HUNDED: An approach to Remote Music Collaboration

ANEESHA LAKRA, Indraprastha Institute of Information Technology, aneesha18277@iiitd.ac.in
JASMINE KAUR, Indraprastha Institute of Information Technology, jasmine18287@iiitd.ac.in
SRIJAN JAIN, Indraprastha Institute of Information Technology, srijan18319@iiitd.ac.in

Music has existed since time immemorial, and there have been several technological advancements in the way we produce, record, and edit music. Rapid globalization led to the distribution of musical talents across the globe, making in-person collaboration increasingly difficult. After extensive user-research, the above hypotheses were validated, and additional key challenges faced by musicians while remotely collaborating/jamming were identified. These include audio latency issues, complicated interfaces, lack of visual cues, and the absence of a conducive environment. The above challenges have made it difficult for musicians to collaborate and jam remotely; this project explores the use of Virtual, AI, and Social contexts to build solutions for the same.

Additional Key Words and Phrases: music, social, collaboration, jamming

ACM Reference Format:
Aneesha Lakra, Jasmine Kaur, and Srijan Jain. . HUNDED: An approach to Remote Music Collaboration. In *Proceedings of (Design of Interactive Systems, Monsoon 2020).* ACM, New York, NY, USA, 47 pages.

1 MOTIVATION, PROBLEM STATEMENT, PROOF OF ITS SIGNIFICANCE

The COVID-19 pandemic has brought a global shift in trends; everything is happening remotely. Art forms where collaboration is crucial, like theatre, dance, and music, have been affected negatively. While professionals have been finding alternatives to keep working during these times, other individuals have taken up these art forms to connect with themselves and their peers. On exploring these forms of art, we found that remote music collaboration has been of great interest, even before remote collaboration was a hindering factor. This motivated us to explore the shortcomings of current solutions and investigate how collaboration platforms could be made intuitive and cater to the needs of people today.
To streamline our problem area in focus, we materialized the following problem statement:

Musicians and enthusiasts are unable to remotely jam and collaborate due to audio latency issues, complicated interfaces, lack of visual cues, and the absence of a conducive environment.

This problem was legitimized through the user study we conducted. After conducting a thorough analysis of the interviews, surveys, and market research, we have summarized the needs of users in our demographic spectrum. We have also concluded that no solution currently caters to the former. The following findings prove our hypotheses:

- *Lack of visual/auditory feedback:* Current solutions are restrictive in terms of active feedback when the users want to interpret actions/parameter tweaks of others quickly.
- *Intuitive interface:* Current interfaces are too confusing for the users, and it is difficult for them to navigate through the interface to reach their goals.
- *Real-time collaboration:* Current solutions are makeshift, using outdated technologies to facilitate real-time communication. Since audio is a priority, even a minute amount of latency affects results.

2 REQUIREMENT GATHERING

To gather information for our project, we conducted Interviews, Surveys, Academic Literature Review and Competitive Analysis. They have been elaborated in the following four sections.

3 SURVEYS

After coming across the problem of remote collaboration in music, we wanted to gain a broad perspective of the problems related to such a collaboration. We formulated a rough problem statement to give us a direction to proceed with the interviews and keep our goals in mind:

Novice musicians and enthusiasts are unable to access instruments, learn and explore how to make palatable music due to absence of an environment which enables them to understand the nuances of music.

Surveys were an appropriate choice for quantitative data collection for us since these reach a wide variety of people. This could provide us with a plethora of problems associated with collaboration in a remote environment and the needs of a diverse set of people. Such a diversity would help us gain insights into the user needs that could have been skipped by us. We circulated our 2-minute questionnaire among 110 people. Click here to view to Survey.

3.1 Survey Analysis

- Our sample audience was from 14-53 years of age, most people of whom belonged to 18-23.
- People prefer in-person tutor vs. online tutorials.
- The technique remains the most commonly face challenged.
- We got a scattered response on the comfort with the music learning environment.
- The majority (85.5%) of the participants felt that a proper learning environment plays a vital role in learning.
- The majority of the participant's population preferred playing and learning theory simultaneously.

3.2 Learnings

We concluded that despite the availability of resources online, people prefer an in-person mode for their music practices. We also observed that collaboration is pursued by people with some experience in music, and not novice players. This led us to rephrase our problem statement:

Musicians and enthusiasts are unable to explore how to make palatable music due to the absence of an environment that enables them to come together and collaborate.

To understand the reasons behind our observations and get some qualitative insights, we conducted interviews.

4 INTERVIEWS

"Users may not be able to spell out what they want." -Aric Rindfleisch We conducted five interviews to understand users' aims, behaviors, and pain points. In user-centered design, it helps build a starting point for a product or service. We conducted semi-structured interviews, where the conversation with the interview is open to new points and discussions. We kept the focus of our interviews to find out about users' expectations and experiences with collaboration.

4.1 Semi-structured Questions

- How long have you been involved with music?
- What are your objectives?
- Have you tried collaborating with anyone on a project? How?
- While collaborating, what things are of utmost importance?
- What are the challenges you face while collaborating with another musician? the parameters which are important
- Have you ever tried to collaborate remotely? (if yes) how?
- If yes, how was the experience?
- If bad, what do you think was lacking? What else would you prefer? [If good] What made it good?
- How close have you come to accomplishing your goals remotely?

4.2 Interview Analysis

The interviews helped us gain more specific knowledge in comparison to the survey. Needs of people who have some experience in music were noted, irrespective of whether they had experience in collaborating remotely or not. We found that there were a few common needs of people which should be translated from in person collaboration to remote collaboration.

- *"I have a friend with whom I love jamming but that's been hard to do since we're stuck at our homes. We try video calling but can never tune in together, it's just not effective, there is no synchronization."* [U1]
- *"Collaboration requires not just motivation and determination but effective cooperation and creativity, which I think is hard to establish if you're not sitting together in person."* [U2]
- *"There's always a minor communication gap which adds up and the result isn't always good."* [U3]
- *"We found a way around our problems by making sacrifices in terms of collaboration. We divide their music piece into parts and assign parts to each other. Then one of us sits and stitches them together to create a final piece."*[U4]

We refined our problem statement as follows:

Musicians and enthusiasts are unable to remotely jam and collaborate due to audio latency issues, complicated interfaces, lack of visual cues, and the absence of a conducive environment

5 ACADEMIC LITERATURE ANALYSIS

Before beginning the literature analysis, we posed the following questions:

- How is remote collaboration tackled?
- Is remote music collaboration an area of interest of exploration?
- What work has been done in this field?

Our literature review aimed to answer the questions mentioned above.

Remote collaboration is not a new area of exploration. Researchers have investigated the extent of remote collaboration since the era of GUIs [1]. With the popularity of this domain, new technologies are worked on to make the process of collaboration intuitive and engaging. General understanding also forms regarding how such a collaboration should take place. Online collaboration should complement the real world with guides and nudges to make complicated tasks simpler for the user to complete. Moreover, this rendering of guides should be in real-time to ensure maximum output. [2]

Collaboration has a significant impact on music. Music collaboration is a form of contemporary social networking and crowdsourcing. This domain has been under study even during times when social interaction was possible. Daisyphone introduced collaboration in music in 2004 by targeting mutual awareness among users and allowing them to modify each other's work [3]. Netjack focussed on delays and enabled remote electronic musicians to perform as though they were operating the same computer in person [4]. Machine Learning techniques have also been tested to make up for losses during transmission [5]. MirrorFugie explored different interfaces to remotely communicate gestures [6].

