AI4media

BARCELONA SUPERCOMPUTING CENTER

Exploring AI music composition tools for humans

White paper - October 2022

Artur Garcia Saez

Executive Summary

This document explores the interconnection between tools and technologies produced recently in the Machine Learning (ML) research community intended for music creation, and a community of composers and creators and their respective audience.

This connection aims to produce novel and interesting music, using advanced Artificial Intelligence (AI) tools to further explore creative possibilities of human creators. This is another step in the line of technological advances applied to musical creation, from the invention of novel instruments such as the violin or the synthesizer, to the creation of music recording equipment.

This use case, developed within the AI4Media project, is integrating tools and techniques created in the context of the AI4Media project to investigate the possibility of developing a set of tools that humans can employ to easily use complex ML tools. This approach has to overcome an enormous distance between communities working on the same topic but with opposed views: creating novel musical content. One approach explores the artistic potential of these techniques, while the other aims to develop novel techniques that improve generative tools.

The development of novel AI tools brings together the work of two completely different communities formed by researchers in the field of AI and musical composers, respectively. This relation is further explored with the definition of requirements that users identify at the moment of facing a new technology with deep implications to their work, but also to the music industry as a whole. While the music industry has been subject to deep changes in the connection with the audience and the impact that distribution of online music had on the musical format, having novel composition tools will change the relation between creators and all the other stakeholders of the musical production industry.

This work explores the advances happening in the field of AI in relation with advances in composition tools and techniques used by human creators. Among these one may include non-experts, using the recent advances in technologies for musical creation to perform tasks otherwise out of reach, but also proficient creators willing to explore the new paths opened by these techniques. While artists may benefit enormously from novel tools that may help the exploration of novel artistic directions, their use may pose dangers related to authorship and royalties derived from the artist's creativity.

Key messages

This white paper identifies the following fundamental aspects of this relationship between artists and technologists:

- Providing advanced tools to creators in music is a technically complex problem as interesting tools are developed in a restricted community of ML developers. Also, communication of the potential of these tools acts as a barrier for non-experts.
- Tools for music co-creation go beyond learning models and should include the architectural requirements that a user needs to execute a full application. This means access to powerful computing infrastructure.
- A creative process cannot be formalized, and a key element is the balance between powerful tools with the freedom to use and combine them. This is the basic requirement for the co-creative process.

Table of contents

1

Introduction

Music is a human language. As any other human creation, it includes a complex description that expands along a number of different scales and dimensions, and their relation enriches the resulting creations.

Classifying music, in any form, is a difficult task. Additionally, the analysis of music from different perspectives sometimes results in contradicting views, while successful composition can break most of the premises that one may have originally formulated. While one can have a formal or technical approach to music, ultimately it is an art and as such may evade any standardized approach to its formulation as a set of simple rules.

Machine Learning applications usually have well defined tasks and objectives. Classification or generation of data is measured with precision via mathematical expressions, and the overall goal can be precisely assessed. If music evades this approach, and we cannot formalize general properties of the obtained products, it is extremely difficult to certify the validity of a given approach.

Human co-creation is viewed in this context as the relation of human creators guided by artistic interests, with advanced tools appearing recently in different fields . This relation is dynamic, as novel and improved techniques of interest for artists are expected to appear in the next few years from the emerging set of novel techniques continuously created. The final result of this interaction is not intended to be a self-generating content machine, but to define the possibilities of interaction between humans and tools, continuing a path already started by advanced digital programs already used in the field. As such, the research program also explores the integration possibilities of new techniques emerging from research labs, to easily reach other communities.

This is not different from previous moments in music history, from the creation of novel

Figure 1: Two recent works using AI tools for human creation, targeting unfinished fragments by Beethoven (left) and Schubert (right). These works used an AI system to assist a team of musical experts to complete the composition, and performed in front of an audience in a classical concert hall.

instruments, including the irruption of electronic devices, to the appearance of recording techniques that translated music to any conceivable moment and space. However, the emergence of automatic tools for music creation raises a completely novel concern, that of human authorship.

