AI4media

AI for Video Game Testing and Music Processing

White paper - October 2022

Executive Summary

In the last decade, the game industry has seen a lot of changes to accommodate the continuous need for new content. Games changed from being a product that is created by small teams to allow for a couple of hours of experience to a continuous service that requires updates and additional content all the time.

The current development methodologies only consider hiring more humans to cope with this increase. This is not sustainable as the game production cost increases exponentially. The industry needs to start adopting AI methods to help with the increased demand for new content which will lead to a decrease in production costs.

For this to happen, we need to understand the challenges and needs that are facing the industry. We ran two surveys with our industry partners to explore the idea of using AI-powered tools for quality assurance (QA) testing and music analysis/synthesis. The surveys showed that these tools should not replace humans but should rather enhance their work experience by helping them to discover bugs faster or inspiring them to create music. These AI-powered tools should be able to integrate easily with the current production pipeline. In QA testing, industry partners need to have fine-grain control on the testing conditions and allow for multiple different output methods ranging from summary to replay videos for the discovered bugs. On the other hand in music analysis/synthesis, the industry partners didn't care much about composing music for production but they were interested in creating inspiring samples based on user input tags. Finally, we also investigated industry views on Trustworthy AI features. The survey results showed that the game industry cares mostly about privacy and robustness of AI tools rather than bias, transparency, or explainability.

Key
messages

- AI-powered tools shouldn't replace QA and music analysis/synthesis processes done by humans but rather enhance existing practices and help humans in achieving their tasks.

- Industry partners don't mind spending more time to get AI-powered tools working but they must be able to easily integrate them into their production pipeline.

- It is important to have fine control over the input of the automated AI systems and provide a variety of methods to showcase their output.

Table of contents

Introduction

The game industry has changed in the last decade. Game companies started focusing on online games, digital platforms, and mobile and social games. This caused a shift in the development paradigm.

To retain players in the game, the industry needs to keep producing new content and adding new features continuously. To keep up with the market and stay on top, they have to be ahead of the curve and make sure that the released content is optimized and fully tested. When millions of players are playing these games daily, companies have to make sure that any change or new feature to the game won't affect existing content negatively or risk losing players. However, as games grow bigger with continuous updates, companies have to hire bigger teams of testers to make sure that any changes are not disruptive before releasing them to the users. With the increase in the demand for new content, the pressure on game companies is bigger than before, however, investing in increasingly larger quality testing teams is unsustainable over the long term.

Artificial Intelligence (AI) can provide a radically new approach to testing and optimizing games by letting developers test thousands of different game configurations, possibly in a fraction of the time required by human testers. While AI methods might not replace all human labor in testing, they can augment and enhance existing practices by providing product evaluations faster and more efficiently than current methods both in terms of human work hours and data. Beyond optimizing game testing, AI methods can be used to generate novel content either automatically or in co-creation with a human designer. As procedurally generated content should fit seamlessly into the game experience, these methods can also be evaluated by the same techniques used for assessing human-generated content, providing a faster turnaround on the iterations of game design and development.

In parallel to automatic testing, the need to generate more content with equal effort implies the need to have more music backgrounds to diversify the audio atmosphere of the game. Either audio designers select the music pieces from a catalog of recordings, or they can use a sound synthesizer that produces music from digital scores. In both cases, the selection of music must be done according to the target ambiances of the game levels. This task may become difficult when the development duration is short or when the volume of generated content increases exponentially. In these cases, AI methods can provide automatic tools to help game audio designers in the selection and production of music for video games.

AI4Media is an EU co-funded research initiative with 30 technology and media partners aiming to explore diverse aspects of AI in the media sector, advance AI technology research, and develop specific solutions for seven use cases. Led by MODL and IRCAM, one of these use cases focuses on AI for the gaming industry.

This use case focuses on two main features of using AI

1. "Automated testing for Games" implies the use of AI agents that can automatically detect bugs and glitches in video games by autonomously playing them and compiling test reports as human testers would do. The testing process usually involves manually playing a videogame and reporting the problems encountered. However, the future of the game market sees an increasing adoption of the game-as-a-service model. Therefore, quality assurance (QA) becomes a vital part of the game development cycle, as the new content produced has to be continuously tested.

2. "Improved music analysis and synthesis for Games" aims to help developers to choose or to conceive background music or sounds. In video games, as in cinema, the audio track has the role to strengthen the atmosphere of the present scene, and it has an impact on the emotion felt by the game user. This aspect is important, because an unsuitable audio track may break the mood of a game. Developers can either choose music from a catalog, or they can delegate this task to music composers or sound designers, but automatic tools may help, especially in the case of dynamical changes and procedural content generation.

