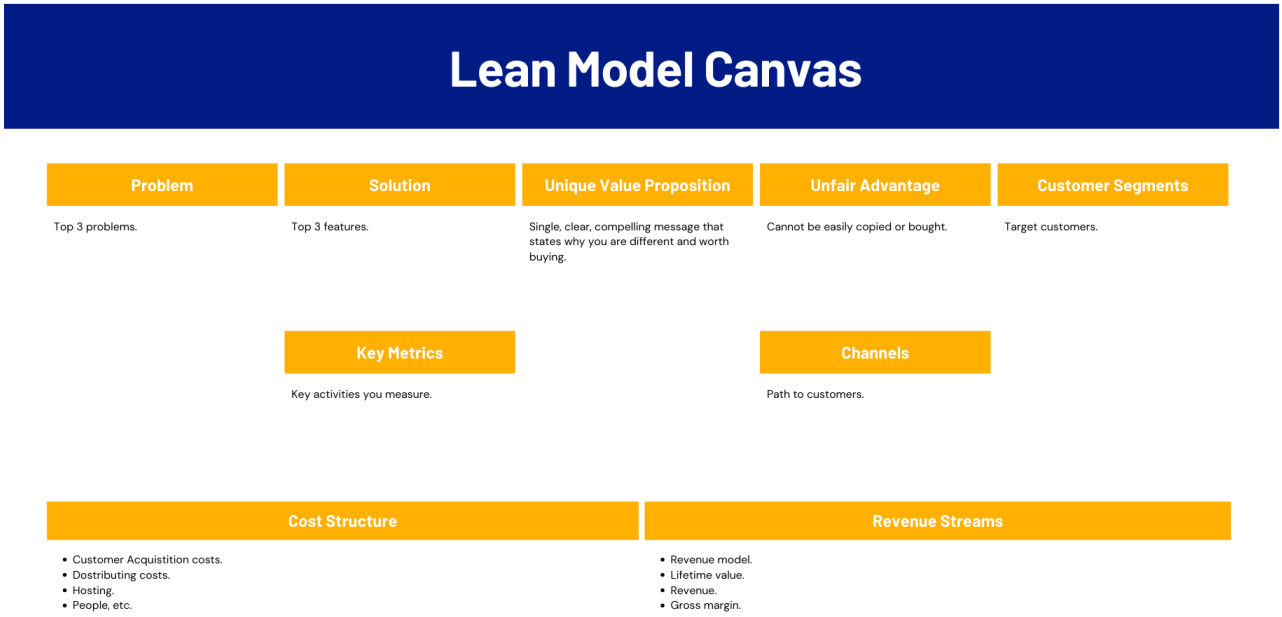

Figure 2: Lean Model Canvas with source text (adapted from: [Mau12])

and performance. The other elements that are included in a Lean Model Canvas are channels, unique value proposition, customer segments, cost structure, and revenue streams. These elements also occur in the Business Model Canvas and will therefore not be discussed in great detail in this section[Mau12].

3.3 Lean Model Canvas versus Business Model Canvas

Both models are created to structure certain aspects of a business, and both share a lot of components that are implemented in both models. The emphasis of the Lean Model Canvas is on startups and relies on novelty and problem solving rather than the daily activities of an existing firm. Important aspects are problem, solution, and unfair advantage, which solicit the problem of a certain customer segment and how the product solves it. Unfair advantage describes how the solution provided by the company is unique and difficult to recreate for competitors[Mau12]. A Business Model Canvas is more useful for an existing firm that prioritises strategy in components such as key activities and key partners, for example. Components like key activities, key partners, and key resources facilitate long-term planning and give a more complete view of the company as a whole rather than a solution it offers for a problem[OP10].

3.4 Unified Model Canvas

The Business Model Canvas and Lean Model Canvas have proven themselves as useful models and are widely used for all kinds of industries. However, both models cannot cover a company as a

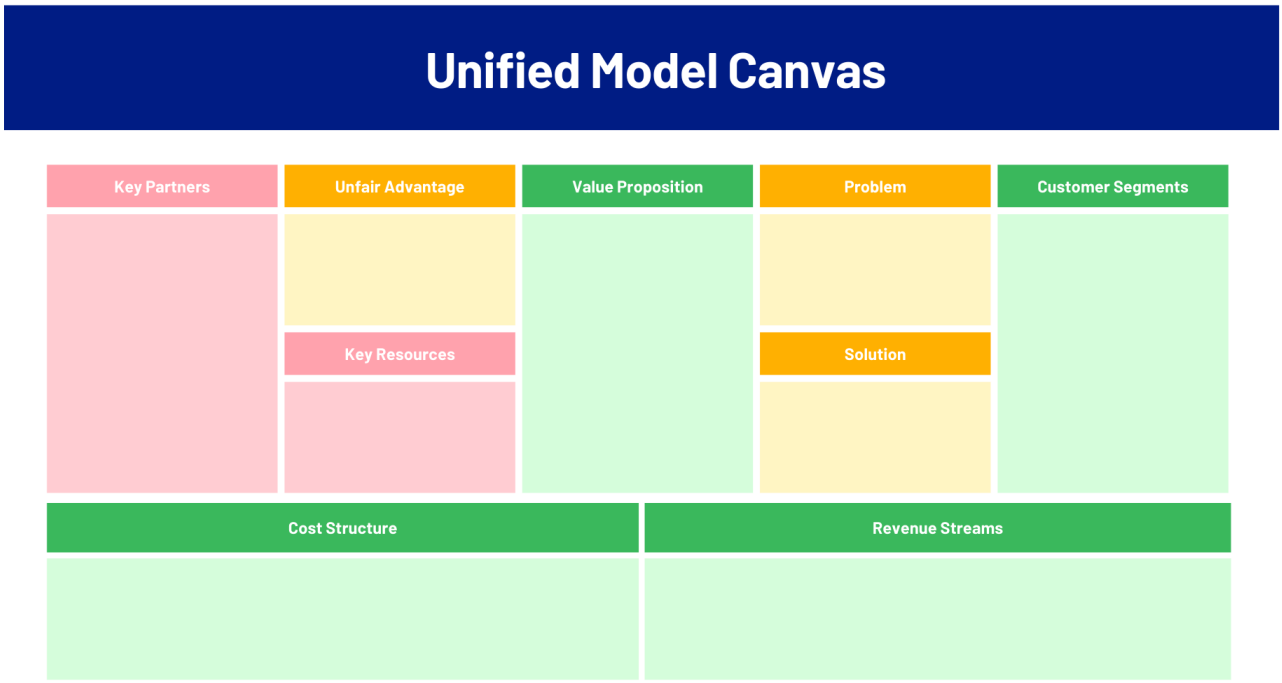

Figure 3: Unified Model Canvas

whole, and limitations do occur. While the Business Model Canvas can be used to support the brand's strategy, it does not take into account exogenous factors as well as competition that affects their expected quarterly results, as an example. The Lean Model Canvas is aimed at startups and problem-solving but lacks long-term practicality because the model is difficult to scale for an entire company or its changing environment. This new model combines aspects of both models in one new model. In figure 3 a visual representation of the model is given. The items that are coloured green occur in both the Business Model Canvas and the Lean Model Canvas. The components that are coloured pink originally only appear in the Business Model Canvas, and the elements that are coloured yellow initially only appear in the Lean Model Canvas. This model aims to create a bridge where important parts of the Lean Model Canvas, like the problem and the solution for a value proposition, are captured, as well as unfair advantage since the first exogenous factors of the company's competition can be tackled here. This whilst also providing aspects of a Business Model Canvas like key partners and key resources/assets that are less susceptible to change quickly.

The Unified Model Canvas excludes some features from the Business Model Canvas and Lean Model Canvas. There are multiple explanations for eliminating some components. The first argument is related to scope. If every detail was included, the canvas would not be concise. Moreover, elements that are considered less relevant are left out, such as key activities. This is because the components problem and solution are more appropriate for this developing industry. Customer relationship is left out because there are two things to consider. First of all, some customer groups use these tools to support other means, such as creating games or social media content; in this case, AI tools are merely means to an end where the use of these tools stops when the right output music is generated for their content. Secondly, even though most brands opt for a subscription-tiered model, where

retention is indeed important, currently substitutes are scarce, and therefore a combination of a specific value proposition matched with the right customer segment is prioritised. The last element that is not included is channels. In this industry where developing and providing software is the main business model and users buy the tool in a SaaS fashion, it was considered of less importance to include this component compared to others. The reason for this is that these companies operate in a technology-centric market; success relies more on technical capabilities, superiority, and innovative features than on the channels in which they are delivered. Furthermore, these SaaS tools are often distributed directly to the customer through digital platforms.

