Adoption of AI technology in Music Mixing Workflow: An Investigation

Soumya Sai Vanka[1]

Maryam Safi [2]

Jean-Baptiste Rolland [2]

George Fazekas [1]

[1] Centre for Digital Music, Queen Mary University of London

[2] Steinberg Media Technologies GmbH, Hamburg, Germany

Queen Mary University of London

AI + MUSIC

UK Research and Innovation

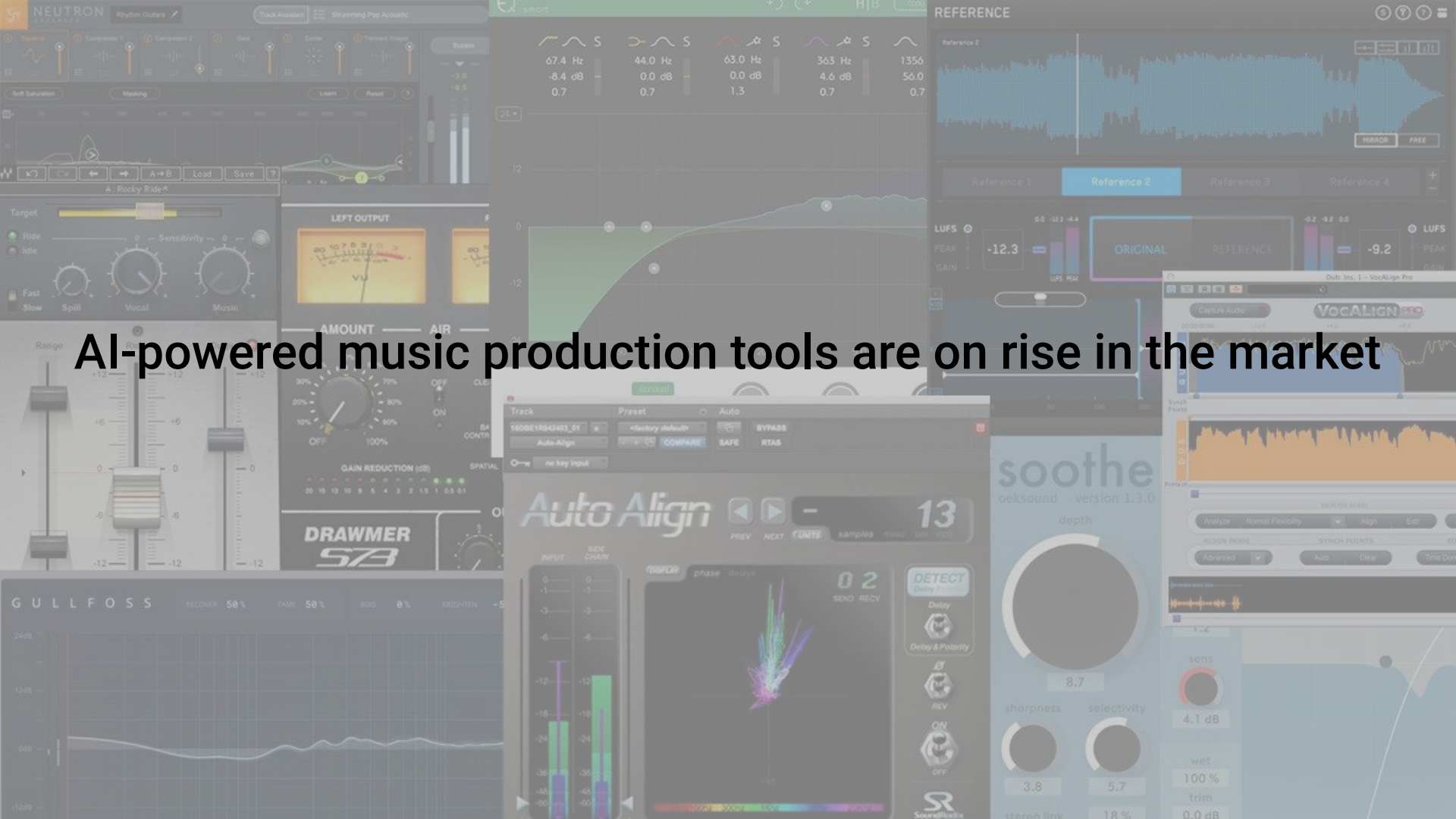
steinberg

“**I place automatic mixing alongside magic wallets, time machines and transporter beams.**”

- The Elf (user on Gearspace forum, 2013)

Man and his environment participate in moulding each other. Man is now in the position of actually creating the total world in which he lives, or what the ethologists refer to as his biotype. In creating this world he is actually determining what kind of an organism he will be.” - Edward T. Hall

Introduction



AI-powered music production tools are on rise in the market

Observed skepticism and doubts about these tools

Reasons:

- Fear of being replaced
- Doubt on subjective tasks
- Lack of understanding of AI reasoning

Towards a Human Centric Design Framework for AI Assisted Music Production
Tsiros, A. and Palladini, A. *NIME 2020*

Objectives of the paper

- Examine the attitudes and perceptions of users
- Figure use case scenarios and expectations from AI tools for user type
- Motivate the creation of suitable AI-based tools catering to the needs of different users