The interaction of creators with technology should not be regarded as happening in a single direction, as their feedback is of great value to guide unexplored paths on the research side. We explore in this work the opinion received from authors, to understand the view that the community has of this new technology, and the potential impact to their work.

As examples, we motivate most of our discussion in recent advances and examples of the interaction between AI, authors and performers. Notably, authors may be considered playing an active role by iteratively interacting with an AI system. However, it soon becomes clear that the same technology can be used for the recreation of an author style, in order to revisit unfinished works. Notable examples are the completion of unfinished works by Beethoven and Schubert (see Figure 1 above) after centuries of conjectures about their artistic directions. While these works have been completed using assistance from AI tools, their final authors are human artists.

2

Problems & Challenges

The use of AI tools is disrupting the music industry, its impact appearing in different areas such as recommendation systems, audience analysis, and content generation.

In an era where content is ubiquitous, having access to automatic creation tools brings powerful support to creators and industry actors. Music is not only an object of consumption by itself, but also almost any media format such as movies or video games or even news videos will always appropriately be accompanied by music.

Machine Learning tools have an impact on different roles in the music creation process and its distribution. While music can be labeled and selected using automatic tools, a relevant aspect of new generative techniques is the way they can be introduced in the creative process of authors. This means **integration with current tools,** rather than the need to include novel tools in the pipeline of composers.

Having advanced tools developed by technical teams integrated under existing programs used by composers is one of the main challenges of the path separating the community of creators from the research environment where technological advances take place. Identifying the target set of tools to complete the integration is a difficult task given the limitations raised by different vendors and the wide variety of tools employed by creators.

The quality of the creative process assisted by AI tools is deeply related to the quality of the tools employed. But in order to broaden the access to advanced tools, one has to provide access to advanced computing architectures. Current **generative models** perform complex operations using advanced computing infrastructure which may not be easily accessible to non-expert users, and out of reach pricewise. Besides, this equipment may become outdated at a higher cost than may be replaced by occasional users.

The formation of interdisciplinary teams, where authors are tightly working alongside technology developers, may become a key strategy to enhance the capabilities of generative models used in the industry. Other fields have addressed similar difficulties by introducing experts in the development field. An example is the collaboration with linguists* in the development of generative models for human speech generation and text to speech applications, such as automatic translation.

Generative models used in audio generation can be adapted to the general problem of generative music. However, most of these models are developed in the industrial context of i.e. automatic translation, with a large market and commercial interest. On the contrary, the range of application of generative models in music is not well defined, from the perspective of the development of novel creative tools in a very experimental environment.

*See for example https://taia.io/

3

Industrial needs

The community of creators identifies the easy access to complex and advanced tools as a fundamental requirement for the widespread use of AI tools for music composition.

QCreators and composers that show interest in novel forms of creation to complete their creative process need to converge into a common set of tools defining a common language.

Integration of components developed in the context of AI4Media is an example of the development of advanced techniques happening in AI. This development is likely to continue in the next few years, increasing the demand for integration efforts.

The results of fundamental research have acted as an attraction mechanism for artists interested in the published results. Reproducibility of these landmark experiments may pose a difficulty to industrial development. While in the context of artistic creation the issue of reproducibility may not be as relevant, in industrial applications the assessment of quality and control of the outcome may become central. As with other AI tools, the uncertainty of the outcome is related to technological and environmental factors sometimes out of the control of the users.*

A clear definition of the creation rights and a **modification of copyright laws may be necessary**, in order to address the challenges that creation and ownership of compositions are facing in the wake of automatic composition. A major challenge to address in this direction is the relationship between the industry as a whole and the artists as creators. Using generative models raises the question of property and ownership in a field where authors delegate their royalty rewards to companies and labels**.

Content creators approach novel tools in an exploratory manner. The application of generative models in the creative process is subject to circumstances that the industry cannot foresee, and tools need to be adapted to this variability to become widespread. As an example, the performing quality of these tools will determine the degree of involvement of humans in the process, and it is expected to improve over the years. It is then necessary to interact closely with a diverse community of creators that may perform an initial exploration of the applicability and possibilities of the interaction with the new tools. From here, along very general lines, creators and industry may come to an agreement on a formalized setup, based on standards and protocols. This is a key element to overcome limitations that different vendors may pose over the development of tools across different platforms and systems.