However, during the recent COVID-19 times, when all work has to be done remotely, it is essential to provide an efficient remote music collaboration platform which caters to the needs of musicians effectively. Video conferencing is the adopted media of communication between artists and enthusiasts. When remote performances occur like physical world scenarios, it is called Networked Music Performance. However, musicians face the lack of an immersive and efficient medium of remote, collaborative jamming or playing and production of music. [7]

The literature review not only answered our questions, but pointed out what could be done to make the process of collaboration better. This motivated us to pursue the problem of collaboration in remote environments for musicians and understand the underlying factors which affect this process.

6 COMPETITOR ANALYSIS

"You should learn from your competitor, but never copy. Copy and you die." - Jack Ma

Why did we do a competitor analysis?

- To understand the general landscape in which our product will compete.
- Compare the unique qualities of the competition.
- Compare visual and language styles.

6.1 How did we select competitors?

We selected our competitors based on the market share captured by these applications in the remote music collaboration domain. We also diversified our selection by opting for various styles of remote music collaboration: live instrument based, digital music production and video based conferencing.

6.2 Competitor Comparison

Jammr:

- A barebones Desktop application that facilitates synchronization of audio between multiple users through a distributed metronome (a time keeping instrument to help musicians play in the correct time signature and beats per minute).

JamKazam:

- Cross-platform application with synchronization, recording, and limited support for audio effects.
- Users are able to discover other musicians and create sessions based on latency.

Zoom:

- Standard video conferencing applications that utilize real time communication to facilitate video/audio transmission.

PiBox:

- Web application for comments and suggestions on particular segments of music.

Feature	Jammr	JamKazam	Zoom	PiBox
Platform	Desktop	Cross	Cross	Desktop
Audio Mixer	N	Y	N	Y
Metronome	Y	Y	N	N
User statistics (latency, audio drivers, etc)	N	Y	N	N
3rd party Instrument support (Virtual Studio Technology instruments)	N	Y	N	N
Audio Effects (reverb, pitch modulation, etc)	N	Limited Support	N	N
Live broadcast	Y	Y	Y	Y
Discover nearby users (social)	N	Y	N	N
Spatial audio (volume/panning)	N	Limited Support	N	N
Chat support	Y	Y	Y	Y
Video	N	N	Y	N
Comments on specific timeframes	N	N	N	Y

Table 1. Comparison between four competitors

6.3 Learnings

Competitor analysis is an absolutely essential step that allows us to identify market trends. We learned:

- All competitor apps have a dark interface and support essential features such as a chat support and live broadcast.
- There are 2 forms of music collaborations: live and recorded.
- Latency has been given a lot of importance while designing these softwares

7 PERSONA SCENARIOS

We categorized our potential users into the following two categories, based on the assumption that both of them have experience collaborating with others in real life: [8]

Fig. 1. Rhea does not have experience collaborating online

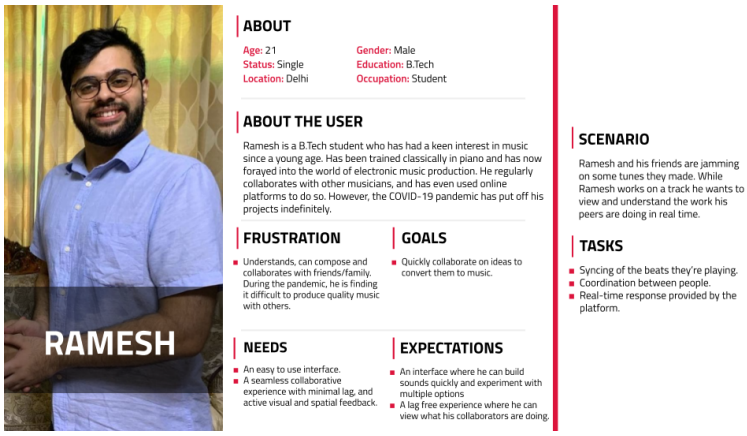

Fig. 2. Ramesh has experience collaborating online

By creating these two personas, we were able to classify our users in two broad categories. This allowed us to understand user needs, goals, and behaviours better This would help us visualise a solution which would satisfy the users and fit their expectations well.

8 AEIOU FRAMEWORK

We realized that recording observations and small details in the appropriate context is an important step to design interactive systems and hence we decided to apply the AEIOU framework.

Using the AEIOU Framework, we were able to understand the scope of our problem in various contexts.

Activities	Environment	Interactions	Objects	Users
Collaborating with others in real-time	Home and Music Studios	Playing a music instrument	Musical Instruments	People with music experience
Producing music, Exploring its different properties, Editing music digitally	Distractions in the background	Instructions being given on how/what to play	Music production software	

Table 2. AEIOU Framework

9 TYPES OF REQUIREMENTS GATHERED

Requirements Gathering equipped us with essential information about the various requirements needed to formalize a solution for remote music collaboration. We gathered four major types of requirements-

9.1 User characteristics:

Needs:

- Collaborating with others
- A seamless collaborative experience with minimal lag and active visual feedback

Expectations:

- An interface where sounds can be built quickly
- A lag-free experience where the user can view what their collaborators are doing.
- An easier interface
- A systematic environment

Goals:

- Collaborate with friends virtually.
- Quickly convert ideas into music.

Frustrations:

- Understands and can compose music, but wants to collaborate with friends/family.
- Finds digital music tools hard to use.

9.2 Functional/Data requirements:

Functional:

- The product should be able to react actively to user interactions.
- The product should have minimal complexity while still retaining most of the state of the art features currently available in most Digital Audio Workstations.
- The product should be able to create a channel of communication with multiple users to transmit both user actions/sound with minimal latency.

Data requirements:

- We require an audio database with temporally stable samples, which will be used to generate music and also learn how to mix different tracks to create palatable music.
- From the user, we'll need gestures/actions interacting with the application and current user state (sequencer, audio samples, etc.).

9.3 Usability and User Goals:

- An environment where user actions are converted to both visual and auditory feedback to enables a better understanding of the user actions.
- An open channel of communication between multiple users to allow greater collaboration. The user should be able to see changes in real-time and manipulate their system accordingly.

10 IDEATION

Reiterating on our vision elucidated in Section 1, the COVID-19 pandemic has brought a global shift in trends; everything is happening remotely. Art forms where collaboration is crucial, like theatre, dance, and music, have been affected negatively. While professionals have been finding alternatives to keep working during these times, other individuals have taken up these art forms to connect with themselves and their peers. On exploring these forms of art, we found that remote music collaboration has been of great interest, even before remote collaboration was a hindering factor. This motivated us to explore the shortcomings of current solutions and investigate how collaboration platforms could be made intuitive and cater to the needs of people today.

To streamline our problem area in focus, we materialized the following problem statement:

Musicians and enthusiasts are unable to remotely jam and collaborate due to audio latency issues, complicated interfaces, lack of visual cues, and the absence of a conducive environment.