Audio Engineering Society

Convention Paper

Presented at the 154th Convention
2023 May 13–15, Espoo, Helsinki, Finland

This paper was peer-reviewed as a complete manuscript for presentation at this convention. This paper is available in the AES E-Library (http://www.aes.org/e-lib), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Adoption of AI Technology in the Music Mixing Workflow: An Investigation

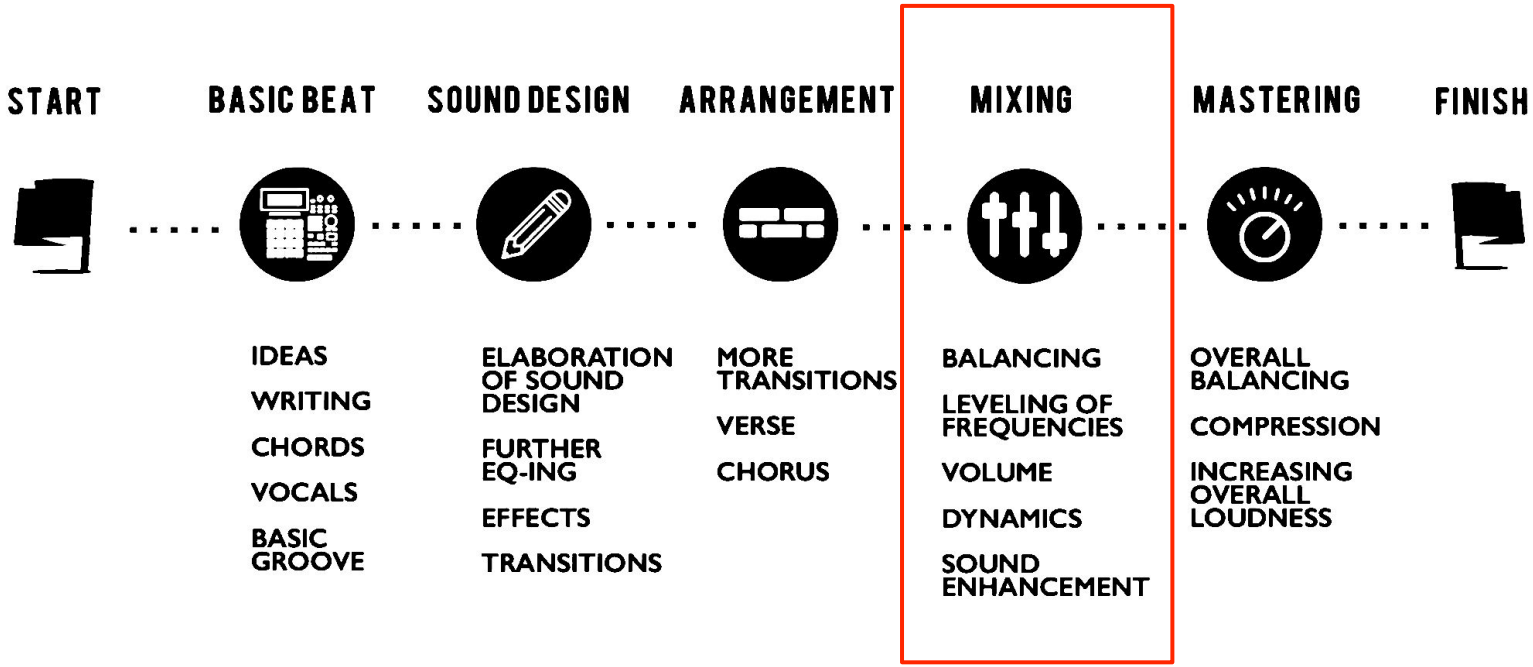
Soumya Sai Vanka[1], Maryam Safi[2], Jean-Baptiste Rolland[2], and György Fazekas[1]

[1]*Queen Mary University of London, London, UK*
[2]*Steinberg Media Technologies GmbH, Hamburg, Germany*

Correspondence should be addressed to Soumya Sai Vanka (s.s.vanka@qmul.ac.uk)

Background

Music Production



Mixing

Audio mixing is the process of blending multitrack recordings

- Technical considerations together with creative, artistic or aesthetic decisions

Achieved with audio processes/effects

- Gain
- Panning
- Equalization (EQ)
- Dynamic range compression (DRC)
- Artificial reverberation

Democratisation of Music Production

- Less cost of production because of digital technologies like personal computers, internet, DAW
- Lower cost and easier access to music production tools = more people producing music
- Access to tools for a more diverse group of users

A creative industry in transition: The rise of digitally driven independent music production
Hracs, B. J. *Growth and Change, 2012.*

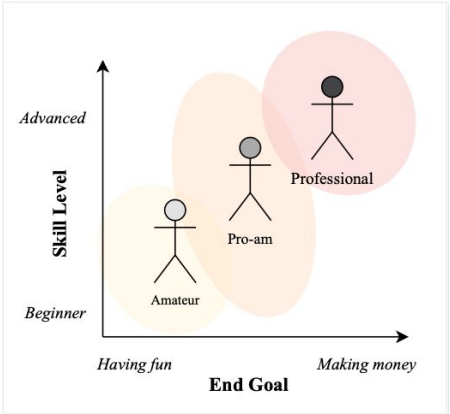
Democratisation of Music Production

- Three different categories of users from the technical literature
    - ➔ Amateurs
    - ➔ Pro-Ams
    - ➔ Professionals