4 Qualitative market analysis of tools for music co-creativity

Key information on tools						
Tool name	Release year	Developer	Method	Music sort	Maximum output lenth	Unique key features
Magenta Music Transformer	2018	Google	Text-to-music[1] and audio-to-music	Audio instrumental	a minute	Allows for input audio and is easy to implement due to google colab file
Jukebox	2020	OpenAI	text-to-music	Audio with vocals and audio instrumental	"multiple minutes long"[2]	Artists names and styles can be incorporated in the prompt
MusicLM	2023	Google	Text-to-music	Audio instrumental	5 minutes[3]	Suggested prompts are provided, generates 2 outputs simultaneously and is downloadable
Magenta Studio V2.0	2023	Google	Audio-to-music and text-to-music[4]	Audio instrumental	32 bars[5]	Works as plug-in for Ableton(DAW), Combine drum melodies in a mash-up
Google Lyria	2023	Google	Text-to-music and audio-to-music	Audio with vocals	30 seconds[6]	Partnerships with artists and YouTube
MusicFX	2024	Google	Text-to-music	Audio instrumental	70 seconds	DJ feature which allows mixing by combining genres, suggesting alternative descriptive words to enhance output
Udio Beta	2024	Udio(ex employees of Google DeepMind)	Text-to-music and audio-to-music[7]	Audio with vocals and audio instrumental	2 minutes	Unique settings such as clip start in context of full song and how much influence the lyrics have on the song and extensions of audio files of up to 15 minutes
Suno AI V3.5	2024	Anthropic	Text-to-music and audio-to-music[8]	Audio with vocals and audio instrumental	4 minutes	Usability, output of full music length, genre wheel
Apple Logic Pro V11.0	2024	Apple	Text-to-midi[9]	Audio instrumental	8 bars(per generation)	Customisation options, versatility and an active community providing tutorials

Table 1: Overview of various tools

In the industry where music is co-created with artificial intelligence, several companies and researchers have developed their own creations. This section will discuss several tools and their key capabilities. Google produces the majority of models, followed by OpenAI, and lastly independently developed research and tools. All models use a kind of input that is used in the algorithm to generate a song. Several examples are text-to-music, audio-to-music, text-to-MIDI, and audio-to-MIDI. While the most used method to create music is by a text-to-music prompt, several companies have different ideas about what kinds of inputs they prefer and how they are implemented in order to create a desired output. Customers evaluate tools on aspects such as usability and unique features in order to find the tool that matches their needs the best. In table 1 you can find an overview with the highlights of each tool. Below, several brands are discussed in chronological order, primarily showcasing the transition from research to commercial tools.

The first example is Magenta Music Transformer, which is research done by researchers at Google. This model was released in 2018, and the use of AI to create music was still in its beginning stages. Therefore, the possibilities for commercial users are quite limited compared to other models that were developed more recently. Some unique features that Magenta Transformer possesses are the ability for users to insert a WAV file that is altered to a MIDI file that can be incorporated in the network. Its easy accessibility for users and developers in Google Colab is another advantage that Magenta has over other tools in this section[CAH].

The next example is Jukebox, one of OpenAI's models. This research builds upon previous work by being able to generate raw audio directly and for a couple of minutes. OpenAI's Jukebox can generate music with or without lyrics, depending on the user preference. However, if the user does provide lyrics, the model does take longer to generate a song. As stated on their own website, "It takes approximately 9 hours to fully render one minute of audio through our models, and thus they cannot yet be used in interactive applications"[Ope]. Ten musicians that were provided with Jukebox said that they currently do not see application in the current music creation workflow due to these limitations. Besides the long rendering time, there remains a significant gap between human-created music and Jukebox's creations due to the lack of training with a longer output sample length as in a full song, which results in the model never generating melodies that were repeated. The context window in which the model remains very coherent is approximately twenty-four seconds. Therefore, compared to the other models, Jukebox performs worse due to recent and ongoing developments in this field[Sut20].

Another research project by Google is MusicLM. Compared to Magenta Transformer, this tool is more fine-tuned, and the capabilities are more developed. While both tools still only create instrumental music, MusicLM is able to generate up to five minutes of audio and has a more complete interface. If users need assistance because of a lack of creativity, MusicLM provides three suggestive prompts before the user starts writing their prompt. When the user is finished with its prompt MusicLM and its successor MusicFX both provide two outputs at the same time[Fra23].

[1] in text-to-music a few pre selected priming sequences or melodies are provided
[2] as stated in [Sut20]
[3] as stated in [Fra23], however user output does not reach such lengths but remains at maximum 30 seconds
[4] temperature and variation settings, no further input
[5] in the mode "Continue"
[6] in the mode "Dreamtrack"
[7] audio-to-music as option in paid subscription tiers
[8] audio-to-music as option in paid subscription tiers
[9] instrument and style selection, no further input

The following example is Magenta Studio, which mostly uses an audio-to-music prompt that reads a MIDI file and generates music in response. There are five distinct modes available to the user, each with a unique focal area. The first mode, "continue," is used to predict what a possible progression of an instrument or melody can look like. The input file can be elongated up to thirty-two measures using their Recurrent Neural Network[Mag]. Another feature is called "interpolate" and allows to combine two input files that are given in sixteen different files, perfect when the user wants to combine two drum strokes, for example. There are three other modes available in Magenta Studio; only "generate" will be discussed further in this section because it is the only mode that does not require an input file. Generate allows users to create four bar progressions based on only two characteristics, namely temperature and variations, which decide the randomness and the amount of variations used by the algorithm.

Another perception of an input prompt is Google's Lyria, which enables users to create music in a sound-to-music style called Music AI tools. This can especially be helpful for users that have a specific rhythm in mind. Music AI tools take the humming of a tune as input and generate instrumental sections on it [Goob]. Besides the sound-to-music prompt, a unique digital watermark called SynthID sets Lyria apart from the rest. The watermark is directly embedded in the waveform of the file. It scans the audio file looking for the watermark; all content created by Lyria can be detected by this, which is helpful for users to solve copyright issues or to find out the origins of the song [Gooa]. The text-to-music version of Lyria, Dreamtrack, is made in collaboration with Youtube and various artists like T-pain and John Legend. It allows users to create thirty-second-long music for Youtube shorts in the style of one of the pre-selected artists. This version of Lyria could definitely fulfil the needs of content creators without copyright problems[Goob].

On the contrary, MusicFX, which is also created by Google, will not generate music with certain queries asking to mimic specific artists or include vocals[Mus]. This is done to protect original artists voices and styles. MusicFX is a text-to-music model that creates two thirty-second songs by default, with the option to create one song of up to seventy seconds. A substantial upgrade is made to generate higher-quality music as well as faster song generation. Furthermore, when writing the prompt, it automatically helps by suggesting alternative descriptions that could better match with the desired output. What makes MusicFX stand out is its DJ feature, which allows users to mix beats by combining genres. The beats evolve by changing the composition of the selected genres and changing the importance of each genre with a slider. Since MusicFX and Lyria share the same developer, they both benefit from the digital watermark SynthID.

Udio Beta is an AI tool that is very similar to SunoAI. Udio is made by ex-employees of Google DeepMind. Both share similar qualities in the way that they offer text-to-music and audio-to-music capabilities. While they have similar characteristics, Udio does have some features that set it apart from the competition. Clip start is a feature that allows users to choose where their created clips start in the context of a full song. 10% corresponds to the beginning of the song, whereas 50% results in a song that starts midway. This is especially useful in combination with the extension feature. The duration for these extensions can lead up to fifteen minutes, and this context window allows the placement of the extension to create a lot of opportunities. Lyrics strength is the last unique feature and allows users to choose how much influence the lyrics have on the output song. A low value can result in a higher coherence but result in lyrics being ignored[Udi].

While the text-to-music format is the same, every company's approach is unique because of specific functionalities in the prompts or parameters that can be given in addition to the input. Suno AI allows users to pass on lyrics that have to be incorporated in the song. The amount of

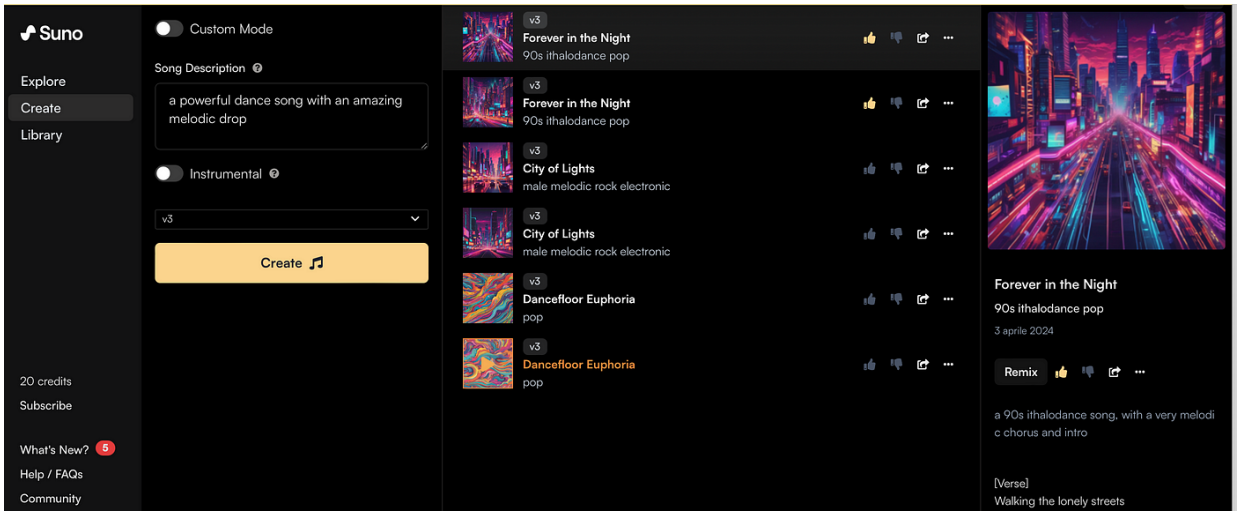

Figure 4: Create interface of Suno AI

parameters is quite limited, with only the style of music and the title being customisable to the user. Figure 4 shows the interface of Suno AI where users can create a song with the use of a description. Custom mode has to be activated if the user wants to pass lyrics or change the style of music. Other improvements have been made in the last months, for instance, in its functionality. Since the development of version 3.0 in February 2024, refinements have been made, which include better audio quality, more genres, and better coherence in the songs. The last update is V3.5 and builds upon V3.0 but enables users to generate songs of up to four minutes and allows for two-minute music extensions[Sun].