The ideation process for the problem of remote music collaboration was tackled in the four steps elucidated in the following sub sections.

10.1 Ideation Planning

To understand the scope of the problem, we mapped out the observations made from the survey, interviews, literature review and the analysis of existing collaborative solutions for music in Figure 4. A mood board helped us in gathering and visualisation of inspirations for concepts as shown in Figure 6. Further, the idea of collaboration was diverged into sub-parts, namely feedback, music, interface and social aspects (Figure 5). The activity helped us understand

the subparts of collaboration which our potential solutions may cover. 5Ws + 1H (Figure 7) was followed to reiterate on the problem, and give a direction to ourselves for idea generation. The members of the group participated in the ideation techniques, where the stopping criteria decided was time. A hundred ideas were generated. The selection technique adopted was clustering, where we found underlying patters and formed groups, formulating nine concepts, and converging into solutions which catered to the broader problem of collaboration. [9]

Note: Kindly refer to Miro board cited for higher resolution frames of the following images.

Fig. 3. Defining scope of the problem

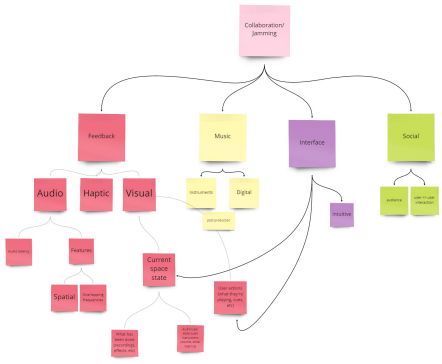

Fig. 4. Division of subparts

Fig. 5. Inspiration Moodboard

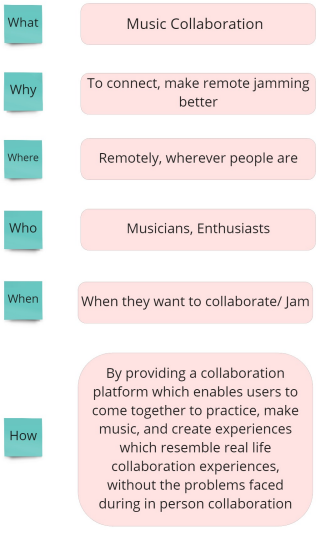

Fig. 6. 5Ws, 1H

10.2 Idea Generation

For idea generation, we followed the technique of brainwriting. We diverged with over 100 ideas generated, keeping in mind aspects of collaboration [9].

Fig. 7. Brainwriting

10.3 Idea Selection

We followed the method of clustering to converge to nine ideas which may be potentially prototyped. We grouped sticky notes to identify the following themes as shown in Figure 8:

- **Interface:** These ideas included solutions involving the interface for collaboration. Solutions were formed to target novel and easy to understand interfaces which enable intuitive interactions.
- **Interactions:** These solutions mention various interactions to target the subproblems of collaboration. They aim to make the process of collaboration more efficient and easier for users. Interactions to make collaboration more interesting have also been explored.
- **Social:** This caters to the various social interactions to facilitate collaboration and introduce novel ways of collaborating for jamming/creating music.
- **Environment:** This category includes some ideas which involved the creation of particular environments for collaboration.
- **Tech:** This specifies various technological ideas to improve the process of online collaboration.

The above allowed us to look at the problem of remote collaboration in music from various perspectives. We were able to form connections between different ideas which catered to the subproblems of jamming and collaboration. We created a dialogue on how ideas from multiple categories could be clubbed together to form solutions which address the target problem as a whole. This led to the formation of 9 ideas elaborated in the following subsections.

Note: Kindly refer to Miro board and Google Drive cited for higher resolution frames of the following subsections.

10.3.1 ***Concept 1: HAWLO.*** The first concept is inspired by holograms. Musicians will be able to see live, life-sized projections of their band members and will share a central, common mixer to control audio properties. Players see statistics along with every holographic representation of the collaborators which assist the participants while jamming. These encompass various audio, logistical information (key, scale, latency, etc). A performance can also be viewed by an audience. The audience views the session in the third person, and they're shown as personas, so as to not create information overload for the players [9] [10].

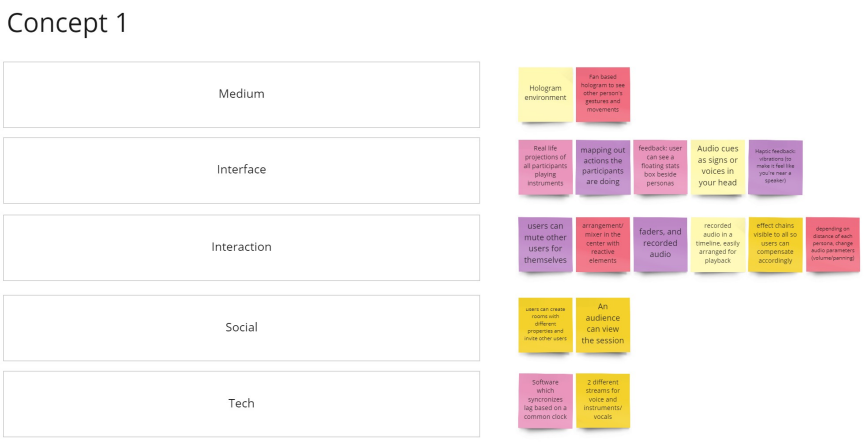

Fig. 8. Concept 1

Fig. 9. Concept 1 Moodboard

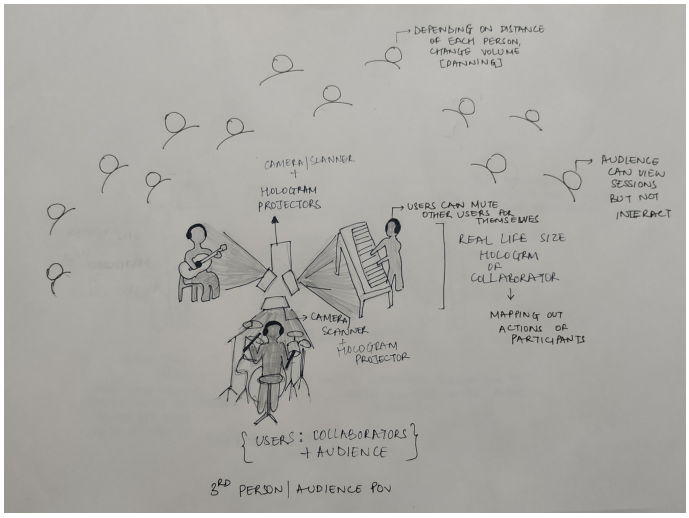

Fig. 10. Concept 1 Visualization (A)