A creative industry in transition: The rise of digitally driven independent music production
Hracs, B. J. *Growth and Change, 2012.*

Amateurs

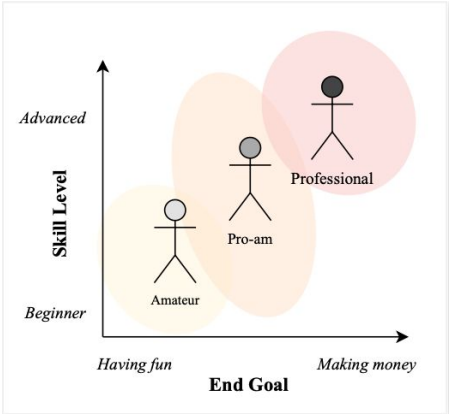
- New to mix engineering, without formal training or experience
- Hobbyists or musicians learning the basics of mixing
- Typically less skilled, unpaid
- May remain in this category for a prolonged period if mixing skills are used occasionally



Coming in from the margins: amateur musicians in the online age
Hoare, Michaela, et al. *SIGCHI 2014*

Pro-Ams

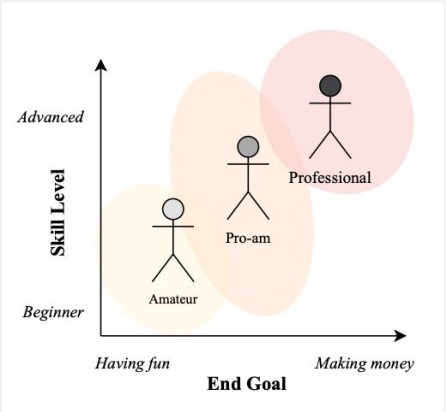
- Work to professional standards without the same infrastructural support as professionals
- Passionate about mixing music and desire to improve skills, but not necessarily making a living from it
- Mix engineering may not be their main focus in the field



New amateurs revisited: Popular music, digital technology, and the fate of cultural production.
Prior, Nick. *Routledge handbook of cultural sociology, 2018*

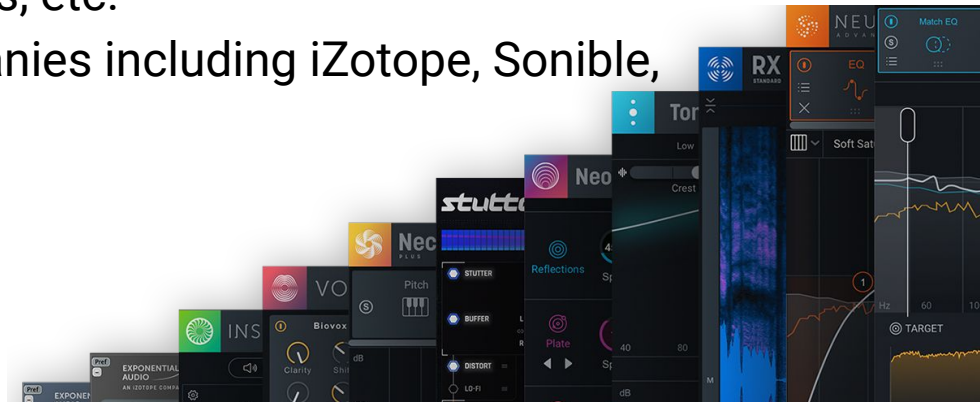
Professionals

- Highly skilled, paid for their services
- May have formal education or tacit knowledge gained from traditional studio environments
- Have good connections and infrastructural support for equipment and facilities
- Expected to continuously improve and update their knowledge and skills
- Considered experts in mixing and held to a higher standard than amateurs or hobbyists



Smart Tools for Mixing Workflows

- **Automatic mixing**: Any system that can create a mix given the stems or raw audio.
  - RoEx and iZotope Neutron offer automatic mixing services and tools, respectively.
- **Assistive mixing tools**: Use AI technology to assist in mixing workflows.
  - Includes AI-based plugins and processors such as smart equalizers, smart reverbs, smart vocal riders, etc.
  - Offered by several plugin companies including iZotope, Sonible, Focusrite, and more.



Investigation Methodology

Step 1: Semi-structured interviews with Professionals: *for theory building*

Qualitative analysis of the transcripts to generate **themes and inferences**

Step 2: Generate questionnaire: based on inferences
Check agreement amongst larger pool