Apple Logic Pro is a DAW made by Apple. With the launch of the new version 11.0, AI has been implemented. The key features of implementing the AI spread further than the generating purposes only; however, they will not be discussed in this thesis. These session players generate audio that can be used to complement existing music in the DAW or on their own. There are three different instruments that the user can choose from. While drummer has been released as a session player in 2013, with the update, piano and bass have been added as options. Figure 5 shows the new session player of the piano with several settings that the user can use to tweak the generated output with. The algorithm generates eight bars of music per cycle. Users are able to select the kind of instrument and choose one of the multiple styles offered. While these inputs share the simplicity of other tools such as Suno AI or MusicFX, there are several things a user can do in order to change the result. Manually selecting the chords and how it progresses changes the instrument sounds as well as other settings such as complexity and instensity without having to generate a full new output. A higher complexity results in a more intricate beat, while intensity results in a higher velocity. The session player is not only useful to add instruments that complement the music that already exists in the DAW but also to explore different paths of music creation that otherwise would have been skipped in the workflow[Appa].

Figure 5: Interface of Apple Logic Pro DAW with the piano session player(from [Appb])

5 Applying Canvas to AI in Music Production

A breakdown of the Unified Model Canvas's components is provided in this section. New entrepreneurs can investigate these aspects in order to properly apply the canvas in this industry. This benefits the market orientation of the brand. A qualitative market analysis has been conducted to identify the state of the art in order to validate this framework. Each element will be discussed, and compelling examples will be provided.

Before every component is discussed, a short introduction is given as to how a company should implement this model. This includes an explanation why the model has to be filled in a specific order. Understanding how different client segments distinguish themselves due to their problems and characteristics is critical for any brand. Prior to developing the next segments, a new startup should focus on these elements. If this is done poorly, a brand may be tempted to concentrate on the solution too soon. This may lead to a brand that is more focused on finding a solution than completely understanding the issue and finding a comprehensive solution. Once the brand has identified its intended customer base and identified the issues, it needs to evaluate if the problem is worth solving. The brand should describe in the value proposition and solution sections what makes their product unique, why it is valuable, and how it addresses the issues. The last four segments give better insights into how the company aims to create and bring the product to the market. Key resources and key partners are needed in order to create the product. When combined, these provide the knowledge, resources, and tools needed to produce the product as effectively as feasible. The final sections define the inflow and outflow of cash. In revenue streams, the company should

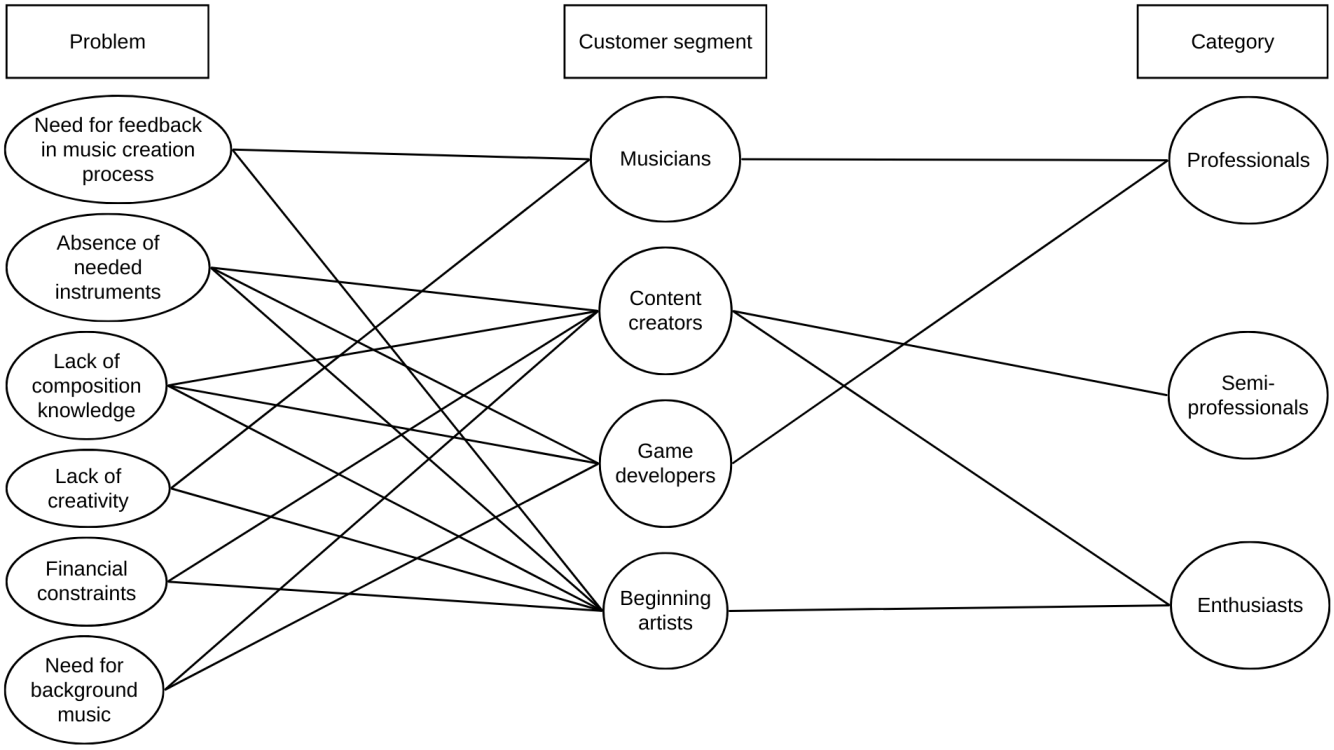

Figure 6: Problem Customer segment relationship diagram

document how the product can be monetised while the cost structure documents operational costs like the time, resources, and money that are needed to take the product to the market[Mau12].

5.1 Customer segments

It is crucial for a company to understand its customers. Customers are segmented into groups. A group of customers that share the same attributes, pain points, and behaviours belong in the same customer segment. This means that customers in the same segment share similar needs. As seen in figure 6 the problems relate directly to a customer segment. Even though needs overlap, particular characteristics within each segment ensure that they belong to distinct customer segments.

Customers of a company that allows for co-creating music between humans and AI could be segmented as follows: professionals, semi-professionals, and enthusiasts. The professionals comprise people who could use this tool in a professional setting, supporting a customer's main source of income. Semi-professionals are skilled customers in which this tool can support them in their hobbies; however, this hobby is not their main source of income. Lastly, the category of enthusiasts comprises hobbyists and beginners who use this tool for amusement rather than a way to earn money from. These categories can be split into smaller subcategories until all the users needs can be combined into one customer segment.

A first category in the professional segment could be musicians. What differentiates musicians from other segments is that musicians have the required knowledge and facilities to create music. Co-creativity with AI could lead to regained creativity and stimulation of the brain when an artist is stuck. With the possibility to interact with an online AI bandmate, an artist can share musical

ideas and generate complementing chord progression based on existing input music[SB24a]. This technology can not only increase creativity in the studio but also be used in live performances like the Dutch artist Reinier Zonneveld. With the help of software engineers, he developed software trained on over two thousand hours of his musical data, gathered from live performances over the past fifteen years. The first performance that is to showcase this software is in a back-to-back show in August 2024 [Wel]. A performance is called a back-to-back when more than one DJ works on the same show, taking turns playing songs[Lyn].

Content creators could be an example segment of the category semi-professionals. What is unique about this segment is that their needs are more widespread, in the sense that customers from this segment could be more easily satisfied in comparison to the segments in the professional industry. The people that fall into this category are people who do not necessarily want to create music as an end goal but rather want a generated song to support another need, for example, as background music under a live stream or YouTube video. This category therefore has no need for extensive hands-on composition options, and because of a lack of composition knowledge, a tool that generates a song quickly without a lot of hassle is an ideal solution. Dreamtrack, a subcategory of Google Lyria, allows users to generate thirty-second songs based on a text prompt. This tool has been made in collaboration with several artists, like T-Pain and John Legend. Users can generate songs in the style of these artists and use the output under YouTube shorts[Goob].

Beginning artists and students could be one of the segments in the category enthusiasts. What makes this category stand out from the other ones is that the people in this category do not commercially use this tool. These people want to create music but do not have the facilities or the knowledge to compose music. Even though the distribution in this group is large, a tool that allows for growth through lessons and tutorials could be beneficial for everyone.

5.2 Problem

Several kinds of problems arise with music creation. Physical problems that limit users in generating music could include equipment and knowledge. Musicians who compose music have a deep understanding of music creation, timbre, and harmony and have extensive knowledge of specific instruments, but not all. The production and recording of music have evolved rapidly due to digital improvements, with DAWs becoming a common tool.

These systems allow producers to add virtual instruments, MIDI effects, and tweak instruments separately. It can be costly and time-consuming to become skilled in these instruments and obtain a large library of these effects. Learning these skills and instruments can take years of training, which can be discouraging for aspiring musicians. Besides knowledge, limited resources and financial constraints create another problem for beginners. Hiring instruments and additional electronic devices, as well as a professional recording studio with sound engineers, can be expensive, which results in a certain standard of professionalism in order to create music.

What if the issue isn't a lack of knowledge or financial restraints but rather a poor singing voice? While it is possible to obtain equipment and knowledge, even with all the supporting tools, someone is either born with a singing voice or not. Famous artist Phil Collins used to play the drums and sing at the same time in the 80s; during that time he was part of the band Genesis[Wikb]. Had he not been born with this advantage, his professional trajectory would have been significantly different.