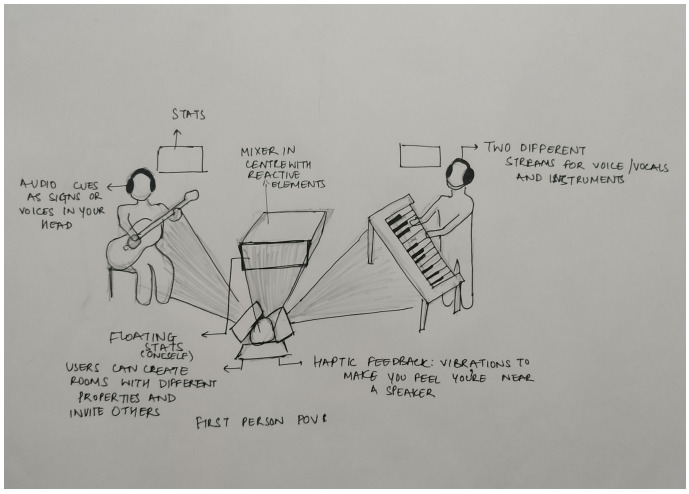

Fig. 11. Concept 1 Visualization (B)

10.3.2 ***Concept 2: JamFam.*** This concept represents users as avatars in a game-like environment which resembles a studio. Position of users relative to the room and other players, which can be manipulated using keyboard keys like arrows, directly affects their audio perception and the final recording. To collaborate using physical instruments, users can move their avatars in the environment and get associated with an instrument. The audio input is used to enable collaboration. A central mixer displays user effects and properties which aid in asynchronous editing of audio for every collaborator. This leads to a cleaner mix and better collaboration. If a user doesn't have a physical instrument, they have the option of utilizing a simplified sampler to create music. Users can join sessions either as an audience or as

collaborators which gives users the flexibility to explore the platform. They can also create bands/groups which then references pre-saved sessions. This leads to quicker setups and people can get right into making music. Users can also find other musicians/bands based on their common interests [9] [10].

Fig. 12. Concept 2

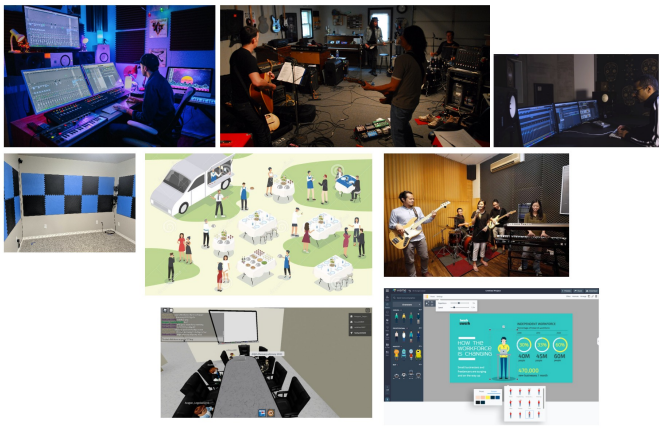

Fig. 13. Concept 2 Moodboard

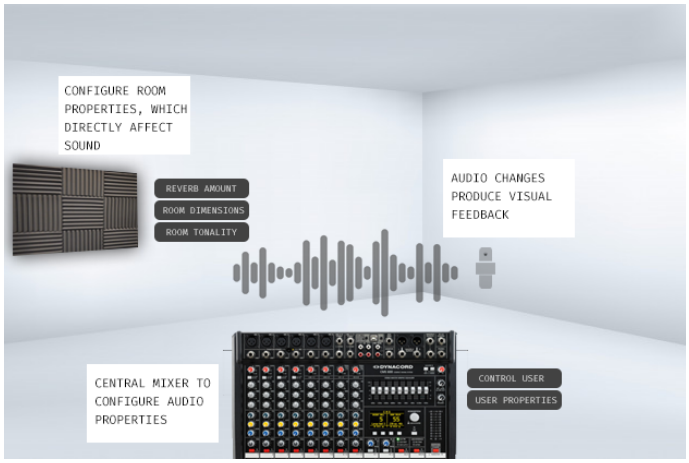

Fig. 14. Concept 2 Visualization

10.3.3 ***Concept 3: Band Baja.*** The following idea emulates a real-world music studio, where the users are mapped to their personas in the environment. Their gestures and actions can be tracked corresponding to their actions in real life. Eye/hand tracking for cueing is present so that real-world studio communication can be emulated efficiently. Users can create rooms with certain acoustic properties that affect the sound in different ways. This concept is inspired from Virtual Bricks [11], tangible objects can be used as instruments in the sessions. These objects are recognized and given their own unique sound which stems from their shape, position, etc. They would be given representations according to their sound. Users have the ability to join and create rooms, both as audience and as musicians. Depending on the style of music, room lighting changes which is another layer of feedback for both the users and audience. Movement of users in real life is mapped to their movement around the room, and the audio changes in real-time [9] [10].

Fig. 15. Concept 3

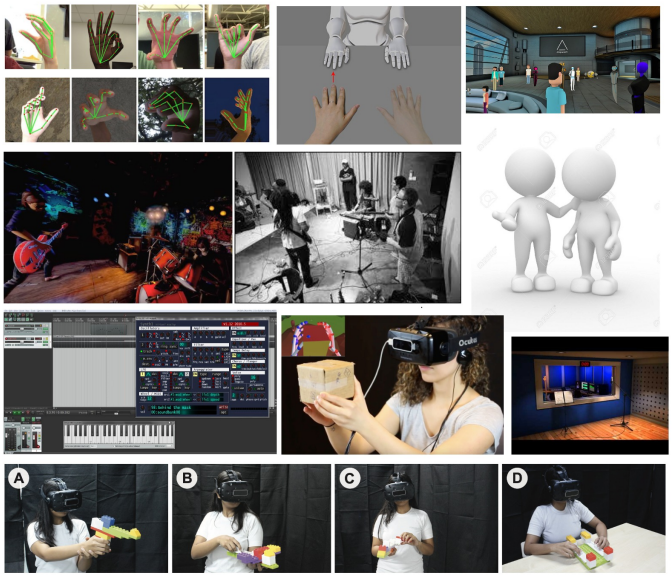

Fig. 16. Concept 3 Moodboard

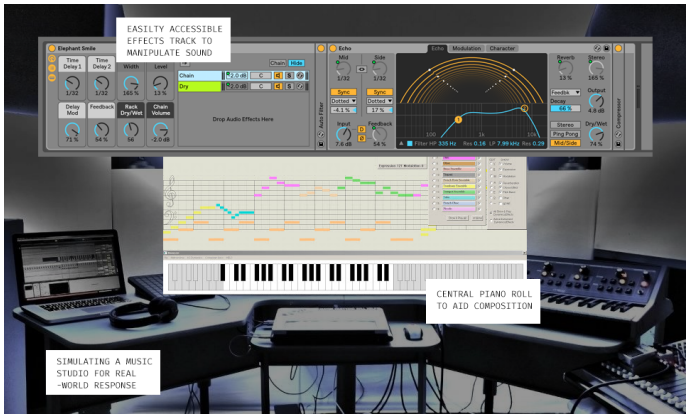

Fig. 17. Concept 3 Visualization

10.3.4 ***Concept 4: Jajantram Mamantram.*** This solution is based on tracking AR tags (stickers) for music production. A user can draw out specific pre-defined shapes on a surface in any sequence. With the camera overhead, they are scanned. Any gesture made on/interacting with the stickers can change parameters of sounds and music can be generated. Multiple users can create bands or groups, public or private jamming sessions. While jamming, users can also see others' tags, which are highlighted on cue. Music produced in jamming sessions may be posted on a social forum of the app [9] [10].