Professionals and Pro-ams

Step 3: Going through internet forums:
Collecting opinions of wider user groups

Amateurs and Pro-ams

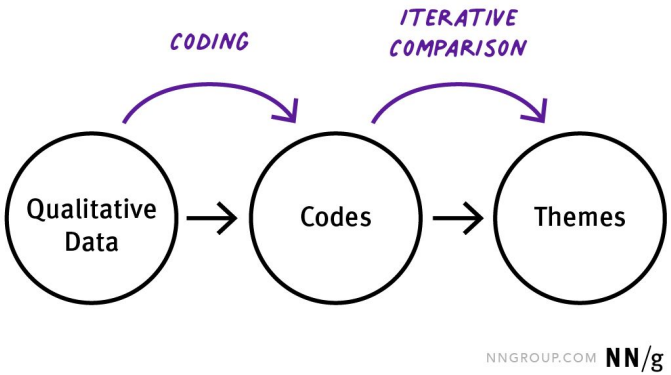
Semi Structured Interviews

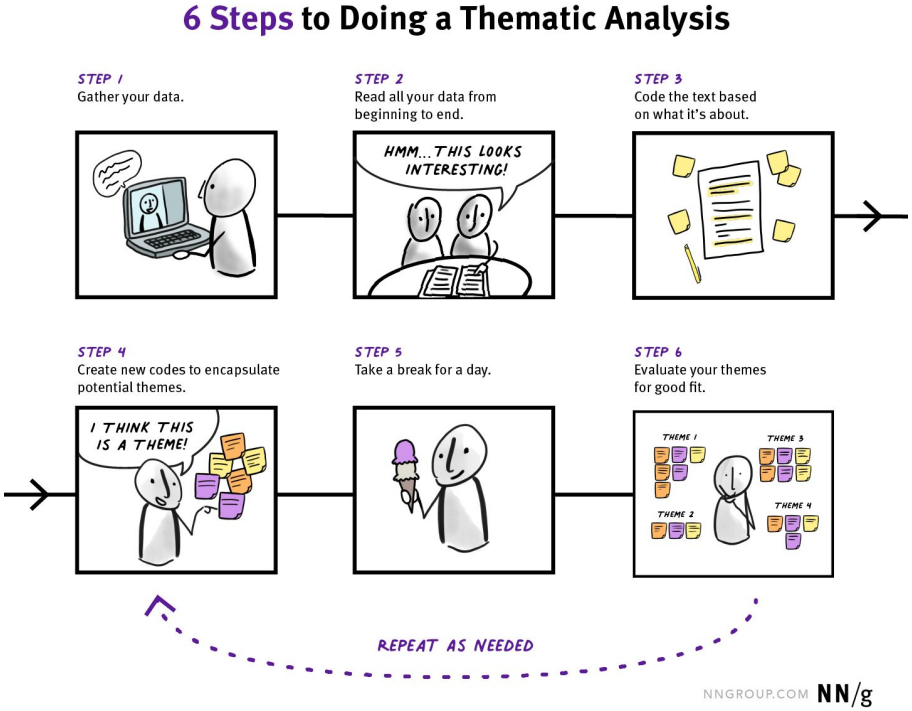
Participants: Professional Mixing Engineers

Criteria: More than 5 years of experience, Advanced to expert mixing skills

- Semi-structured interview
  - qualitative research method
  - predetermined set of open questions (questions that prompt discussion) with the opportunity to explore responses further.
- Purpose: Preliminary theory building
- Interviews:
  - Theme: Use of smart technology in mixing workflow and expectations
  - Online on Zoom calls
  - Recorded and transcribed
- Data Analysis Method: using MAXQDA
  - Qualitative analysis using grounded mixed method approach
  - Generate themes and inferences

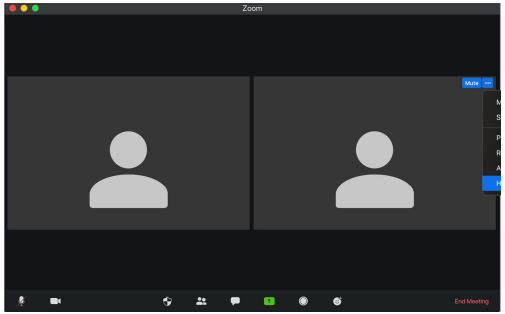
Thematic Analysis





Thematic Analysis
Braun, V. and Clarke, V. *American Psychological Association 2012*

https://www.nngroup.com/articles/thematic-analysis/



Online Interview

smart technology to assist in mixing

smart technology to assist in mix

creativity

about.
416 80 --> 10
[redacted] (Guest)
The idea that.
417 70 --> 20
[redacted] (Guest)
With Co creativity,it provides more time for creativity.
418 0 --> 60
Soumya Sai Vanka
Here.
419 0 --> 0
[redacted] (Guest)
It's actually any and if the system can make suggestions
That you wouldn't have thought of as a human cause.Ho
you possibly think of the you? You maybe can think of fiv
amazing creative ideas in the mix if you're lucky,whereas
system can maybe.
Hit you with 100 great ideas and you can choose but,but
where the computer system is interacting with the humai
where it's really interesting where where the computer sy
could say,ah,OK,so I see that you're doing this.How abou
420 0:48:0.150 --> 0:48:0.430
Soumya Sai Vanka

Transcribing the interview and qualitative coding the transcript

I just say Ohh it's an intelligent tool. The reason they get so threatened is because they think it's going to stifle the creativity, so we need to reeducate them and say no, this gives you more time for creativity because actually,

Code: ● smart technology to assist in mixing Weight score: 0
Transcripts\[redacted] Position: 413 - 413