Beyond the incapacity to sing professionally, an existing musician could be dealing with a lack

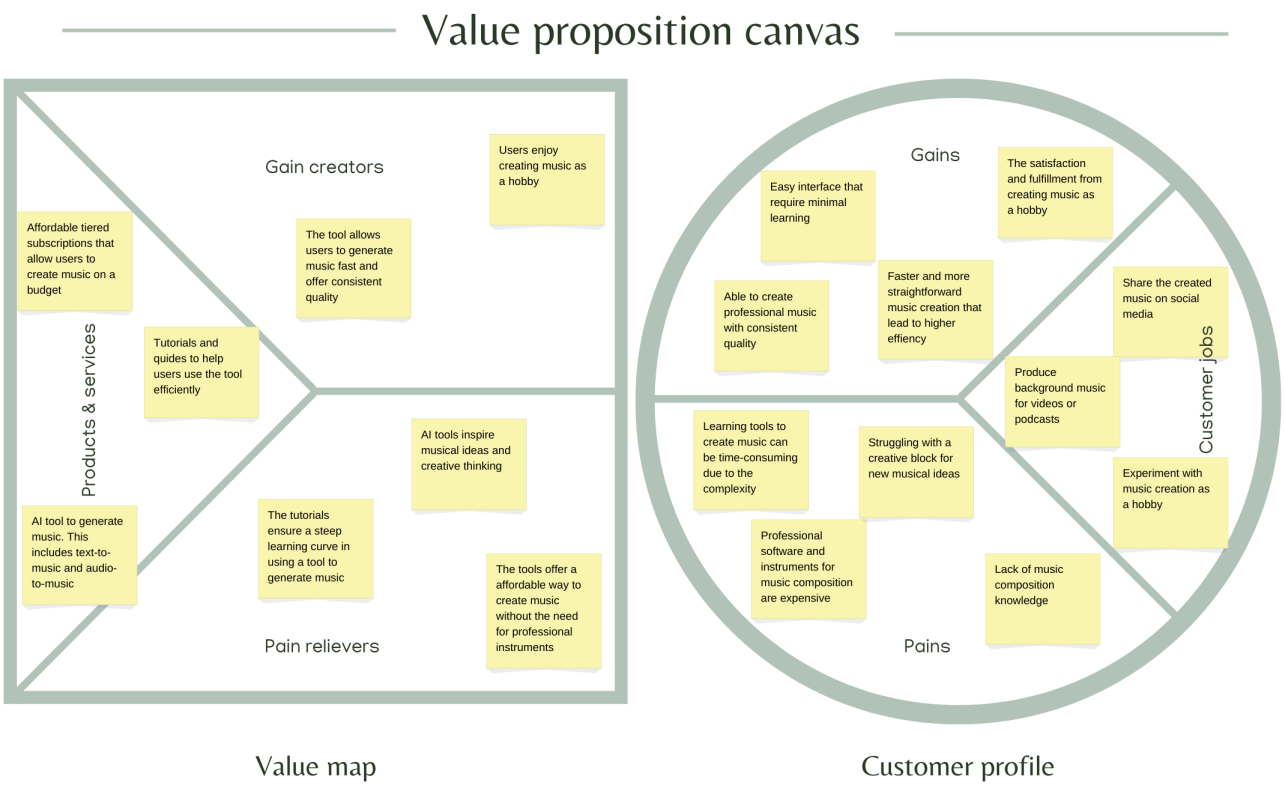

Figure 7: Example of a Value proposition canvas

of inspiration or a creative block. Musicians use creativity in combination with their knowledge to create something new and unique in their songs. When they get stuck on repetitive bar creation and struggle to create novelty in music, it limits an artist's growth.

5.3 Value proposition

A value proposition is the way a company captures its value for customers and, more specifically, "Why you are different and worth getting attention" [Mau12]. It combines the tools and services to satisfy the customer's needs. The value propositions should be catered to a specific customer segment. It provides value to customers problems by being innovative or able to outperform the competition. Several examples of different value propositions will be discussed in this section.

A value proposition canvas is a quick way to visualise how the product creates value. Figure 7 is an example for this industry. This canvas provides brand insights by visualising and identifying the components of the customer profile that contribute to the greatest improvement in value proposition and assesses which gain creators should be incorporated in order to complete the customer jobs the best. The canvas consists of two segments, the customer profile and the value map. Customer jobs document the tasks that the user has to fulfil, pains log unnecessary hassles that

the user experiences whilst completing the task, and gains describe benefits that the user expects from the value proposition. On the left side, the value map consists of the products and services needed to fulfil the value proposition, and pain relievers describe how and which customers pains are eliminated or reduced. The gain creators include which expected benefits from the customer gains have been implemented in order to create a better value proposition and could also include unexpected conveniences [dig].

In this industry, Music FX offers a unique value proposition with its reformulating capabilities. Users create a prompt, which will be used to generate a song. MusicFX automatically suggests alternative descriptions that could better match the desired output. Since the algorithm is trained on a big dataset, outputs can change drastically based on a slightly different prompt. The nuances in word description can be vital to generating the desired output song [Mus].

Whilst not available to the public as a tool, the software Reinier Zonneveld has created with a friend he met during his time at the university is unique in the way it was trained. Most models that are developed to make music are trained on immense datasets, often using copyrighted music from artists that have no knowledge of this and have not agreed to its use. Reinier Zonneveld has only used his own music and live performances as training data. This results in the model creating very accurate progressions when used in a live show. The model can benefit from a very unique and accurate dataset, resulting in better use for specific needs. In Reiniers case, that would be live performing a DJ set. During an interview with the Dutch radio NPO Radio1, he said that while he could recognise his own transitions of live performances in the beginning, when the dataset evolved and became larger, it became much harder for him to recognise. Something else he said is that whilst these tools can speed up the process, they still do not replace him and his creativity. The essence is that whilst these tools can be very good at creating something and exploring other options, they are currently always trained on human-made music, and musicians still have to decide how their music ends up sounding[Vee24].

Another example of a unique value proposition is that Google's AI department, DeepMind, developed a digital watermark that is directly embedded in the waveform of the output file. SynthID scans these watermarks and knows if a song is generated using their AI tool. Users are able to find out in an easy way if the music was made by AI. This can be beneficial if misinformation has happened or copyright infringement has occurred, like the song that seemed to be made by Drake but was actually made by an AI [Cos23][Gooa].

The final example could be Apple's Logic Pro with its innovative way of music creation. It allows users to generate music in a DAW that complements existing music. The user is able to put MIDI files as input in the DAW. With a single click, a new band member in a session player is created, and if a little tweaking is needed in some settings, the users are presented with an intuitive display that allows for easy customisation. Whilst being unique and innovative, the easy way in which users can create and customise the outcome further adds as a unique value proposition[Appd].

5.4 Solution

This section describes the solution the canvas offers to customers pain. Solutions are linked directly to the problem but can offer additional benefits besides taking away the main problem of a customer segment. The solution offered by a company that allows users to generate music with an AI tool stretches further than music creation alone. Besides music generation, it sparks stimulus for beginning enthusiasts to play an instrument, for example. The bar that is set to create (professional)

music is lowered by using a tool that generates a song from scratch with only a prompt. Furthermore, an online AI tool offers a wide variety of instruments and styles to choose from. Not only beginning musicians benefit from this, but experienced musicians can adapt their style to a completely different style of music. This offers a solution to the problem given in the problem section by taking away the requirement of extensive instrumental knowledge.

An AI tool that generates audio offers the solution to a lack of creativity in a musician. In a study by (H. Chu et al.) the researchers found that the tools that allow users to generate music with AI were inspired creatively, with one of the participants saying the following about a created song: "a desire to add a piece of lyrics to it" [CKK+22]. EZDrummer3 is a tool that can generate a rhythm for an instrument based on an input audio file. The output can be used directly to complement the existing instruments in the song or used as inspiration. Interaction with such an AI bandmate can help create a breakthrough in the music production process [SB24a]. This can be an example of a solution since it solves the financial problems given as an example in the problem section. The instruments that otherwise have to be rented are substituted by this tool.

What if a person wants to create a song but their original voice is unsuitable for singing professionally? In the problem section, Phil Collins was discussed because of his unique career, which would not have been the same if it were not for his ability to not only drum but sing as well. On the 8th of May 2024, SoundID released Voice AI, an AI tool that lets users capture their voice and generate a song using their lyrics but sung in one of twenty-three pre-set studio-grade voice models. Not only that, but the humming sound of the user can be transformed into twenty-one instruments. It is currently available as a versatile DAW plugin[Son].

5.5 Unfair advantage

In this section, unfair advantages are being covered. What keeps the solution a company provides to its customers from being replicable for competition and therefore, a sustaining advantage? This industry has a few elements that can be challenging to replicate.

First of all, technical capabilities could be a moat against competition. This results in an unfair advantage if your company's technology is just too advanced and comprehensive in comparison to rivals. Users could be convinced immediately if the tools output is superior in this field compared to others. On the other hand, rivals may be able to duplicate the level of quality if it is not patent protected.

Besides technical superiority, reputation and brand image can form advantages that cannot easily be copied. It takes years for a company to obtain a certain reputation, and after that, it is not easily changed. A demonstration example is Spotify. While a lot of music streaming platforms exist, Spotify has been the clear market leader compared to others like Deezer or Apple Music, while Deezer allows users to listen to music in FLAC quality. A link in people's minds is made with Spotify as a good streaming platform.

Another example are partner deals. These deals are exclusive deals between the company and the agent. In this industry, artists could be the most crucial agents in these partnerships. A partnership could offer an algorithm that is trained on their songs. If the model is trained on the curated data of the artist, copyright issues are less likely to happen compared to competitors that train on the artist's musical data illegally. Besides training on this data, users could use the artist's name in the prompt or select them as the style in which they want music to be generated. Google Lyria's Dream Track allows users to generate music in the style of certain artists [Goob]. Besides style and

copyright benefits, collaborations with artists can elevate the reputation of the brand if people see that these artists use the tools of the brand in their music creations. This cannot be copied by competitors and therefore makes for an unfair advantage.