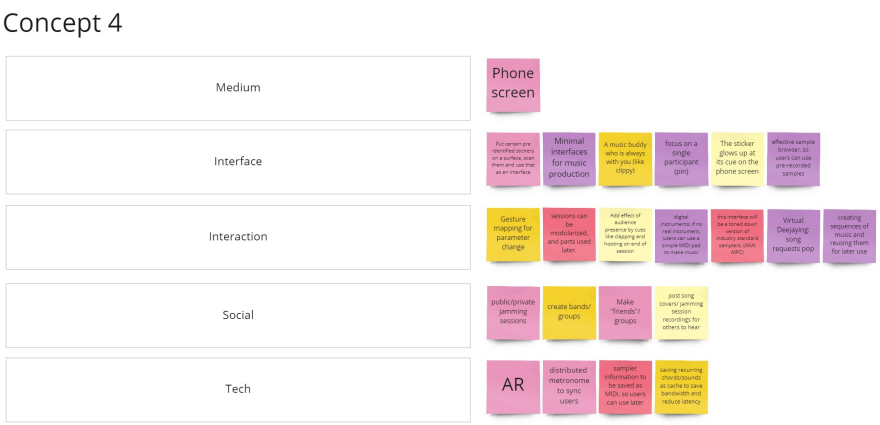

Fig. 18. Concept 4

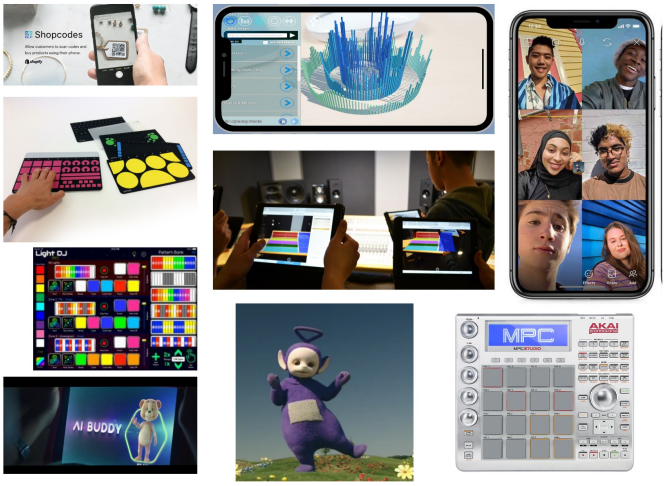

Fig. 19. Concept 4 Moodboard

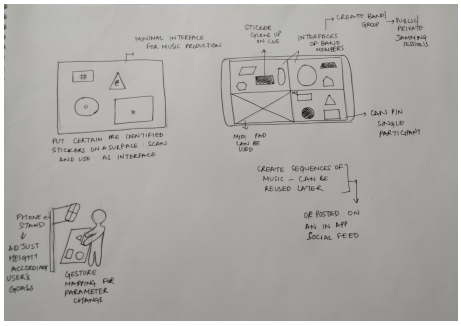

Fig. 20. Concept 4 Visualization

10.3.5 ***Concept 5: Tinky Winky.*** This concept is focused on gamifying a particular aspect of collaboration: accuracy. Based on what the users play, and how accurately they play, they can get live feedback on their style. A common music companion/avatar provides users with real-time feedback and facilitates enhanced collaboration. Users can watch other people's progress, as well as their own while reaching their personal goals [9] [10].

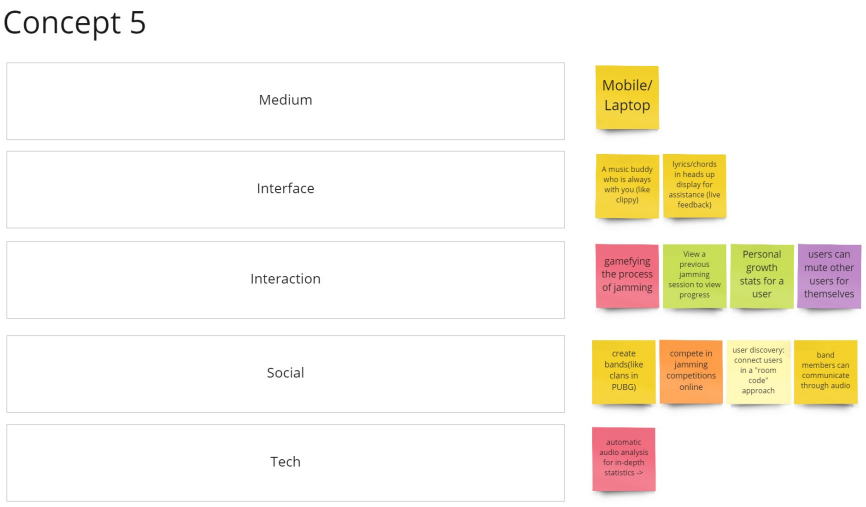

Fig. 21. Concept 5

Fig. 22. Concept 5 Moodboard

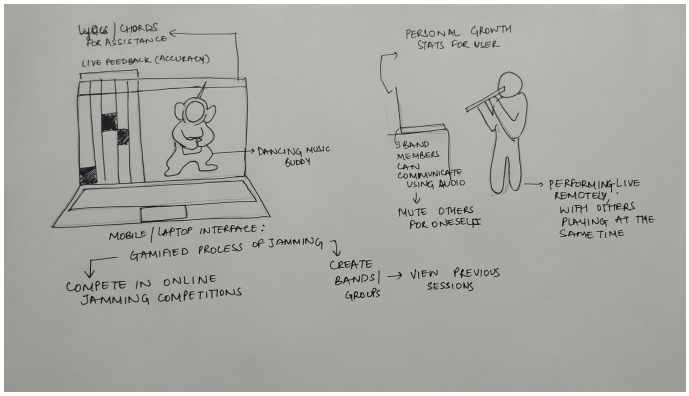

Fig. 23. Concept 5 Visualization

10.3.6 ***Concept 6: VirtualStudio.*** The sixth concept will emulate real-world arenas for concerts, mapping sonic properties and spatial features. In this idea, musicians will be able to perform around the globe-remotely. The musicians will appear on a common background with the backgrounds from their videos cropped. The audience will appear as avatars. The following concept aims to emulate real-world arenas, factors such as sonic properties and spatial features [9] [10].