*See i.e. https://petapixel.com/2020/08/17/gigapixel-ai-accidentally-added-ryan-goslings-face-to-this-photo/

**See i.e. McCormack, J.; Gifford, T.; and Hutchings, P. 2019. Autonomy, authenticity, authorship and intention in computer generated art. In Proceedings of the international conference on Computational Intelligence in Music, Sound, Art and Design. Craig, Carys J., The AI-Copyright Challenge: Tech-

4

Survey results and summary

In order to better understand creators' perceptions and requirements for these novel tools, and the impact they may have on the creative process, we have performed a survey using personal interviews and an online survey to a selected group of tens of content creators with or without prior technical experience, located mainly in the Barcelona (Spain) area.

Our methodology aims to provide a diversity of opinions from different creators. The interviews have been performed with authors with a background either in AI and/or music, so a clear bias towards a favorable opinion in this interconnection is observed. However, the responses received on the online survey are a good example of the diversity of opinions and potential impact these AI tools may have over the wide range of profiles among musical creators.

Our exploration using the online questionnaire aimed to identify the relationship between AI tools and creators, and also its potential impact on the industry. We track 3 faces of this relationship, namely (i) the potential impact of AI technology to the current work of artists and creators, (ii) the market disruption by novel tools, and (iii) the quality metrics applied to the musical outcome of this relationship.

The explored population appears to be formed by a relevant number of creators with a technological background, and interest or even experience in AI tools. Among these, there seems to be a mixed use of AI tech, with a potential impact on future work.

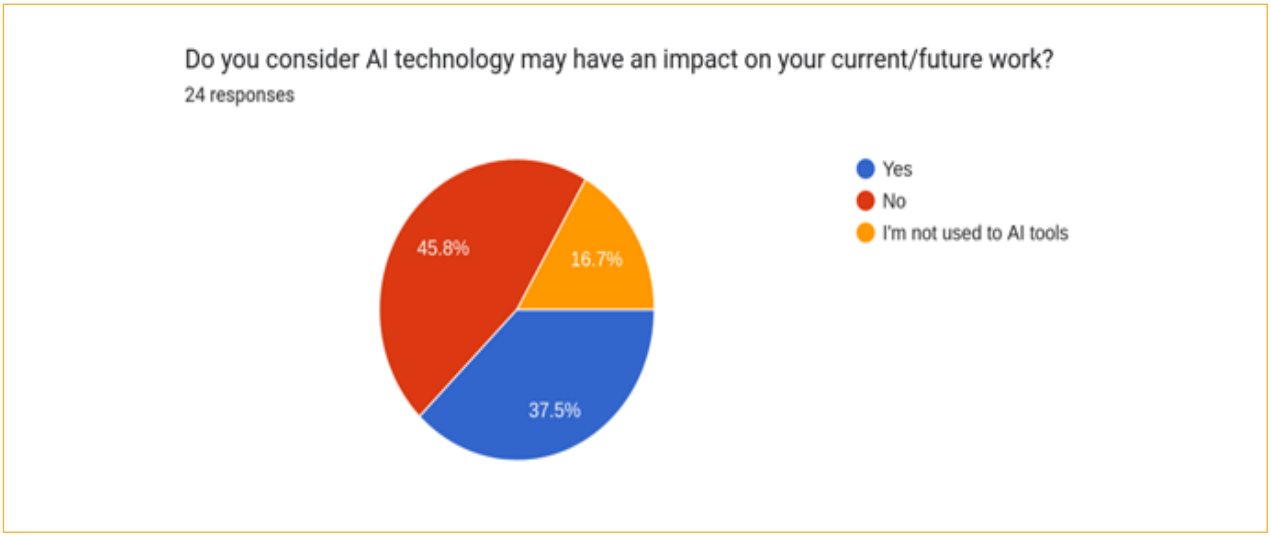

Figure 2: Opinions on the impact of AI technology on current/future work of music creators