In this white paper, we explore and discuss the key challenges and problems encountered by the industry during the game development process and showcase a possible solution using AI, for both features: automatic testing and music synthesis.

In order to build these features, we need to understand the essential requirements for these tools. For this reason, we organized a survey among our industry partners asking them about their development process, the challenges they face, and their disposition toward AI-assisted tools. We focused on their opinion of using and integrating AI solutions into their current development pipeline and their main concerns about AI-assisted tools.

While we are able to get valuable answers, unfortunately, due to how development pipelines are treated as trade secrets in the game industry, we were not able to collect more than 9 responses for our survey on automatic testing, and 2 for music analysis/synthesis. These responses are collected from 5 different small to medium-sized companies or laboratories located in Denmark, Malta, France, Portugal, and a "worldwide" company. These 11 responses are collected from several different job titles in the game industry: 4 are programmers, 3 are working in quality assurance (manager, staff, or tester), 1 is a designer/project manager, 1 is a researcher in audio processing, 1 is a director in audio design, and the last one didn't label themselves.

Problems & Challenges Faced by the Video Game Industry

The Video Game industry has always been depending on humans for all facets of the industry. Companies hire people to do all types of jobs from creating content (music, 3D models, textures, etc) to testing and marketing the game.

This worked well since the beginning of the industry as games were small due to technology limitations and could be created by a small number of people. These days, however, this does not work anymore as games have become big in size and scale with a lot of game content that takes the player between 20 to 60 hours to experience. Other games changed the model to become a service that requires continuous production of content. This scale in the industry added more challenges for them and required them to hire more people over time to keep up with the demand. This is not sustainable anymore as the cost of production has been increasing exponentially to the point that if a game fails on the release, it might be the end of the development studio. **We believe that it is essential to integrate AI in the process of creating and testing games.** In this section, we focus on the different challenges that face industry partners during creating and testing games. We mainly look at the reasons preventing them from integrating AI-powered tools in their work process.

Challenges for QA Testing

Most glitches in video games are visual and easy to validate as a human. The main issue is that they are hard to discover as the searching process is similar to looking for a needle in a haystack.

QA testers need to repeat the same task under different settings to validate that a small area of the game doesn't contain any glitches. We believe that having an AI tool to support the QA tester will increase their efficiency and effectiveness while doing their job. By automating repetitive tasks with AI tools, QA testers can focus on analyzing the results and reporting the glitches to the development team.

We believe that the main reason that QA automated tools are not widely used is that **integrating these tools in the production pipeline would take a lot of time and money and might not be easy to achieve**. In addition, a lot of designers and developers don't want to use these AI-powered tools because they believe that this will cut human jobs. We validate these insights through our user study by inquiring what management tools companies usually use, the willingness to integrate these tools, and the purpose of their use.

When asked about their companies' **potential to invest in integrating automated QA testing tools,** 8 out of 9 users confirmed that their company is willing to do such investments. They further indicated that their company is **not looking to replace humans but to enhance their testing (100%) or replace an in-house automated testing tool (25%).**
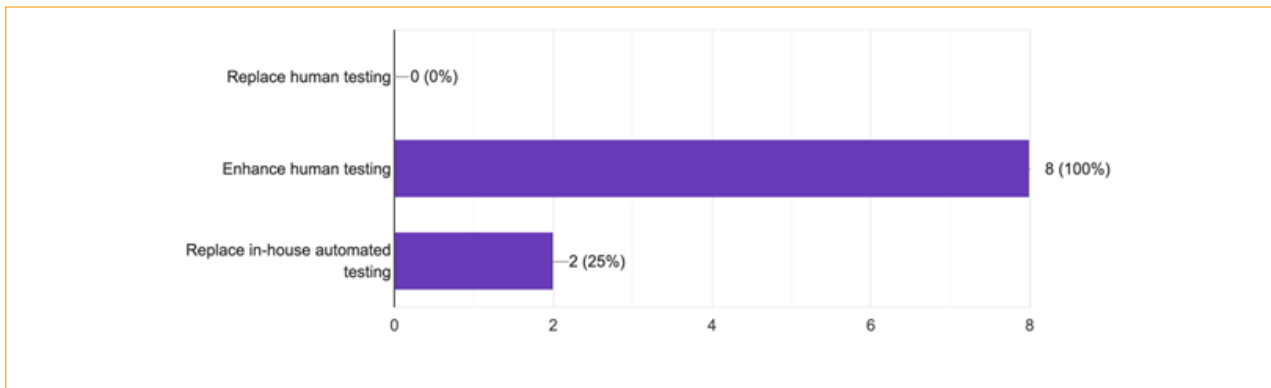

Figure 1: The survey results for the following question, "How would you integrate a QA solution using AI technology into your existing QA infrastructure?"