I prefer the people not to send me the entire session on how they created it or rather sending me the stems as you said just because then I can organize my own session without losing as much time as I would get in somebody's Association on re-aducting it to my idea so yeah but at the same time when um organizing that and so uh getting stage in everything and coloring the coloring the trucks and putting them on my on the order I've got set up in my hair in my head uh to make it easier for me to work on it it's I always play the song even if it's not mixed at all and it's all a mess but in that mess I can oh I can already kind of get used to how it sounds it can kind of get uh uh understand the vibe and all those things so it is time consuming uh most of the time it is boring but at the same time can it's I feel like it's already part of a job and can get me uh uh to a point where I'm done with this preparation but I kind of already know the uh where the song may lead to and I always I never do like one task of a time so let's see well if I if I'm call or calling and not playing a song because that's just waste of time and a uh but if I'm if I'm already getting staging it if that is needed obviously uh uh I do it again stage but at the same time I try to balance uh the the instruments as well yeah to like what it could be called like a rough mix or just you know just balancing find them a place in the in the panning Fields this kind of thing some things that I do uh at the same time

Code: ● smart technology to assist in mixing Weight score: 0
Transcripts\[redacted] Position: 52 - 52

I think it must be questioned at all times the role and I think it's great at doing things which give you more time for creativity, so labeling and chord recognition instrument, all of those things which are really boring and maybe doing rough balance

Code: ● smart technology to assist in mixing Weight score: 0
Transcripts\[redacted] Position: 62 - 62

Interesting aesthetically, it then gives you more time to then do the thing that maybe the computers not so good at doing, which is to do the the wrong stuff

Code: ● smart technology to assist in mixing Weight score: 0
Transcripts\[redacted] Position: 65 - 65

Drawing themes and inferences from the codes

Structured Questionnaires

Participants: 22 pro and pro-am mix engineers

Criteria: Intermediate to expert mixing skills and more than 3 years of experien

- Structured Data Collection
  - Useful to gather quantitative data
  - Questions are asked in the set order to every participant
  - Used when there is clear understanding of the topic
- Purpose: Standardised enquiry to compare agreement of results among a larger pool
- An online questionnaire-based study :
  - Questions based on themes and inferences generated from first study
  - Hosted on Microsoft Forms and circulated online amongst handpicked engineers
- Data Analysis: Quantitative/ statistical analysis of results

37. Would you be inclined to use a technology that can apply plugin effects for you? *

- ◯ Yes
- ◯ No
- ◯ Other

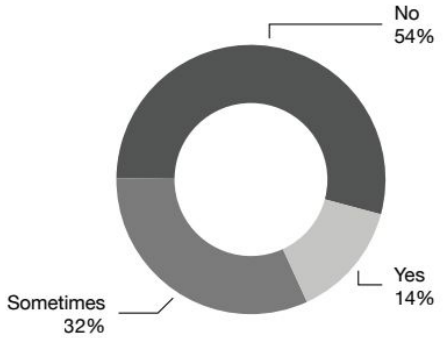
38. Would you be inclined to use a technology that can create mixes for you? *

- ◯ Yes
- ◯ No
- ◯ Other

39. Would you be inclined to use smart technology that can create a mix in the style of the reference song provided? *

- ◯ Yes
- ◯ No
- ◯ Other

Online Questionnaire Form

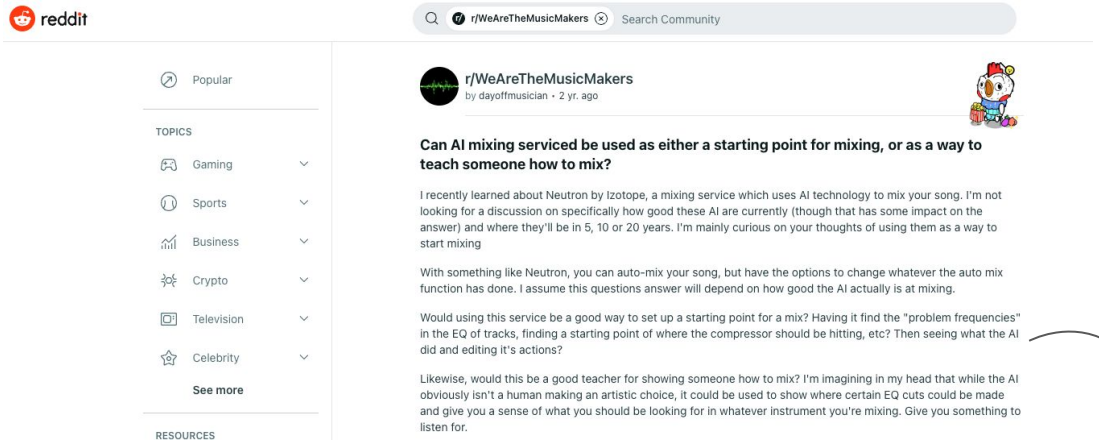


Quantitative analysis and visualisation of collected data

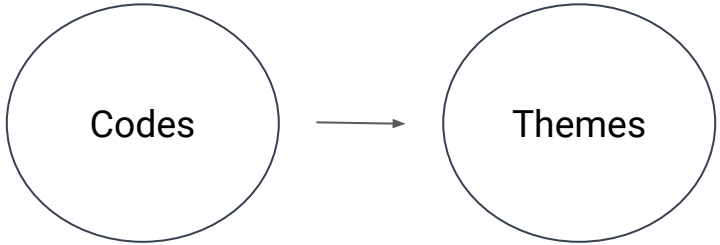
Unstructured Data Collection from Forums