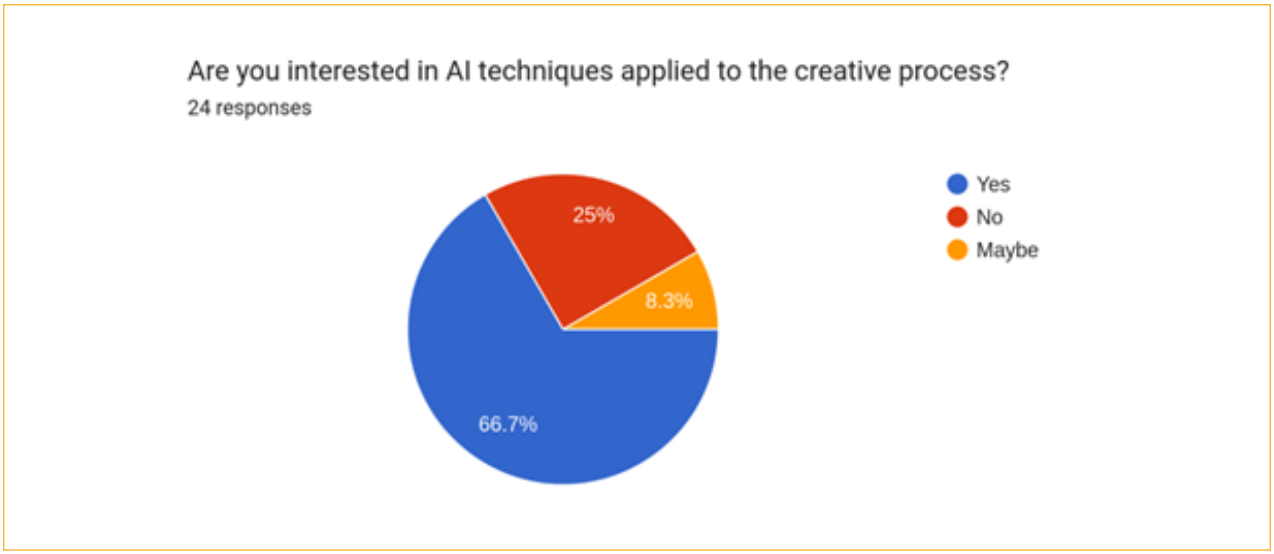

Figure 3: Level of interest in AI techniques applied to the creative process

When asked about the potential impact of AI in their creative process/work, only around 37% indicate some impact (see Figure 2). However, 66% of users show interest in AI tools (see Figure 3). As only 16% indicate not to be familiar with the use of AI tools, we suggest that the little potential impact of AI tools expected by users interested in the technology may appear as a result of both the current status of AI methods not being mature enough for users, and also of the limited access that artists may have to this technology.

In order to better establish and define the relationship between humans and algorithms, we explore the perception that authors may have now on the final impact that AI may have in the final form of their work. These may include original compositions where the generated audio is used as the initial inspiration for a novel work, or a generated audio that fits a visual content.

In particular, it is important to identify the potential tasks where an industrial solution may cover a niche looking for relevant solutions. In this direction, responses (see Figure 4 below) point towards tasks where creativity may not be the most relevant component, such as Harmonization and Sonorization, and related basic compositional tasks. On the other hand, a full composition obtained from AI tools is not expected as a solution for creators.