5.6 Key resources

In the constantly evolving AI-generated music industry, several key resources are essential for maintaining a competitive advantage and facilitating innovation. Examples of these resources are human resources, financial resources, physical assets, and intangible assets.

The crux in this market is human resources, where a group of employees with a special expertise in creating AI/ML software or having an inherent comprehension of musical aspects like timbre, logical chord progressions, and other subtleties is key. Some present-day AI models, like Musenet, find it difficult to reflect this combination of knowledge accurately in the tool[WZL+23].

This expertise is linked to the research and development (R&D) department, which are considered intangible assets. In this industry, where the business model is to create software, a strong R&D infrastructure is crucial. Ongoing investments in this department in order to improve model coherence, enhance interaction, and elevate music quality are important to remain relevant and not fall behind competition. In order to test the advanced AI technology created in this division, hardware like CPUs and GPUs are needed. This hardware is used to run the demanding computing processes required to create music and to implement the complex algorithm in the cloud. While the hardware is part of the R&D department, it is considered to be part of the physical assets.

Access to large and varied datasets is essential for successful music production. Intangible assets such as these datasets, which come in a variety of audio file formats (WAV, MIDI, FLAC, and MP3) with a lot of metadata providing details on bands, musical genres, beats per minute (bpm), and tonality, need to be carefully selected and labelled. With a less precise and extensive dataset, optimising the algorithm becomes harder.

Another example of intangible assets are patents and copyrights, which play a significant role in this industry. These intellectual properties are used to support strategic goals like horizontal integration in addition to protecting inventions. As an example, consider Google's SynthID digital watermarking technology, which can be tailored for a range of uses in diverse industries, expanding the scope and versatility of the business's artificial intelligence algorithms[Gooa].

5.7 Key partners

Each company has a group of strong partnerships that are essential for the brand to thrive. These partners ensure high-quality music output, an efficient workflow, and expand the market reach. These partners can be categorised in a number of key areas, like musicians, content creators, and brands such as technology companies.

The first category of partners are musicians. These professionals offer insightful feedback and creative input, which help refine AI algorithms. With their expertise in rhythm, timbre, and musical nuances, they are helpful in the process because this knowledge is not often possessed by software engineers. These collaborations also make it easier to train AI models across a wide variety of musical genres and styles without copyright infringement issues. The wide variety of data enhances the model's versatility and output quality.

The next example is partnering up with top technology firms like NVIDIA and Intel, which offer specialised gear, which is necessary. These businesses provide the high-performance CPUs and GPUs needed for the demanding computational tasks connected with creating music and implementing complex AI systems. Besides the deals with these hardware companies, additional partners to effectively run the cloud services are needed. Examples of these could be Google Cloud and Microsoft Azure. Google and Microsoft have their own tools in Google's DeepMind Lyria, and Microsoft has a partnership with Suno AI for its development. This makes for better forward vertical integration in the supply chain.

The last group of partners that are discussed are content creators and influencers, which can be useful partners. These influencers are useful to boost brand awareness on various platforms. These influencers are active on social media platforms like YouTube, Instagram, and TikTok and could reach a young audience. By co-creating content using the tool, engaging in promotions, or featuring AI-generated music in their videos, they boost brand awareness.

5.8 Revenue streams

The revenue streams section includes the ways a company monetises its business model. Companies that allow customers to create music with AI can generate revenue in various ways. Some examples could be subscription tiers for the software, followed by licensing, royalties, one-time purchases, and a marketplace.

A tiered subscription model offers various levels of access and features based on the user's chosen plan. A user that chooses for a higher subscription tier is able to generate music quicker in a higher quality (e.g., FLAC versus MP3). Other benefits could be that the user is able to create more songs and select song duration compared to a free or lower subscription tier. This model encourages users to choose a higher tier to unlock these features. Besides the fact that a higher subscription tier offers a higher cash flow, encouraging long-term commitment by providing options for monthly and annual billing cycles provides a consistent cash flow of income.

Licencing and royalty agreements for music produced by the AI can be used to provide an extra cash stream. This method is very useful for commercial use cases, such as music for video games, movies, and public spaces like cafes. Based on the scale of the venue or length of use, the licensing rate can vary. With digital watermarking technologies like SynthID, usage can be detected[Gooa] and if disproportionate use is observed, the company can act on it. Licensing not only creates a revenue stream but also increases the spread of AI-created music.

Another example is a one-time purchase. This could be to unlock extra features, longer generations, or the software as a whole. An advantage for the consumer is that if the person buys the rights to use the software, additional costs to improve the experience are avoided. Apple offers the Logic Pro for a one-time purchase of two hundred dollars [Appa].

The final example is to open an online store where people can purchase and sell songs made by AI. Users are able to sell their songs with the company charging a small fee for facilitating the transaction. A library that matches with what the users need might improve this marketplace even more. AI-powered solutions, like Spotify's AI DJ, could create playlists in response to users's preferences[Spo]. For content creators, game developers, and other creative workers, such a library can be valuable since it offers ready-to-use music.

5.9 Cost structure

The cost structure comprises all the costs that are incurred by a firm. Examples could be operational costs needed to run the business and overhead costs, which are not directly linked to a product. The nature of the costs can differ, with some being fixed costs and others being variable costs. It is important for a firm to decide what kind of business model they have and what kind of cost structure matches with it. A company that is cost-driven is likely to give in on product quality in order to keep the costs low, whereas a company that focuses on delivering high-quality products is less likely to cut expenses in the trade-off between cost and quality.

The first category are operational costs. These costs are necessary to run the firm. In order to create an algorithm to generate music, a lot of development costs are made since the main product of a company in this industry is the service of music generation. This includes the costs of research and development to create and fine-tune algorithms to their specific needs, as well as continuous development and upgrades. The collection and acquisition of data and formatting it in a way suitable for training the model is also included in this category.

The second example are physical assets such as hardware. The CPUs and GPUs required to run all the software need to be of high quality. These components tend to be very expensive, and even though such a company does possess firm infrastructure where this hardware is run, companies in this kind of industry do not sell physical products and therefore do not require a big factory. The hardware is therefore one of the most crucial aspects of the physical cost structure.

Maintenance is the next big cost expense. In order to maintain a relevant firm and not fall behind, systems need to be checked regularly. These costs vary from small bugs in the software to large malfunctions in the hardware. The costs in hardware and software involve improvements in safety and reliability.

Overhead expenses are costs that are not directly linked to a product. Examples of this category could be salary expenses for the administrative staff and human resource management, as well as equipment used by cleaners.

6 Validating the unified framework with two examples

To validate the Unified Model Canvas, it is applied to different AI tools in this industry. Intending to showcase clear differences between the business models to become evident. We identified two different kinds of tools. The first category consists of companies offering customers a service by creating ready-to-use music directly from a prompt. The second category are brands that use AI to generate audio; however, it offers users a wide variety of customisation options during the process of generating the music or after. With a hypothetical example inspired by a blend of real-world examples in the industry

6.1 A lazy mans tool

The first type of company models that will be discussed are models that create ready-to-use songs. Users are able to create music without the hassle of hours of composing music in a professional studio. The companies offer a generative tool that allows for music creation in a matter of minutes. With a single text prompt, these tools generate songs of up to four minutes in the case of the latest versions of Udio Beta or Suno AI[Sun]. Figure 8 shows an example of a filled-in Unified Model

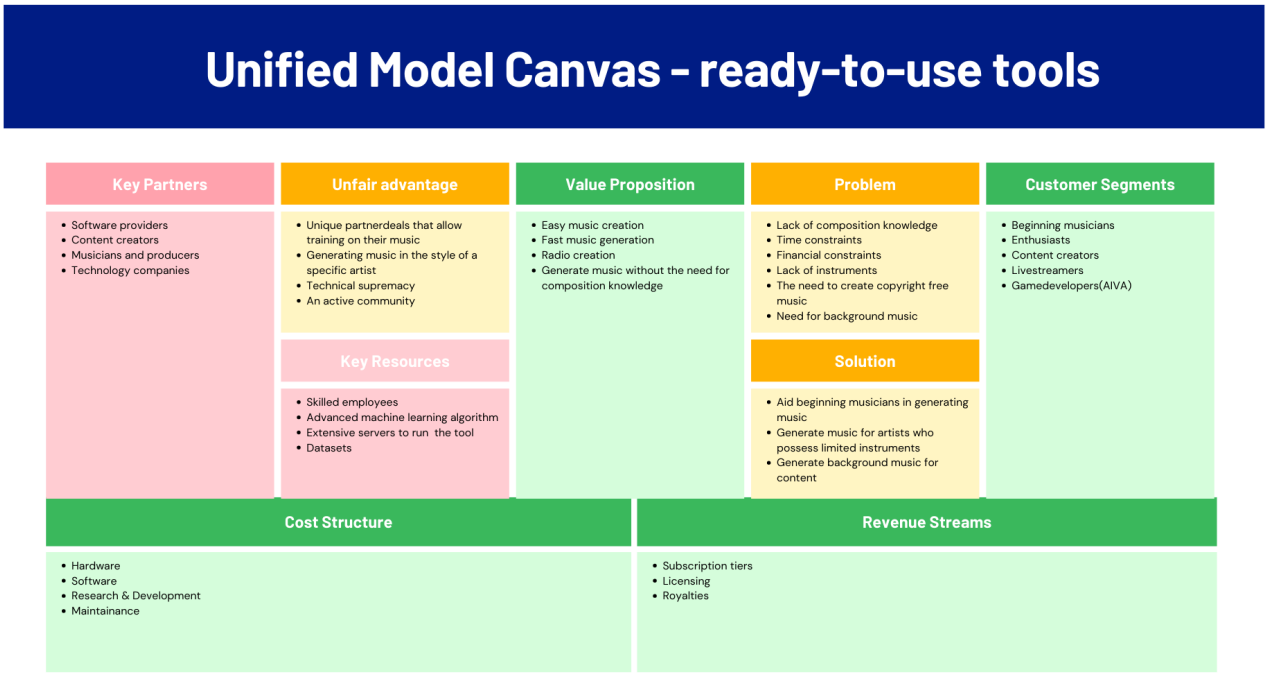

Figure 8: Example of filled-in Unified Model Canvas for a Lazy Mans Tool

Canvas for a generic brand that offers a tool in this segment. Several instances of elements in each component are listed, such as which customer segments are targeted because of their specific customer pains that are relieved by the product the brand offers. Later in this section, real examples and supportive visualisations for some elements are discussed, which further demonstrates why this canvas is useful for a brand to incorporate.