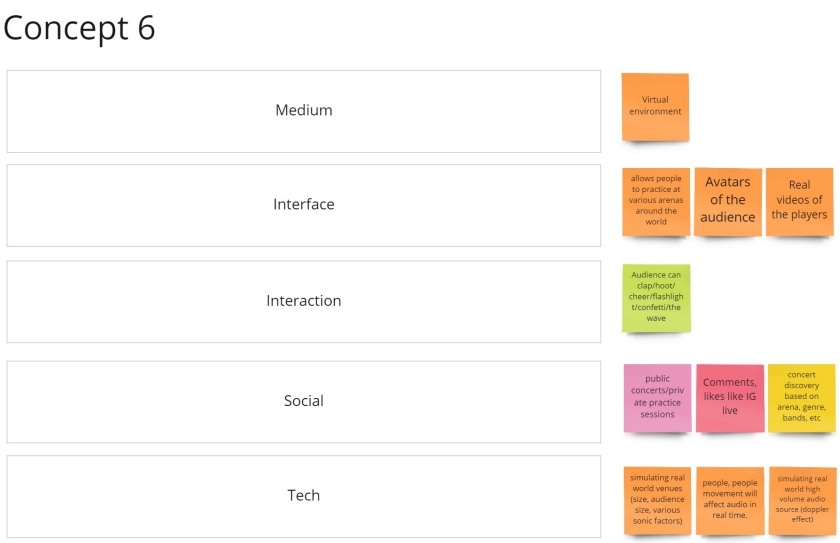

Fig. 24. Concept 6

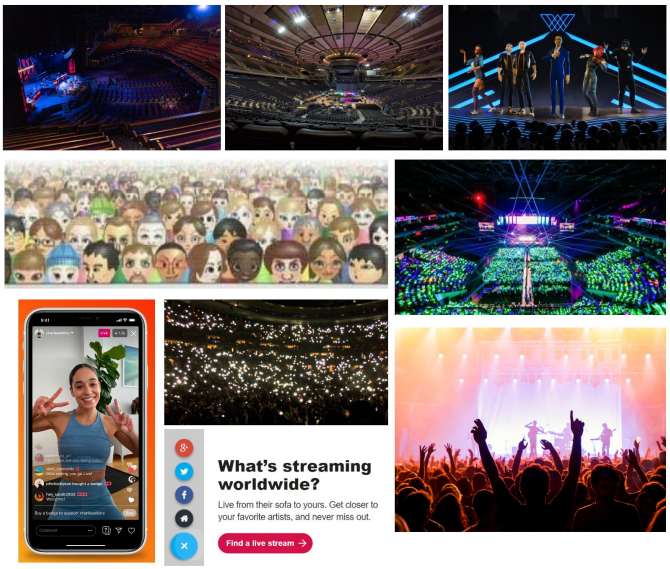

Fig. 25. Concept 6 Moodboard

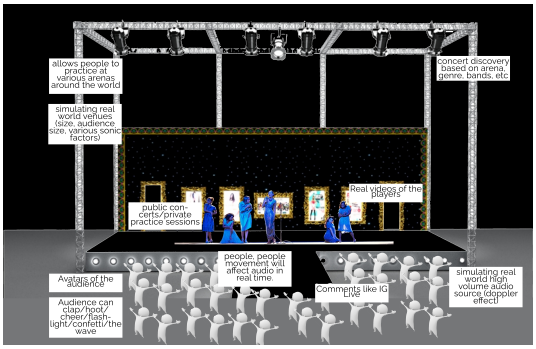

Fig. 26. Concept 6 Visualization

10.3.7 ***Concept 7: BuDjBa.*** This idea aims at making electronic music collaboratively. Current solutions cater to utilizing an instrument, the following solution emulates standard electronic music creation technology. Users will have a simplified version of the standard piano roll to compose music. Users have the ability to load their own sounds or utilize pre-defined sample kits (drums, percussion, leads, etc). A central timeline/audio editor aids in group compositions, users can see changes in real time and compensate accordingly. Users can either draw in notes, or utlize gestures to record sounds, which are directly translated to MIDI. Users can request to join such public rooms, and get straight to composing using the simplified audio editor. [9] [10].

Fig. 27. Concept 7

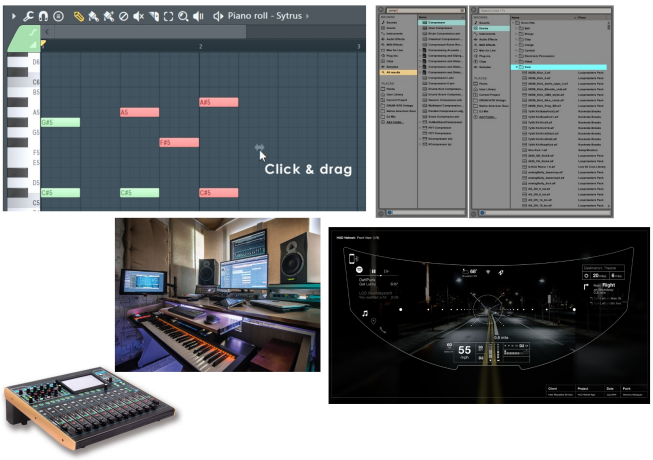

Fig. 28. Concept 7 Moodboard

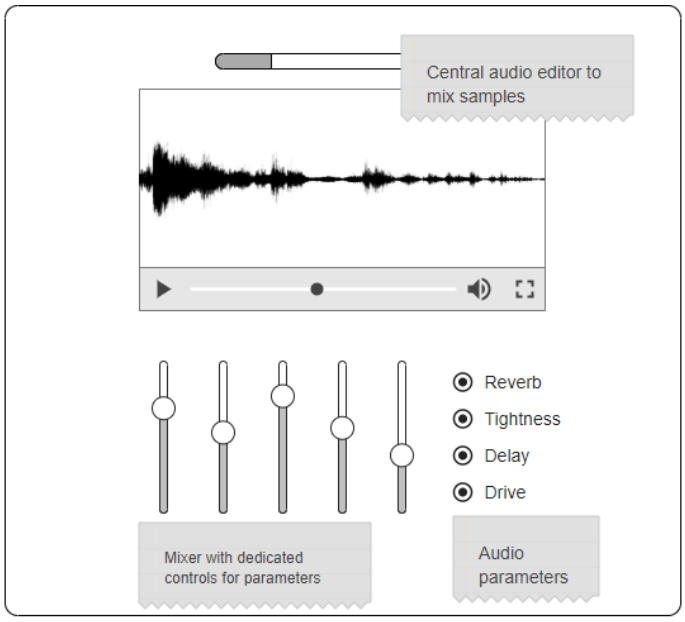

Fig. 29. Concept 7 Visualization

10.3.8 ***Concept 8: Co-opera***. A concert hall with different sections for different instruments in which different families of instruments can be controlled by different participants and a single audio stream will be heard by all the participants. Gestures/buttons will be mapped to different audio properties, which are activated based on the cues given from the conductor. Users can use pre-defined arrangements or have the choice to create their own in real-time based

on the former cues. A single user controls a single quadrant and can change audio properties such as (scale, key, pitch, reverb, etc). There's also the role of a conductor who gives cues to every quadrant, maintaining structure. Users will have the ability to join and create rooms, both as audience and as musicians. The audience views the session in the third person, and they're shown as personas, so as to not create information overload for the musicians. The conductor also gets real-time feedback from each quadrant on the basis of audio properties (tonality, effects, etc) [9] [10].