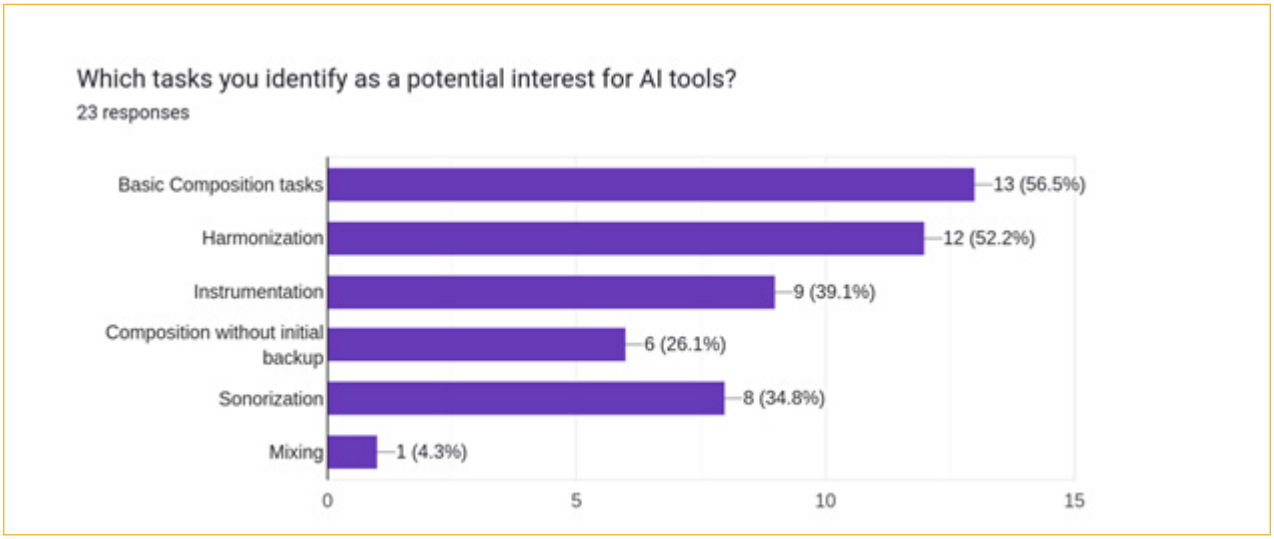

Figure 4: Potential tasks for applying AI tools

Given the current state of the art of generative tools for audio, it is expected that authors and creators will still play an important role in the final creative steps. We concluded above that little impact of current AI tools in the work of creators is partially due to the immaturity state of tools intended for authors. However, authors share the perception that this may change in the next few years, a trend that can be observed from the trend observed in Figure 5 below.

Here, 25% of respondents expect high quality outcomes from AI tools already, which is in contrast with Figure 6, where a majority of answers point to little impact of AI tools in the final outcome. However, there is a relevant perception that high quality creations will appear in the short term (2-10 years) due to novel AI tools. This perception highlights the need for active development of novel tools and resources for artists in the next few years.