Looking into the companys' willingness to integrate automated QA testing tools, we asked questions related to how much time the company will allocate for integration and what management tools these systems need to integrate with. The answers showed that industry partners (75% of the users) **don't mind taking one week or more to integrate the system within their own workflow** instead of having a generic plug-and-play tool or working side to side with the developer of the tools to fit their needs. They also prefer that **the integration goes deeper within their workflow such that the results from the system get reported through their other used tools,** especially Git and Slack, instead of getting emails or checking a custom dashboard. We originally believed that one of the reasons for not using AI tools is the integration time. Although we found that our industry partners are more than happy to spend more time integrating, we think smaller studios would still benefit from having an effortless integration with known game engines as these studios won't be able to afford longer integrations.
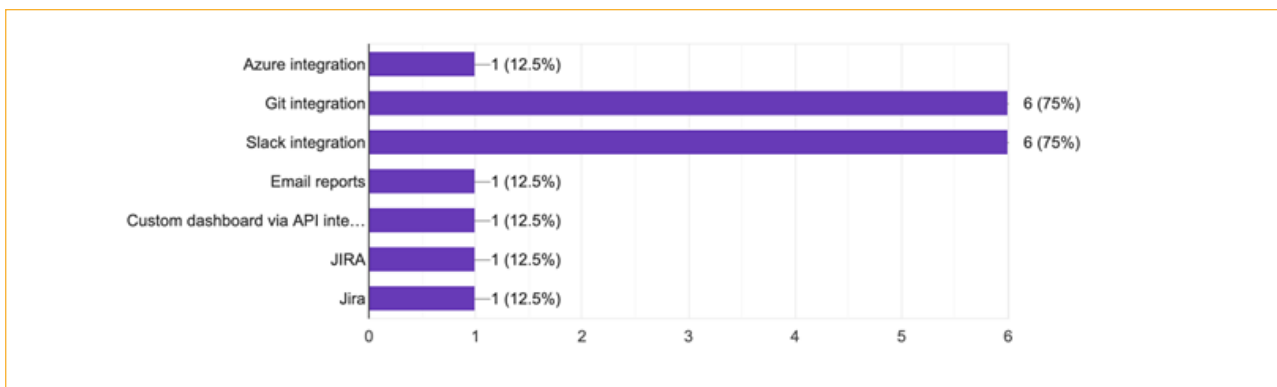

Figure 2: The survey results for the following question, "How would you integrate a third-party QA solution into your workflow?"

Challenges for Music Analysis/Synthesis

Like in cinema, background music is an important part of a video game because it helps to reinforce the atmospheric feeling of a game. However, the generation or selection of music for a game is not an easy task because of many constraints: not only the songs should be from various artists, but they must also fit the mood of the game or game scene.

For that, AI tools can help music designers of video games to produce music within the time constraint of development.

The survey we conducted stands up for these points: the task of manually finding music pieces in a music catalog is considered to be long and quite difficult. Even if the interviewed people have no predefined opinion about the use of AI, any technical tool which can help is welcome. Due to the lack of such tools, current practices use expert knowledge and manual annotations that are sometimes missing or inaccurate. When an automatic annotator of music can be used to detail and automatically recommend music, the most important labels are about the emotion/mood, the mode (minor/major), and the tempo. These labels inform about the ambiance and the "energy" of music, which is important in order to align the music with the target atmosphere of the game.

We believe that the automatic composition and arrangement is not a desired AI feature for the production of a video game, and the survey tends to confirm this thought. Nevertheless, one participant of the survey, a video game producer, commented that the gaming industry is not too interested in AI sound synthesis. According to him (or her), in most cases, game directors take on "human" composers and the music is pre-recorded into audio files using commercial samples-based sound synthesizers. Rather than online synthesis, this method allows a satisfying quality and lighter use of resources without the problem of latency. Nevertheless, with a further analysis of the comments, the door remains open to AI sound synthesizers as we initially believed. First, interactivity is an important aspect of the music for the game. Here interactivity means synchronization of the music with the gameplay, and dynamic change of the music with a quasi-randomization to diversify the audio playback together with a fine alignment to the game atmosphere. AI sound synthesizers can help this interaction. Second, automatic composition based on tags can give inspiration to musicians, and it can help the communication between directors and composers.