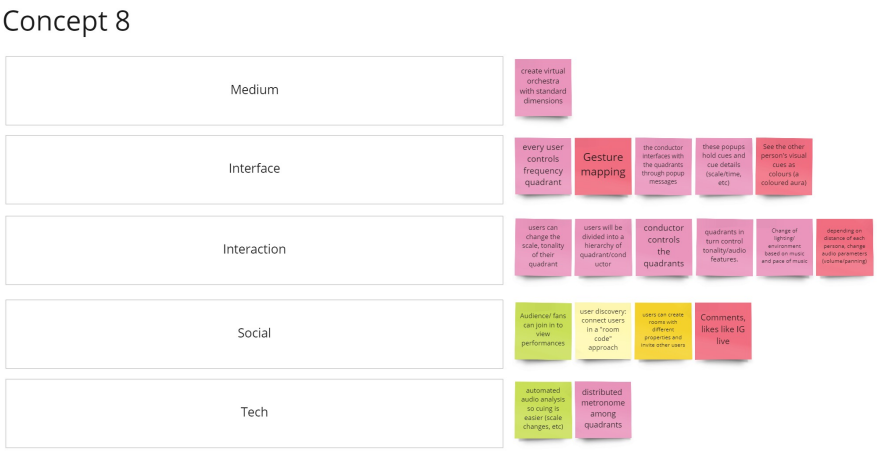

Fig. 30. Concept 8

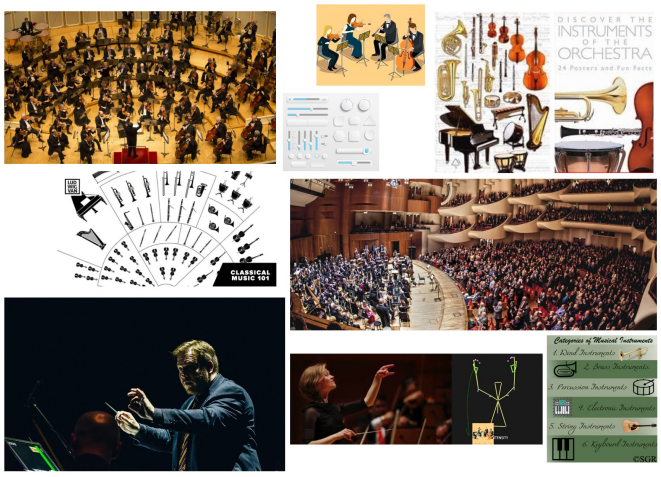

Fig. 31. Concept 8 Moodboard

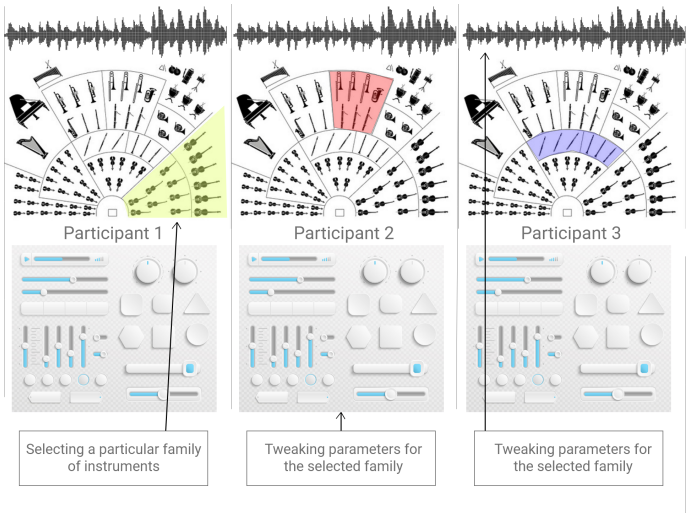

Fig. 32. Concept 8 Visualization

10.3.9 ***Concept 9: StreetMusic.*** Crowdsourced mapping of music/beats to GPS coordinates for creating spatial audio tracks. This will enable users (especially) tourists to experience the culture of a place through music. The medium to visualise these beats would be through AR-enabled mobile app. The users can privately add their own music/beats to coordinates or contribute to publicly available spatial audio tracks. The beats can either be recorded or be created through the app. The app would also feature a global map (similar to google street view) that will enable people to experience the music of a place virtually. The spatial audio tracks can be customised to place different beats in every foot of the jogging track (which would create a different music tempo for different running speeds) or a single track of music be localised to a geographic location like a Snapchat filter [9] [10].