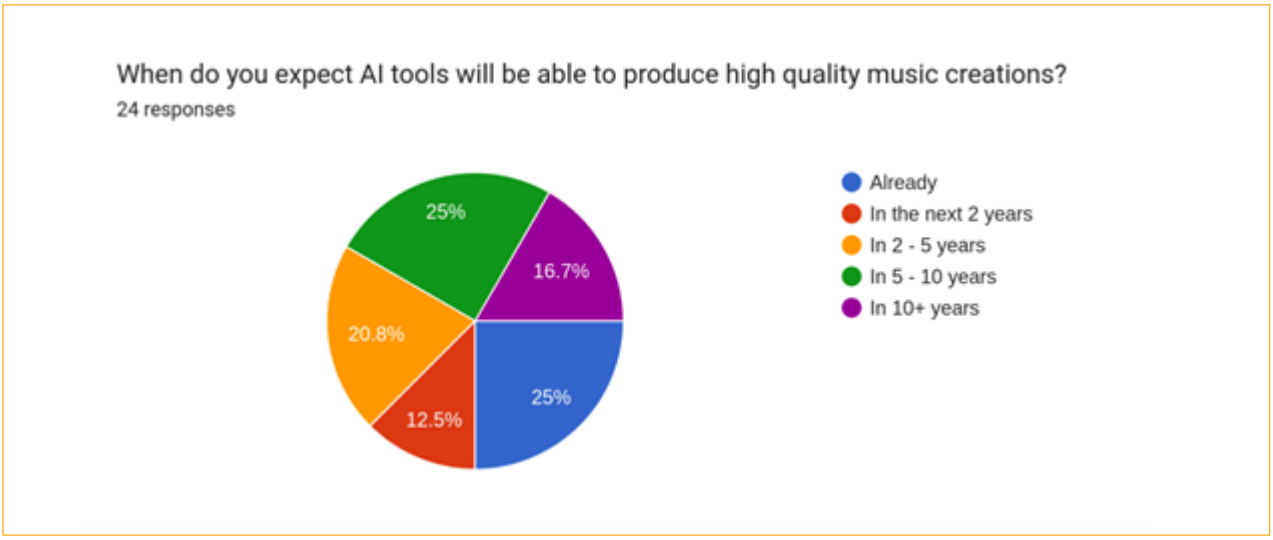

Figure 5: Expected impact of AI tools on the quality of novel music

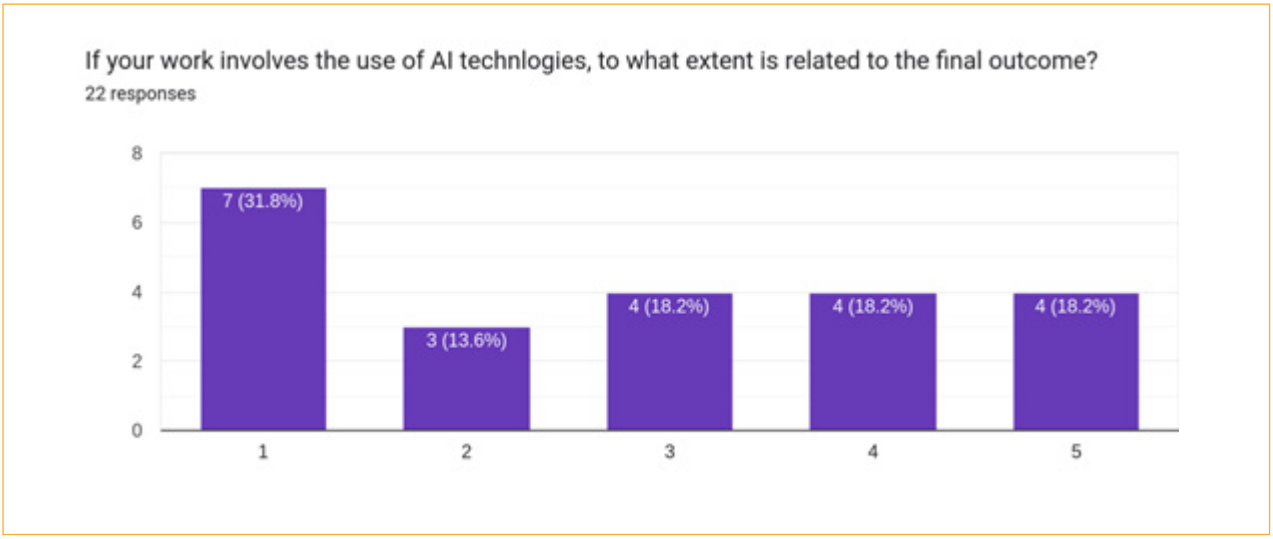

Figure 6: Opinions on the relation of AI technologies to the final outcome of the work of music creators (1 is minimal relation, 5 is a close relation).

Which fields may benefit in the short term from novel AI tools? Observed responses (see Figure 7 below) identify an equal distribution among different media products that include music as a component in a major or minor role. Surprisingly, there is a balance among Movies/TV/Video Games, where music can be regarded as secondary to the visual content, compared to original music. This balance may arise because fewer authors will be inclined to use AI tools to stand alone musical creations due to generative tools' capabilities, but also due to a large population of creators interested in more diverse media formats. For the industry, this represents a diverse range of users and solutions that need to fulfill a wide spectrum of applications (I.e., generative solutions that integrate musical generation with video editing may be required for many users).

Finally, we want to explore the perception that this community has over the future of these generative tools and their impact on novel music creation. The quality of this novel content, as suggested in the introduction, is difficult to assess, so the educated responses of creators may give an idea of the perceived impact these compositions already have in the short and medium term.