Brief Analysis of Industrial Needs

In this section, we focus on the needs of the industry partners. This is essential to understand how to build AI-powered tools. To collect these needs, we focused on understanding what features the AI system must be able to do, what are the controls for these systems, and finally, what outputs should it produce.

Industrial Needs for QA testing

With regard to the input and the output of the AI-powered QA testing tools, we believe that an AI-powered system should enable QA testers to have parameters that control the running tests under different settings.

These parameters should include information about what information should be recorded by the system and the type of glitches that the system is supposed to find. These should be enabled through an easy-to-use interface with predefined parameters and glitches types that the QA tester can choose from. The user study showed that 50% of the users wanted to **be able to define their own glitches and select the data being recorded,** while the rest either wanted to **select from predefined choices** (37.5%) or **have an automated setup that detects these parameters** (17.5%). This was a little counterintuitive to what we believed, but we think the reason for that is each game comes with its own types of glitches and problems. There might be glitches that are new and not predefined so it would be faster if the system enabled that without the need for updates.

Looking deeper into the different types of glitches, **users identified that framerate issues (when**

the game slows down while playing due to a lot of computation happening) and rendering glitches (such as wrong textures, wrong UV maps, etc.) are the most important type of glitches that need to be detected. We were surprised that the "rendering" glitch had a higher priority compared to the "getting stuck" glitch (the player can't move anymore) or network issues (for multiplayer games where the game slows down) because rendering glitches don't cause the game to be unplayable compared to "getting stuck" glitch or network issues. We think that the main reason behind that is that rendering glitches are harder to notice and take a longer time for QA testers to detect compared to other more critical glitches.
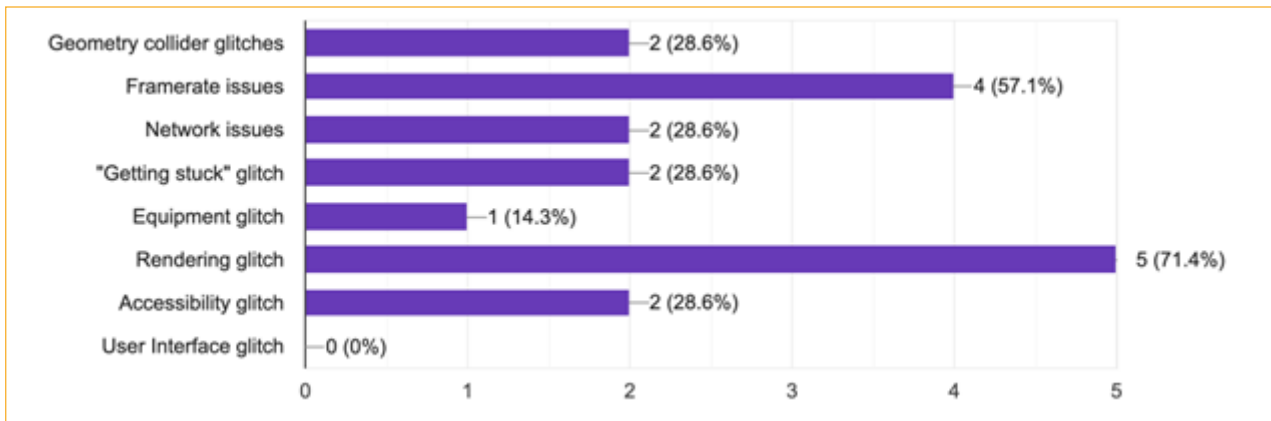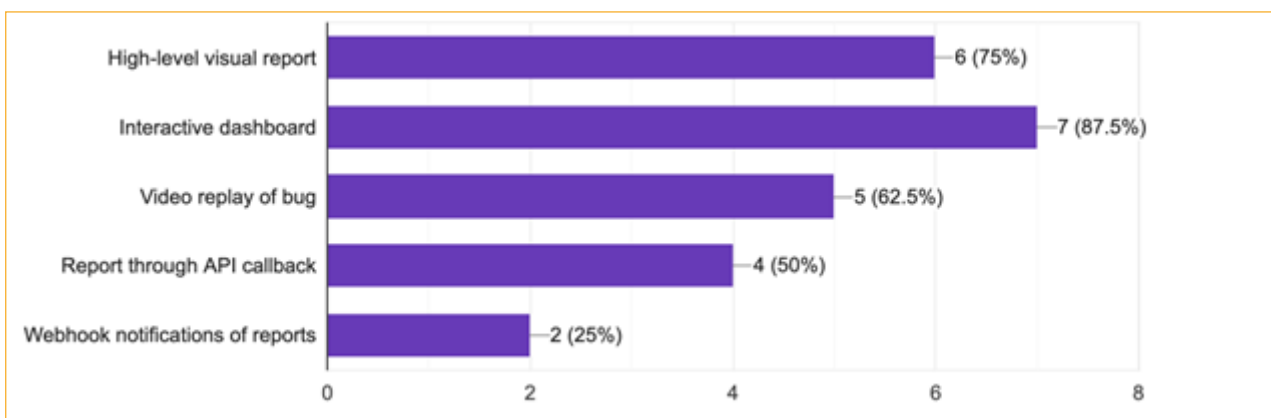