Data from Internet Forums on AI-based Mixing and Mastering

- Collecting threads from forums on Twitter, Reddit, Quora, and more
- Purpose: Collect data about wider user group like the amateurs and enthusiasts that cannot be accessed in a structured way
- Data Analysis: Thematic analysis



Thematic Analysis



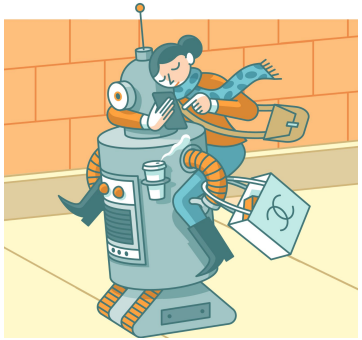
Outcomes

Amateurs

Primarily create and **compose music**

Limited knowledge of music mixing

Mixing: biggest **hurdle** to releasing music



- Expectations: **highly autonomous** mixing system
- Not expecting high quality output
- Using AI mixing systems: **produce a decent mix with minimal effort**
- **Positively embracing** the emerging technology

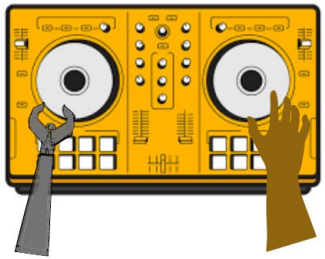
Amateurs

“The biggest barrier to me getting a recording out into the world is always the mixing process. It’s tedious, ... Give me a button I can push to automate most of the work.”

“I love recording my own songs, but I really can’t understand many of the concepts regarding mixing, I get easily lost with EQs, compression, etc. I’m basically an amateur, I respect the art of mixing, but that’s the phase where I lose the fun of making music.”

Pro-Ams

Higher technical skills than amateurs but less experience than professionals.



- **Use cases**:
  - **Improve their skills** and work towards becoming professionals
  - quickly **achieve a certain sound** or style in their mixes.
- Aware of the **limitations** of technology - willing to put tools to best use.
- **Cautiously optimistic** about the future of these tools.

Pro-Ams

“It could be used as a tool to learn the basics, then using the reference method, as well as trial and error, it’ll be another tool in the tool belt of learning. AI being used for basics, other methods to make more exciting mixes.”

“None of this AI mixing is going to mix your music for you. It’s going to make basic suggestions of moves you can make to better process the audio. You, yourself, still have to take the brunt of the decision-making and the work.”

Professionals [Positive]

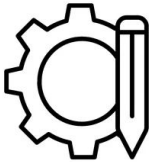
save time on repetitive tasks

experiment with new sounds

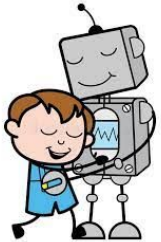
tasks like filtering, peak detection, pitch detection, mastering, equalization, and sound enhancement

Use Case

Accurate and precise



customizable



assistive and co-creative technologies that enable collaboration

Expectations

Phase	Category	Tasks
Before Mixing	Mix preparation*******	Labelling and colour-coding tracks***
		Grouping and arranging tracks
		Importing session data
		Setting up sends, AUX, and buses**
	Editing	Trimming Silences
		Identifying and fixing clicks and pops
		Identifying and fixing phase issues
		Checking mono compatibility
During Mixing	Levels Balancing	Gain staging**
		Setting a rough fader mix***
	Spectral Corrections	Identifying and fixing masking issues**
		Auto-high pass filter for content with low bass content
		Auto EQ and EQ matching***
	Communicating	Getting artist feedback on time-stamps
	Co-creativity**	Suggesting ideas for mix**
		Recommendations for audio effect processing

Table 2: List of smart functionalities that mix engineers desire to ease their workflow; Frequency of the requests for the tool is represented by the number of asterisks (*)

Professionals [Positive]

“You maybe can think of five amazing creative ideas in the mix if you’re lucky, whereas the system can maybe hit you with 100 great ideas and you can choose”

“Machine could process as well and then give me the controls to just adjust everything. It would be much more helpful than starting from scratch.”

“I’ve used Neutron’s masking feature to get some extra clarity but your talking one move of the potentially hundreds of thousands of moves involved in a production.”

Professionals [Negative]

Cannot fully replace the human touch and creativity required in the process.

Traditional methods of mixing are superior - learning by trial and error best way to master mixing.