What will become evident in the following examples of the components is that the Lazy Mans Tool provides a value proposition and solutions for a certain set of problems. Figure 6 showcases example problems that overlap with issues that customers that use this tool experience. The client segments that utilise this type of technology overlap with the elements "beginning artists" and "content creators" in the figure. A conclusion that a brand in this category can draw is that these customers are "enthusiasts" or "semi-professionals" and do not fall in the category "professionals" because of the simple needs that these segments require and the purpose of using such an AI tool. Using this kind of diagram provides useful insights into the problems and the kind of customer segment the company aims to target.

One of the leading companies that use AI's generative capabilities to generate songs is Suno AI. In December 2023, Microsoft announced that it had partnered up with Suno AI to bring the technology to its Copilot[Mic]. On the 21st of May 2024, Mikey Shulman, CEO of Suno AI, published a press release on the website's blog stating that 125 million dollars have been raised to further develop and enhance the development speed of the tool[Shu]. Microsoft is an example of a key partner of Suno, while the cash also represents a key resource for the company's development.

Besides a key partner and a key resource, some unique features are discussed that are part of the value proposition of Suno AI. The first example is the "explore" wheel, in which users can find

inspiration for music creation with unique genres such as “Mandarin Trance” and “Cajun Afrikaner Folk.” Users can select a genre and will be presented with five different example songs that can be played. Users are presented with undiscovered styles and find value in this whilst also encouraging unconventional thinking. In June 2024, two of the most important features have been added, which boost the versatility of the model with song radio and audio input. Song radio allows users to select a song, and by selecting song radio, songs in similar style will automatically be added to the queue. Users are able to enjoy these songs continuously. Just like in Google Lyria Music AI tools, Suno now allows users to base song generation on audio input, with the requirements being that the audio file must at least be six seconds minimum and sixty seconds at its longest, and that the user is part of one of the paid subscription tiers[Sun]. These features are not only part of the value proposition, but combined with the active community, song radio can become an unfair advantage due to the inability to copy this unique combination.

Google revealed Google DeepMinds Lyria in November 2023. What sets this tool apart from other models is the collaboration between Google and YouTube. “An experiment in YouTube Shorts designed to help deepen connections between artists, creators, and fans through music creation”[Goob]. Google developed this tool with the help of artists such as T-Pain and John Legend. Training on the data of such artists can lead to users seeing a better resemblance between the text input and the output song. Users are able to generate songs with the distinct style of these pre-selected artists. The output is thirty seconds long, and users can use it under Youtube Shorts. While this value proposition is unique in the way that users can choose the style of the artist that has collaborated on the project, other companies that offer similar products are often trained on copyrighted data. This, in combination with the inherent properties of these generative models, means that with an elaborate and precise prompt and a bit of luck, the same results can be achieved. However, a bill requiring businesses to provide the training datasets was introduced in the US Congress in April 2024. The datasets should be delivered to the Register of Copyrights 30 days before their generative AI model is made public, whether it is a music, video, or photography tool. Big companies like OpenAI keep the training data private, and whilst the goal is not to fully ban training on copyrighted material, it does aim to create a responsibility to report the massive amounts of training data [RE]. Other brands like Suno or Udio do not have partners like this and have to rewrite prompts if a certain artist occurs in a prompt. If the bill is to go through, this partnership deal also becomes an unfair advantage over other companies. Whether training happens with copyrighted music legally or illegally, in both cases they offer an example of a key resource of such a company.

Revenue streams for a company that provides ready-to-use music like Suno AI can vary. Subscription tiers are one of the cash flows that come in. These tiers are divided into a free, pro, and premier plan. Whilst the free tier does not allow for direct income, the rights for the music and distribution of it lie with the company. This allows for monetising popular songs. Pro and premier tiers cost 8 and 24 dollars a month if opted for annual billing, respectively. If users opt for a monthly billing, the price increases to 10 and 30 dollars[Sun].

Figure 8 sketches a depiction of a brand that could fall into this category. Overall, the Unified Model Canvas successfully depicts these kinds of tools with elements in this diagram that are characteristic for this category. Customers such as beginning musicians and content creators fit well for this category since the value proposition is catered for this segment. The pain points users endure, such as a lack of instruments or composition knowledge, resemble problems customers often experience. While each company has their own unique take on an AI tool, every brand allows

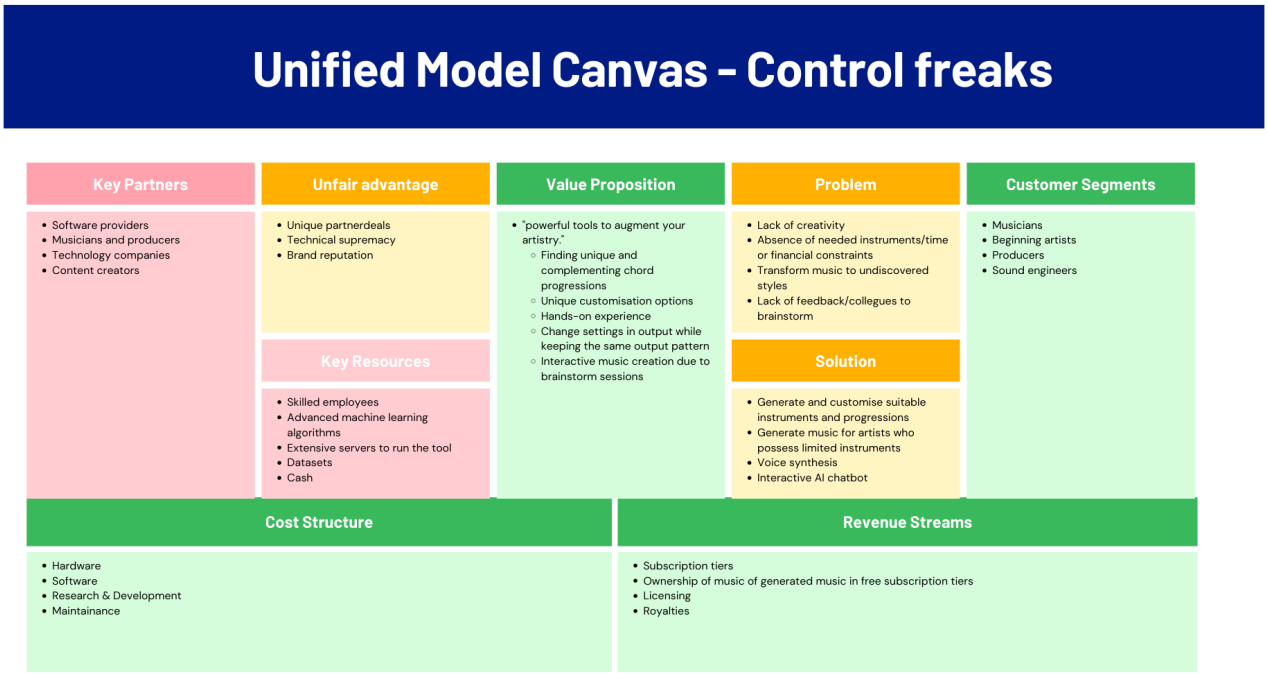

Figure 9: Example of filled-in Unified Model Canvas for a Control Freaks tool; source text for value proposition from [Appa]

users an easy experience in music creation. Key partners that are crucial for this kind of music creation are artists such as John Legend. Without these artists, people would not be able to use their voices, and in this category that allows for music generation with vocals, that would be a big loss. The tools from the qualitative market analysis that monetise aspects of their tool all opt for a subscription tier where users pay monthly or annually. The Unified Model Canvas allowed the user to structure a business model in a fast and meaningful way. It offers a quick overview, and comparisons of different tools can easily be made. Examples of similarities and dissimilarities between companies become evident, making it easy to pinpoint business models that have the same market positioning. New startups can use this tool to structure their business model, and competitors can be identified, further proving its use.

6.2 Control freaks

The other side of the spectrum is the category "Control Freaks." These tools tend to create audio for user groups with very specific use cases. The tool offers a wide variety of customisations in the process of music creation. Unlike the Lazy Mans Tools, which offer limited customisation by regenerating the entire output or changing the style of music, a Control Freaks tool like Apple Logic Pro 11[Appa] or Wavtool[Wav] allows users to interact dynamically by changing aspects without having to regenerate the output as a whole. These tools are examples of DAWs with the implementation of generative artificial intelligence. An alternative name that is used for this tool is Generative Audio Workstations (GAWs)[SB24b].

Figure 9 shows an example of a filled-in Unified Model Canvas for a generic product in this segment. The diagram shows elements that are unique for this segment, and further in this section, several elements are explained with real instances. A single glance at this canvas versus the one for the Lazy Mans Tool shows clear differences, such as the value proposition with the overarching proposition that these tools augment the artistry of the user.