Figure 3: The survey results for the following question, "Which glitch type takes the most time to discover?"

We also examined the preferred method for viewing the results of QA testing. We believe that having an on-demand cloud-based testing tool with a webhook to notify the users through the management tool of their choice when the results are ready, along with a summary, is essential. The user should be able to go to the result page for more details about where the bug happened and how to replicate it through an interactive visual report. According to the survey respondents, we found that **having an interactive dashboard where QA testers can see the different events and glitches that happened in different areas of the game** is the most highly selected method; followed by **a high-level visual report which summarizes all the findings and then a video replay of the bug**. This was expected as developers and QA testers need to know the types of glitches that happened from the high-level report, followed by seeing the location and area in the game where the bug happened through the interactive dashboard and finally validate visually how the bug happened through looking into the recorded play trace video. These three are essential to be able to replicate the glitch, validate it, and be able to solve it.

The survey results for the following question, "What information would you expect to get from a third-party QA solution?"

Industrial Needs for Music Analysis/Synthesis

The survey for music analysis/synthesis has been addressed to people working in the audio design for video games: programmers, producers, managers, sound designers, and musicians.

The low number of responses, from people working in the audio design of games, may be an indication that people working on this topic do not feel concerned about the use of AI for music. Nevertheless, interesting conclusions can be extracted by analyzing the few answers and the free comments. The people who answered are both involved in video games and have a good knowledge of common practices. The survey is structured in two parts related to music: music sound synthesis and music recording analysis.

A part of our prior beliefs has been supported by the responses. The most synthesized instruments are synthetic and electric: Synth Bass, Synth Lead, Synth Pad, and Electric Piano. They are usually generated offline by commercial sample-based synthesizers. Then, the most important features that AI synthesizers may support are the sound quality and the generation of new sounds, while realism is not a key point (defining realism as the ability of a synthesized sound to be confused by a real instrument). As mentioned above, since synthetic instruments are very commonly used, in this case, realism does not make sense. According to an informal discussion, not included in the responses, it can be noted that for mobile devices and for retro-inspired games, this need for synthetic instruments is even stronger than for modern PC/console video games.

Also, two features are not important according to the responses: composition and arrangement, neither in an automatic way nor in an assistive way. An interviewed person has given a really interesting comment: this director is not convinced by AI composition and prefers to take on "human" composers. Nevertheless, when ordering a composition, it is sometimes difficult to communicate the wanted ambiance, mood, emotion, and music style. It is thus suggested that AI can be useful to make "inspiration" examples based on keywords, such as tempo, genre, instruments, composer style, and even based on music similarity.

A really interesting suggestion has been given as a free comment: although exploring a music catalog based on labels can be a helpful tool for this task (mood, tempo, etc.), it could be also really interesting to develop an AI tool that selects the pieces of music which are the most similar to a target ambiance. For example, this target ambiance could be given by the name of a movie: "I want songs with atmospheres which inspire the "Ocean's 11" movie".

Trustworthy AI in Video Games

Besides analyzing the industrial needs for AI for QA testing and music analysis/synthesis, we wanted to understand and analyze the need for Trustworthy AI.