Leads to a loss of control and precision in the final product

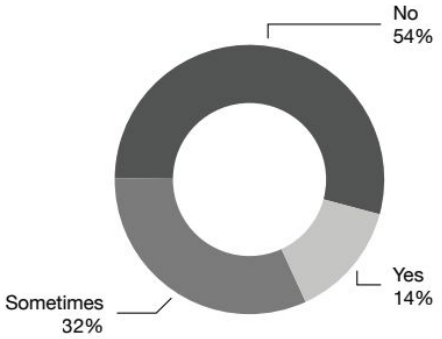
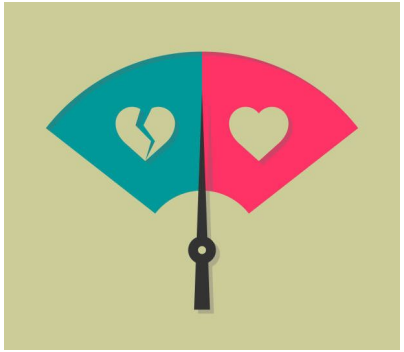
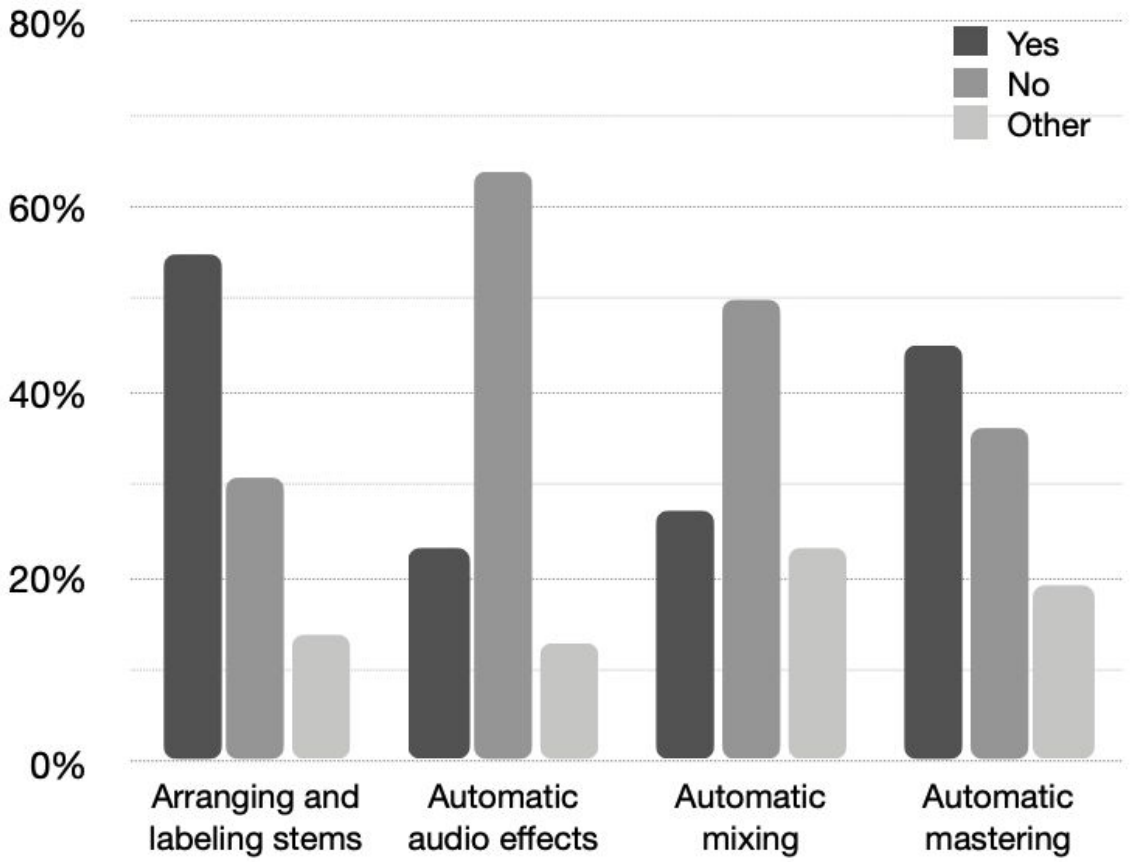


Fig. 2: Responses to the use of AI-powered tools in mixing workflow as reported by pro-ams and pros.

Professionals [Negative]

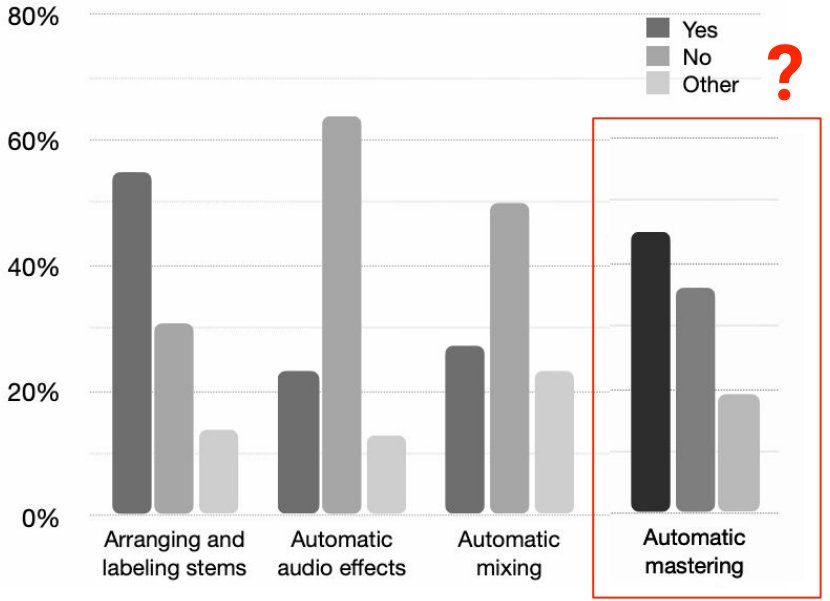


“For a hobby, they can do that at home. On a professional level in a production setting, it won’t fly because client and corporate revisions are demanding and very precise in what they want on a granular level. It’s NOT about doing it faster but with precision.”