The way a DAW uses the generative purposes of AI differs per platform. Wavtool uses GPT capabilities for an AI chatbot. The DAW allows users to create MIDI bars with the interactive chatbot. Users write a prompt about what the DAW should create. If the prompt is created, the conductor chat might ask more in-depth questions about the desired output. While this DAW shares similarities with, for example, Apples Logic Pro V.11 in the way that it allows users a lot more freedom in the process, the chatbot offers other advantages compared to Apples DAW. The chatbot could function as an interactive colleague where artists can brainstorm how new instruments might complement the current music existing in the DAW. The unique value proposition that a DAW with chatbot capabilities offers is the opportunity to grow as an artist quickly. The combination of the accessibility of a DAW with lots of customisation options without losing the easy composition process of a tool that generates music with a text prompt, such as the Lazy Mans Tools[Wav].

On 7th May 2024, Apple launches Apple Logic Pro V11.0. This new version of the DAW includes several implementations of AI in the process of music creation. The first addition is new instruments that can be selected as a session player. A session player functions as a bandmate that complements the music that exists in the DAW. Three different session players currently exist: the drummer, the bass player, and the keyboard player[Appc].

The session player of the drummer has been around for the past decade with improved sound and new functions in the latest modification[Appe]. The session player generates a MIDI track for 8 bars, which is called a region. There are four different styles to choose from: rock, songwriter, alternative, and R&B. A generated output can be edited to the user's desired output. The display is intuitive and easy to use, and new performances are generated by clicking a single button. A lot of different customisations are available. First, users are able to use a slider to change a number of general settings that are available for all instruments. The intensity, complexity, fill amount, and fill complexity are some examples. The session player adapts to these changes by playing at a higher velocity if the user wants the instrument to be more present in the song, while changes in complexity could result in a more intricate beat while maintaining the general pattern that was originally created. The interface allows users to see a visual presentation of the pattern that is being played and which nodes are played additionally if the user opts for a more complex beat. These simple changes offer a solution to the problem of a musician by changing the song drastically, but if an artist really wants to renew their creativity, the last two settings are far more valuable.

For example, a typical phenomenon in jazz is "trading fours," where drums or other instruments like a trumpet exchange solos, in this case for four bars [Bri]. In this style of music where improvisation plays a big role in music creation, the ability to change and generate unique fills by changing how often a solo occurs or the complexity of the created solo is a unique value proposition Apple offers.

An empirical study states that "providing options that enable users to select the listening circumstances or features of the music could improve human satisfaction with AI-generated music" [CKK+22]. These tools that offer a lot of customisation provide a solution to the limited contentment that the Lazy Mans Tool like Jukebox currently offer [Ope].

Although these instruments sound great, and even though the sliders and other unique cus-

tomisation options offer endless possibilities for each instrument, they lack something that cannot be programmed in a computer, and that is a feeling for rhythm. The best artists in the world generally have an amazing feel for rhythm, but even humans can make mistakes. The "humanize" option that is added in this new version tries to mimic that. As stated on the site of the Logic Pro User Guide for Mac[Appd] "Humanize adds a random value to the position, velocity, and length of selected note events. It can add life to strictly quantized material, or manually entered notes. This preset is useful on rhythmic parts, such as clavinet, piano, drum, and percussion. Alter the values for Position, Velocity, or Length to increase or decrease the random factor for each of these event parameters"[Appd]. This offers a unique value proposition in comparison to other tools in both categories, especially when the following case study is taken into account. A case study indicates that people tend to perceive AI-generated music negatively compared to humanly composed music. The "humanize" option creates randomness in the audio, and this could benefit people creating music with Apple's DAW[Maw01].

There are several settings that are unique to the instrument. For example, for the drums, the user is able to select if the snare and/or kick is being played. If the user wants the snare to be incorporated in the song, it offers different styles like a centre shot or a rim shot where the drummer hits the head and rim of the snare or a side stick where the drummer hits the rim of the snare with the back of the stick whilst holding the tip of the stick on the head. In jazz, this function would be very useful because the side stick is often used whilst playing the ride cymbal at the same time[Appd].

The value propositions and solutions that these systems provide offer a lot of freedom for the user, but they are replicable; however, Apple, for example, has some unfair advantages that are not easy to copy or buy. The first advantage Apple has is its brand name. A lot of people use devices made by Apple, and they mostly have a positive experience using them. The products are revolutionary and overall well developed and thought through. Secondly, Apple is renowned for its eco-system, where devices are connected and work together to create a better brand experience. With Logic Pro, there are applications for both MacOS and iOS, which allows for better versatility. Users are able to start a project on their computer and continue working on it on a tablet if they are travelling.

The DAWs in this Control Freaks segment target certain problems. Figure 6 shows that the problems a tool in this segment solves correspond with the problems of musicians and professional game developers. Professional game developers could benefit from the chatbot capabilities of Wavtool, where they can fine-tune and discuss elements with the bot to get the desired outcome. Further changes in settings can be made along the way. Musicians that struggle with creativity could experiment with the session players Apple offers. The relationship between problems and customer segments becomes visible, and such a diagram can support a Unified Model Canvas in order to create a product that is tailor-made for their target segment.

Overall, figure 9 is a great example of a brand that could fall into this category. The value propositions that are focused on the journey of making music rather than just the end result are characteristic for tools in this category. The hands-on experience attracts a smaller customer base filled with professionals such as sound engineers and artists. The types of problems that these customers experience are specific and complex. In contrast, the problems of a Lazy Mans Tool user are often practical, such as financial constraints or the absence of composition knowledge. These needs lead to innovative solutions such as voice synthesis or an interactive chat bot where a user can discuss how the song could progress. Key resources such as extensive datasets and cash are

crucial to developing and training the generative algorithms used to generate audio. The Unified Model Canvas presents the user with a clear overview of a business model. This section provides real-life examples of elements and how organisations differ and how they are similar become clear, making it simple to identify business models with comparable market positioning. New entrants can also benefit from this canvas to position themselves correctly, avoiding mistakes down the line.

7 Discussion

This section reflects on the results of this research as well as limitations and further research possibilities on this topic. This study on the intersection of AI and music creation aims to answer the following research question: What different business models can be identified in the market for AI for music creation? With qualitative market research, the state of the art is identified. In order to identify the different business models in such a rapidly evolving market, a framework has been developed that is validated on real-life AI tools for music production.

7.1 Reflection on results and implications

Section 6 offers compelling examples for two different business models based on characteristics of tools that share similarities. Ready-to-use tools and the business model for a Control Freaks tool are distinct in important features, such as the problems that they solve for their target customer segments. Their differences are easily recognisable in the example figures of generic tools made for each segment. The figure 9 was made for the Control Freaks segment, and diagram 8 was created for the Lazy Mans segment.

The customer segment of a Control Freaks tool is more focused on professionals in the music industry and problems such as lack of creativity or that they lack a colleague in the process of music creation, whereas the ready-to-use tools aim more for the general public, with a wider customer segment from beginning artists that lack composition knowledge to game developers that need background music for their game. The overarching value proposition of a Control Freaks tool makes a case for their difference with "powerful tools to augment your artistry,"[Appa] whereas the value proposition and the problems of the ready-to-use tools are more centred around the end product and how to obtain that quickly rather than the process of music creation.

Even though these business models differ from each other in most components, there are core elements that overlap between the two examples. In order to create these tools, sections such as key partners and key resources are needed. Because the brands behind the business models operate in the same niche industry of generating music using AI, overlapping aspects do occur. Examples of these could be tech companies for key partners and advanced machine algorithms for the key resources compartment. The Unified Model Canvas is supported by visualisations such as figure 6 that shows how a brand attracts a certain customer base directly relates to the problems the value proposition tackles. Figure 7, the value proposition canvas, explains in a concise way how a product creates value for the customer and, after evaluation, what customer pains and gains are incorporated in the final product.

The implications of this research are significant for new entrants in this industry. It is crucial to know how business models differ from each other in order to understand different market segments. Understanding these differences can benefit brands in their market orientation and enhance the

chances of success. Besides offering new entrants a guide to orient and structure their business, this framework provides an opportunity for existing brands to refine their business models and strategies in order to improve tailor-made products for their target customers.

7.2 Limitations and future work

This section discusses several factors that have limited this research.

First of all, while the application of the Unified Model Canvas has been successful in identifying different business models, it was challenging to fill in the entire framework accurately and at a detailed level for a single brand due to limited data. More transparency of the brands could lead to a better qualitative market analysis, resulting in better examples for the proof of concept in section 6.

Moreover, the rapidly changing nature of AI technology, especially in this industry, means that our results may quickly become outdated, though we have aimed to make our framework as robust as possible, allowing for future research to be conducted on the Unified Model Canvas.

Future work could include research into how these different business models evolve over time, resulting in interesting insights as to how this market and the technology evolve as well as market saturation. Collecting more detailed data from a larger number of brands could further validate and refine the Unified Model Canvas; potentially, in collaboration with a brand to get access to internal data could improve case studies. What would happen if the number of brands in the example dataset varies from 10 to 1000, and how does this influence the number of business models that can be categorised? An interesting subquestion regarding the framework could be to what extent business models and solutions will differ for different types of music, such as classical plays, house music, and audio that includes vocals, such as jazz or folklore music using dialect, and how these differences influence the way the algorithms work.

8 Conclusion

With the current interest and applications of AI in creative industries, this thesis explores the impact of AI in the music industry and identifies the various business models of AI-based tools for music generation and production. The research question is as follows: What different business models can be identified in the market for AI in music creation?