Trustworthy AI is an important topic as it helps us to understand the industry concerns related to AI and highlights the parts we need to focus on more during the integration of AI-assisted tools with their systems. The term Trustworthy AI describes AI functions and systems that implement one or more of these features:

- Preserving the privacy of personal data
- Complying with applicable regulation
- Transparency and accountability in handling data and using algorithms
- Explaining how predictions/outcomes are reached
- Listing about processes mitigating bias
- Detailing the level of technical robustness

Looking first at the needs for QA testing, we noticed that 6 out of the 11 survey participants were not familiar with the concept before our description. Everyone except three users agrees that having trustworthy AI is useful in using AI tools in the game industry. The following table contains the distribution of the user choices on the different topics of Trustworthy AI. The user was prompted with the above six topics where each topic has a small paragraph explaining it in more detail to avoid confusion. We asked them to rate each topic from 1 to 5 on how important this concept is for AI tools.

Looking at the data, we noticed that the sheer majority of respondents (90%), agree that privacy

protection and legal compliance are very important concepts. This was not a surprise as the game industry is less transparent about their projects and techniques in order to stay ahead of the competition. Robustness was also marked as important. This could be because of the fact that the industry doesn't want to invest in a system that might skip a couple of game-breaking bugs. Allowing this in a system won't be cost-efficient as the company will still need to hire extra testers to double-check everything.

On the other hand, transparency, explainability, and bias mitigation scored a little less in comparison. It is possible that bias mitigation is not the main concern during QA testing nor in game development in general due to the entertainment nature of video games. When it comes to transparency and explainability, glitches and bugs in games are rather self-explanatory and bugs or glitches found by an automated system are easy to validate with human testers in a relatively short amount of time. An interesting finding is that while the opaqueness of many AI tools might be a point of concern for companies, some development teams are less considerate in this regard. Apart from Legal Compliance, some developers (especially those working in highly specialized positions, such as audio managers) put much less weight on the different aspects of AI trustworthiness.

	1	2	3	4	5
Privacy Protection	0	1	0	4	5
Legal Compliance	0	0	1	3	6
Transparency	1	0	3	6	0
Explainability	1	1	3	3	2
Bias Mitigation	1	1	2	4	2
Robustness	1	0	1	4	4

The survey summary for both sub-use-cases (automated AI-testing and sound synthesis) for the need for trustworthy AI in video games

Conclusions

This white paper discussed the challenges and needs of the video games industry regarding the use of AI-powered tools. This was achieved through running two surveys for both subcases: QA testing and music analysis/synthesis; and also based on our experience in the video games industry.

Most professionals in the game industry are excited about the usage of AI to help in their production pipeline. They want tools that help the QA testers and music composers instead of replacing them but noted that the tools should be able to integrate well with their production pipeline even if it will take more time to integrate.

In QA testing, QA testers need to have fine-grained control over the parameters of the test and not only depend on automatic setup. They also need a multitude of different output formats (short summary, video replays, and interactive board) that enable different people working on the game to figure out if a bug has happened or not, where did it happen, when did it happen, the conditions that caused it, and see a full replay, in order to easily replicate the bug. Concerning AI-based music analysis and synthesis, according to one participant, it seems that the industry is not that much interested in using AI-powered tools to compose new music, but rather in using them as an inspiration for musicians by creating inspiration pieces based on input written tags. Given the role and the company size of this participant, this thought can be judged as reliable.

Finally, we wanted to understand how the industry feels about Trustworthy AI. The survey showed that the most important aspects of trustworthiness for the video games industry are robustness, privacy protection, and legal compliance; while the rest (transparency, explainability, and bias mitigation) are considered less important. This is due to the nature of the game industry, which doesn't lend itself to these problems compared to the rest of the media industry. This does not mean that these other aspects are less important for the whole game industry but only for the usage of AI for testing and music synthesis. For example, using AI to detect cheaters in online games could include bias and it is very important for the game developers to be aware of that.

Looking at the AI tools that we are developing for this project, we can see that our initial intuition aligns with the industrial needs. For the AI for game testing, our tool provides a dashboard that enables users to upload and test their games and view the results of the test. We are also working on an easy plugin for the Unity game engine to allow testers and developers to specify what type of bugs and glitches they are trying to detect and provide them with easy integration. For music synthesis, the current development allows the user to search for new music based on the target game mood and returns snippets of this music. The next integration will focus on music similarity to be aligned with some output of the survey.

The analysis of survey results shows that there is much scope for future research, AI solution development, and potential impact in the areas of QA testing and music synthesis for games. New and trustworthy AI-based solutions are highly welcome and needed to enhance and facilitate game development workflows in a highly competitive industry.

AI4media

Follow us @ai4mediaproject

info@ai4media.eu   www.ai4media.eu