80%
60%
40%
20%
0%
Yes
No
Other
Arranging and labeling stems
Automatic audio effects
Automatic mixing
Automatic mastering

More Demand for Mastering Services?

- To adjust the overall dynamic loudness of draft mixes while sharing it with the client during the mixing process
- Output is not used as final mastered product

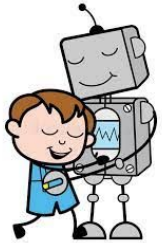


Smart tools\User	Amateurs	Pro-ams	Professionals
Use-case	Create decent mix	Learning and exploratory tool	Automate repetitive and time-consuming tasks
	As a learning tool	To find a starting point	Co-creation and assistance
		Automate technical tasks	To find a starting point/direction for mix
		Creativity and inspiration	Creativity and inspiration
Expectations	Autonomous with less control	Advanced and more control	Highly advanced and wide range of control option
	Cost-effective	Accurate and precise	Accurate and precise
	Easy to use	Assistive	Assistive
		Cost effective	Easy integration in current work-flow
		Easy integration into current workflow	Context-aware
Sentiment	Positive	Cautiously positive	Mixed

Table 1: Comparison of use-case, expectations, and sentiment amongst different categories of users of AI technology in mixing workflows

Future of AI-based Tools

Key Factors for Success of Smart Mixing tools



- Interaction models that facilitate trust
  - lack of interpretability and control - barrier to their adoption.

- High precision and quality of results generated
  - low-quality output not useful in professional workflows.

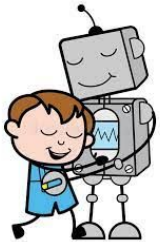
- Seamless integration into existing workflows
  - maximize efficiency and productivity.

Towards a Human Centric Design Framework for AI Assisted Music Production
Tsiros, A. and Palladini, A. *NIME 2020*

Balance of Control and Automation

Black box systems: limiting control and interpretability.

Different user groups desire varying levels of automation and control.

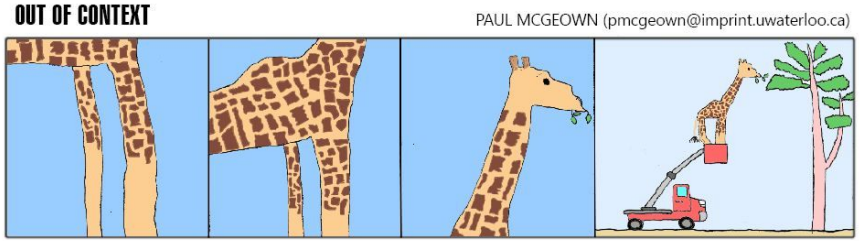


- Offer different levels of control and automation based on user expertise.
  - Same plugin can have different modes, ranging from full to no automation.
- AI-based mixing systems design goal: interpretability and controllability.
- Enhance user experience and productivity.

Towards a Human Centric Design Framework for AI Assisted Music Production
Tsiros, A. and Palladini, A. *NIME 2020*

Context-Aware and Precise Systems

Results are generic and systems do not work well for outliers



Context-aware systems

Providing context using text, audio, and semantic descriptors

Training on diverse music and audio data to work well for diversity

Output has artifacts and the quality not good enough for professional usage

Precise systems

Using black box systems generating audio with artifacts

Predicting processor parameters and applying it to raw audio (also offers interpretability)

Seamless Integration

Amateurs: may not be familiar or well-versed with DAW

- Autonomous mixing tools hosted on web
- Tools with simpler interface and less options to control

Pro-ams: may have established workflows but are open and curious to try new tech

- Web-based interfaces or tools that are simple to use
- Tools that will integrate into their workflow

Professionals: established workflows and familiar tools

- Should integrate into their existing workflow
- build tools that have similar formats and configurations to what these users are familiar with

neutone

by

Qosmo

Development of developer tools that enable the quicker and easier translation of research into technology. *ex: Neutone*

Lightweight systems to work on consumer laptops.

Opens door for remodelling and redesigning existing DAWs to use the increased power of consumer laptops and embedded hardware available in professional mixing consoles.

Context
using text, audio, semantics etc

User Interface
Allowing a way to provide context and control the result

Input

Output
Precise with no artifacts and in line with the context

User Interface
Allowing a way to interpret results and tweak them

Output

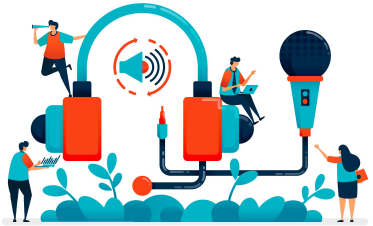
Tool Format
Seamlessly integrating into workflow

Design decisions to be taken based on the end user category in question

Outcomes

Bringing new talent, genres, and diversity into the music industry

Improving mixing workflows, increasing productivity, and enhancing creativity



Some Thoughts

- Being specific on whom a tool is designing for
- Offering control and automation based on target user group
- Optimising the quality of the model output and the interpretability of the output
- Lightweight systems to be used on consumer laptops
- Keeping in mind the three factors for success of AI-tools in the market

Reminder → The user groups are not rigid, rather are fluid