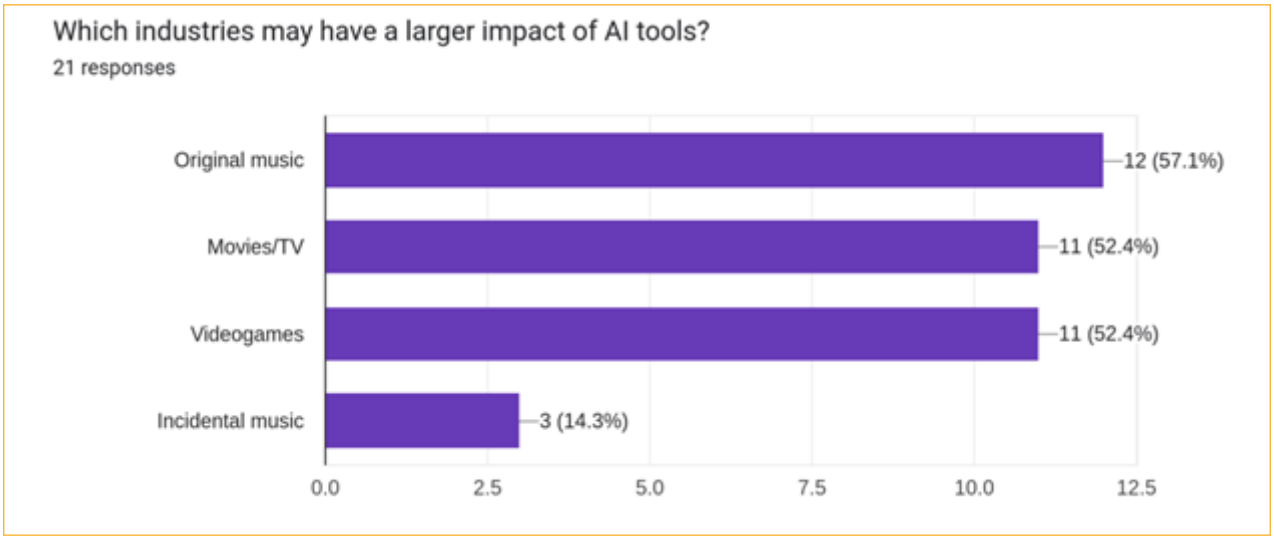

Figure 7: Opinions on the impact of AI tools in different industries

5

Conclusions

When presented with novel tools developed in the field of AI, creators are offered with new roads to explore in their creative process.

These tools are sometimes difficult to identify, as they are developed in a rather technical environment with little contact to an artistic environment. Following the big disruption the music industry is facing in the last decades the potential impact that these novel technologies may have in a continuously changing environment is not clear, especially to music creators.

The music industry requires an enormous quantity of musical content to fulfill the demands of an audience consuming media with music content in many different formats* . The relation of the industrial actors with human creators using AI tools developed by the very same industry, is a key element for the future of music as addressed to a general audience, but also to specialized audiences eager to discover artistic explorations guided by human composers using novel tools. With these novel tools, enhanced creativity may bring new music formats that advance a very long music tradition.

The current status of generative tools requires the intervention of humans to produce quality compositions. The capabilities of these models require enormous computational power, not always available to final users. The outcome needs the supervision and intervention of humans, and be subjected to final human arrangement. In order to overcome the difficulties that accessing this technology may mean for many authors with little technical capabilities, efforts directed at defining a protocol used by model generators, integrators, and final tool developers will increase the interaction and use on upcoming tools.

A key point in the industry is to define the relationship among human creators, the tools under active development in research and industrial centers, and the final audience. This process requires flexibility in the human and AI partnership, defining new tools that are central to the interaction, but also requires an environment of test and development of the methodologies that will allow use of current state of the art by creators. Letting creators lead this process may become central to obtaining a set of powerful and useful tools that will be subject to intensive development in real creative processes.

*See i.e. See i.e. https://www.musicbusinessworldwide.com/files/2020/01/2019-U.S.-Music-Consumption-Year-End-Report-1.pdf

AI4media

Follow us @ai4mediaproject

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951911

info@ai4media.eu www.ai4media.eu