To address the research question, literature research has been done, and a qualitative market study has been carried out in order to find out the state of the art. Alongside these studies, research by design is done, taking components of both the Business Model Canvas and the Lean Model Canvas in order to identify and compare business models by analysing value proposition, problems, key partners, and cost structure, among other elements.

From validating the framework, we identified two distinct business models: “A Lazy Mans Tool” and “Control Freaks,” representing two ends of the spectrum in terms of customisation and user involvement. A Lazy Mans Tool targets a broad audience, including beginning artists and content creators who seek quick and easy music generation. In contrast, tools in the Control Freaks segment cater to musicians and professional engineers where the process and manipulation of music are essential rather than just the output product.

The example diagrams illustrated the core differences and overlaps between business models in this market. While both models depend on similar aspects, such as tech companies providing hardware and resources like sophisticated AI algorithms, their value propositions diverge significantly, resulting in a different customer segment.

Future work should aim at validating the framework in a more detailed environment by applying it to a single brand, possibly by collaborating with specific AI music companies to gather internal data. Another take could be to carry out a longitudinal study to get a grasp at how these business models change over time because of the rapidly evolving nature of AI.

References

[Appa] . https://www.apple.com/logic-pro/. Accessed: 24-06-2024.

[Appb] Get started with Session Players in Logic Pro for Mac. https://support.apple.com/en-lamr/guide/logicpro/lgcpa4324884/mac. Accessed: 26-06-2024.

[Appc] Logic Pro takes music-making to the next level with new AI features. https://www.apple.com/ca/newsroom/2024/05/logic-pro-takes-music-making-to-the-next-level-with-new-ai-features/. Accessed: 01-08-2024.

[Appd] Logic Pro User Guide for Mac. https://help.apple.com/pdf/logicpromac/en_US/logic-pro-mac-user-guide.pdf. Accessed: 11-06-2024.

[Appe] Overview of MIDI Transform presets in Logic Pro for Mac. https://www.apple.com/eg/newsroom/2013/07/16Apple-Unveils-Logic-Pro-X/. Accessed: 28-05-2024.

[art] 200+ artists urge tech platforms: stop devaluing music. https://artistrightsnow.medium.com/200-artists-urge-tech-platforms-stop-devaluing-music-559fb109bbac. Accessed: 22-07-2024.

[Bar12] S. Barber. Soundstream: The introduction of commercial digital recording in the United States. *Journal on the Art of Record Production*, 2012.

[Bri] Trading Fours. https://percussion.byu.edu/trading-fours#:~:text=A%20common%20kind%20of%20improvisation,the%20rest%20of%20the%20band. Accessed: 28-05-2024.

[CAH] M. Dinculescu C. A. Huang, I. Simon. Music Transformer: Generating Music with Long-Term Structure. https://magenta.tensorflow.org/music-transformer. Accessed: 22-07-2024.

[Cas] J. Castañón. The Business Model Canvas. https://www.academia.edu/26101884/The_Business_Model_Canvas_The_makers_of_Business_Model_Generation_and_Strategyzer. Accessed: 10-06-2024.

[CKK+22] H. Chu, J. Kim, S. Kim, H. Lim, H. Lee, S. Jin, J. Lee, T. Kim, , and S. Ko. An empirical study on how people perceive AI-generated music. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pages 304–314, 2022.

[Cos23] J. Coscarelli. An A.I. hit of fake 'Drake' and 'The Weeknd' rattles the music world. *The New York Times*, 2023.

[CRH+23] A. Cai, S. R. Rick, J. L. Heyman, Y. Zhang, A. Filipowicz, M. Hong, M.and Klenk, and T. Malone. Designaid: Using generative AI and semantic diversity for design inspiration. *CI '23: Proceedings of The ACM Collective Intelligence Conference*, 2023.

[Dai24] S. Dai. Towards artificial musicians: Modeling style for music composition, performance, and synthesis via machine learning, 2024.

[dig] How to Use Value Proposition Canvas: The Definitive Guide. https://www.digitalnatives.hu/blog/value-proposition-canvas/. Accessed: 01-08-2024.

[Fra23] A. Agostinelli; T. I. Denk; Z. Borsos; J. Engel; M. Verzetti; A. Caillon; Q. Huang; A. Jansen; A. Roberts; M. Tagliasacchi; M. Sharifi; N. Zeghidour; C. Frank. MusicLM: Generating music from text, 2023.

[Gib13] J. Gibson. Introduction to the MIDI standard. *Introduction to MIDI and Computer Music*, 2013.

[Gooa] SynthID. https://deepmind.google/technologies/synthid/. Accessed: 11-06-2024.

[Goob] Transforming the future of music creation. https://deepmind.google/discover/blog/transforming-the-future-of-music-creation/. Accessed: 16-05-2024.

[Hij] M. Hijink. Voor 2 cent per 'liedje' perst AI de popmuziek uit. https://www.nrc.nl/nieuws/2024/05/17/voor-2-cent-per-liedje-perst-ai-de-popmuziek-uit-a4199269?t=1721735077. Accessed: 23-07-2024.

[HKNR+20] C.-Z. A. Huang, H. V. Koops, E. Newton-Rex, M. Dinculescu, , and C. J. Cai. AI song contest: Human-AI co-creation in songwriting. *International Society for Music Information Retrieval*, pages 1–2, 2020.

[Lyn] Will Lynch. How to DJ back-to-back. https://blog.pioneerdj.com/djtips/how-to-dj-back-to-back/#:~:text=Playing%20back%2Dto%2Dback%E2%80%94,one%20DJ%20behind%20the%20decks. Accessed: 29-05-2024.

[Mag] Magenta Studio. https://magenta.tensorflow.org/studio/standalone/. Accessed: 14-06-2024.

[Mau12] A. Maurya. *Running Lean: Iterate from Plan A to a Plan That Works*. Lean (O'Reilly). O'Reilly Media, Incorporated, 2012.

[Maw01] F. Tigre Moura; C. Maw. Artificial intelligence became Beethoven: how do listeners and music professionals perceive artificially composed music? *Journal of Consumer Marketing*, 38(2), 2021-03-01.

[Mic] Turn your ideas into songs with Suno on Microsoft Copilot. https://www.microsoft.com/en-us/microsoft-copilot/blog/2023/12/19/turn-your-ideas-into-songs-with-suno-on-microsoft-copilot/. Accessed: 3-06-2024.

[Mus] . https://aitestkitchen.withgoogle.com/nl/posts/faq. Accessed: 15-06-2024.

[OP10] A. Osterwalder and Y. Pigneur. *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*. The Strategyzer Series. Wiley, 2010.

[Ope] Jukebox. https://openai.com/index/jukebox/. Accessed: 16-06-2024.

[RE] N. Robins-Early. New bill would force AI companies to reveal use of copyrighted art. https://www.theguardian.com/technology/2024/apr/09/artificial-intelligence-bill-copyright-art. Accessed: 4-06-2024.

[SB24a] E. Sandzer-Bell. 6 AI bandmate apps that interact and inspire you creatively. 2024.

[SB24b] E. Sandzer-Bell. Generative audio workstations: AI VSTs & the future of DAWs. 2024.

[Shu] Mikey Shulman. Suno has raised $125 million to build a future where anyone can make music. https://suno.com/blog/fundraising-announcement-may-2024?utm_source=www.therundown.ai&utm_medium=newsletter&utm_campaign=microsoft-s-new-ai-recalls-everything. Accessed: 3-06-2024.

[Son] SoundID VoiceAI The first voice changer plugin for your DAW. https://www.sonarworks.com/soundid-produce/voiceai. Accessed: 16-05-2024.

[Spo] Spotify Debuts a New AI DJ, Right in Your Pocket. https://pr-newsroom-wp.appspot.com/2023-02-22/spotify-debuts-a-new-ai-dj-right-in-your-pocket/. Accessed: 29-05-2024.

[STC21] E. Suh, M.and Youngblom, M. Terry, and C. J. Cai. AI as social glue: Uncovering the roles of deep generative AI during social music composition. 2021.

[Sun] . https://suno.com/. Accessed: 14-06-2024.

[Sut20] P. Dhariwal; H. Jun ; C. Payne; J. W. Kim; A. Radford; I. Sutskever. Jukebox: A generative model for music, 2020.

[Udi] . https://www.udio.com/home. Accessed: 19-06-2024.

[Vee24] "R. Zonneveld ; W. Veenhoven". Interview with reinier zonneveld. https://www.nporadio1.nl/fragmenten/kunststof/9bc5c481-8048-4408-9e9b-99fd8befcbed/2024-04-10-reinier-zonneveld-techno-dj, 2024. Accessed: 11-06-2024.

[Wav] . https://wavtool.com/. Accessed: 24-07-2024.

[Wel] A. Clare Welsh. Reinier Zonneveld to play b2b with AI version of himself at R² festival. https://djmag.com/news/reinier-zonneveld-play-b2b-ai-version-himself-r2-festival. Accessed: 15-05-2024.

[Wika] Digital audio workstation. https://en.wikipedia.org/wiki/Digital_audio_workstation. Accessed: 29-07-2024.

[Wikb] Phil Collins. https://en.wikipedia.org/wiki/Phil_Collins#:~:text=From%201970%20to%201975%2C%20Collins,on%20Genesis%20albums%20and%20concerts. Accessed: 16-05-2024.

[WZL+23] Lei Wang, Ziyi Zhao, Hanwei Liu, Junwei Pang, Yi Qin, and Qidi Wu. A review of intelligent music generation systems, 2023.

[YZWY17] Lantao Yu, Weinan Zhang, Jun Wang, and Yong Yu. Seqgan: Sequence Generative Adversarial Nets with policy gradient, 2017.