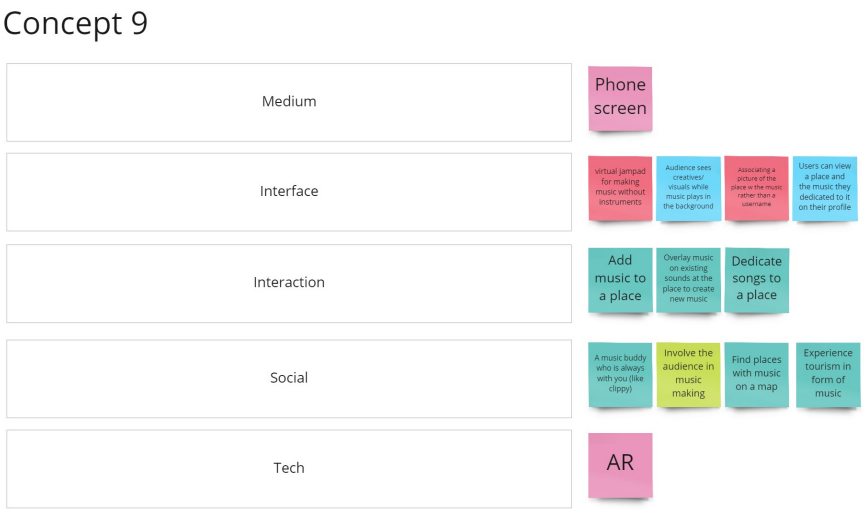

Fig. 33. Concept 9

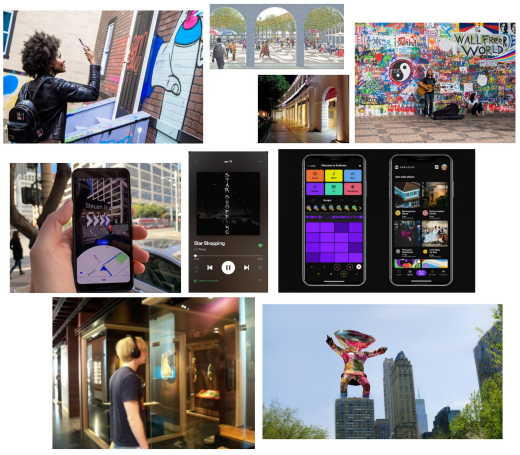

Fig. 34. Concept 9 Moodboard

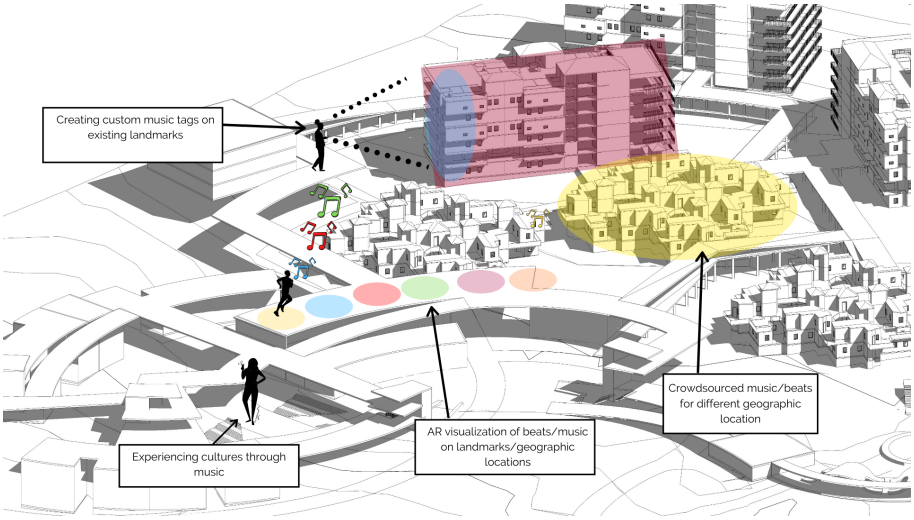

Fig. 35. Concept 9 Visualization

11 LOW FIDELITY PROTOTYPING AND PRELIMINARY USER TESTING

We went over the above stated nine concepts and eliminated a few ideas. Further, we combined ideas from different concepts to converge to the following three solution concepts. We asked ourselves the following questions:

- Which ideas are feasible/ practical?
- In how many different ways are we approaching collaboration?
- Which ideas satisfy the user requirements identified?
- Which ideas give the user a sense of freedom?
- Is there a novelty factor in these ideas?
- Which ideas seem approachable and acceptable?
- Which ideas are we passionate about?

We used a convergence technique called the Dot Voting Method. Each of us considered the above stated 7 questions individually and were given three votes each. All of us could either vote for an existing concept or merge concepts and cast our vote for the new mix. Further, the repetitive ideas made by merging were removed and votes were compared. The ideas with maximum votes were:

- **Concept 2+3:** Enabling synchronous collaboration.
- **Concept 4+6+8:** Enabling synchronous remote performances.
- **Concept 9:** Enabling asynchronous collaboration among different users in different locations around the world.

We visualized these ideas in the following mind map. Figure 37 explains the different approaches to collaboration adopted in this step of the project.
To decrease the scope for vagueness, out of the potential areas explored, we further narrowed down the scope to the following three ideas and conducted preliminary testing with users. We made our users interact with the low fidelity prototypes and asked them to think out loud. We plan to conduct such user testing at every stage of prototyping in

order to ensure that the solution we're working on fits user requirements. This also helped us gain feedback about our concepts for further work to increase their scope.

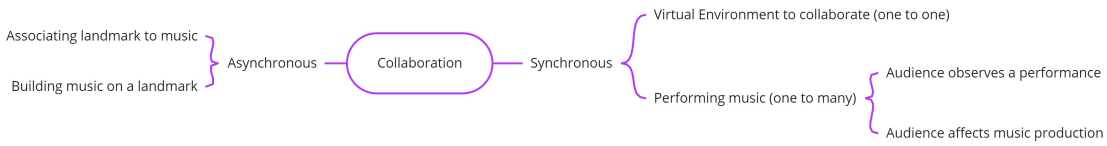

Fig. 36. Scope of solutions

To decrease the scope for vagueness, out of the potential areas explored, we further narrowed down to the following three ideas and conducted preliminary testing with users. We made our users interact with the low fidelity prototypes and asked them to think out loud. We plan to conduct such a user testing at every stage of prototyping in order to ensure that the solution we're working on fits user requirements. This also helped us gain feedback about our concepts for further work to increase their scope.

11.1 Idea 1: E-studio

Type: Synchronous Collaboration
Stakeholders: Musicians
Interaction: One-to-One

Description
Musicians enter a virtual environment to collaborate. User personas represent real-life users. Musicians can select the instruments they want to play. These instruments are visible to other users as physical instruments played by the user personas. For musicians playing instruments, every instrument has a simplified interface with controls for effects, notes, and arrangements to play the instrument. To improve the process of in-person collaboration, floating user statistics are displayed beside each user. This gives other musicians detailed information about what the user is playing. A cue window also helps musicians realize when they have to play. It displays messages from other musicians, which would otherwise not be heard over the sounds of instruments. A central mixer arrangement allows universal controls like audio level controls and provides feedback based on audio.

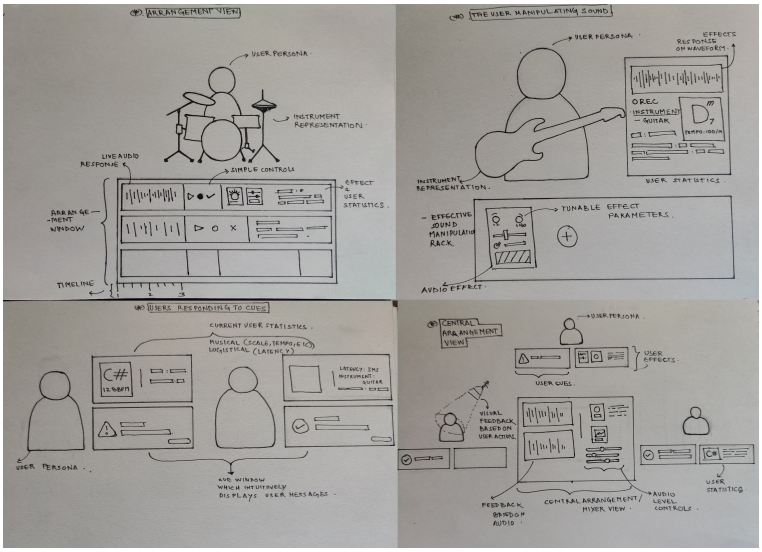

Fig. 37. Idea 1: Paper-Pen prototype

Prototype

To get an overview of the different components in the concept, it was first visualized in two dimensions as a collection of user-action sketches. These sketches gave an overview of the information the users utilize while creating music, and the actions they are capable of doing. Since the crux of the concept is sound spatialization, distance, and the influence of room properties, the two dimensional prototype did not offer a complete interpretation. The concept was programmed as a 3D scene in Unity to help the user experience the interface and possible interactions.

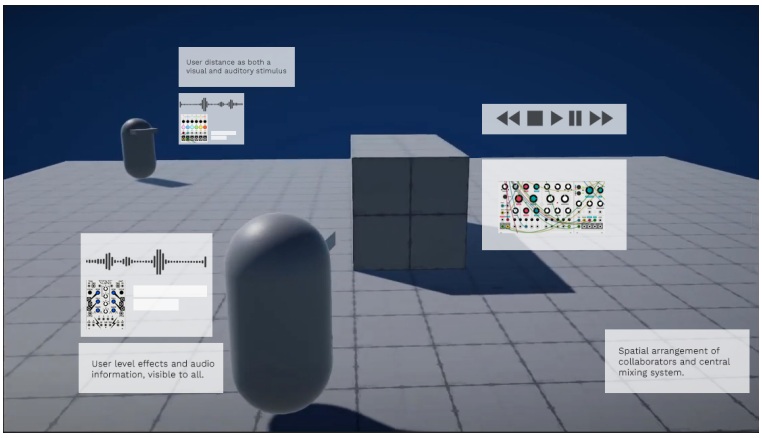

Fig. 38. Idea 1: 3D prototype-1

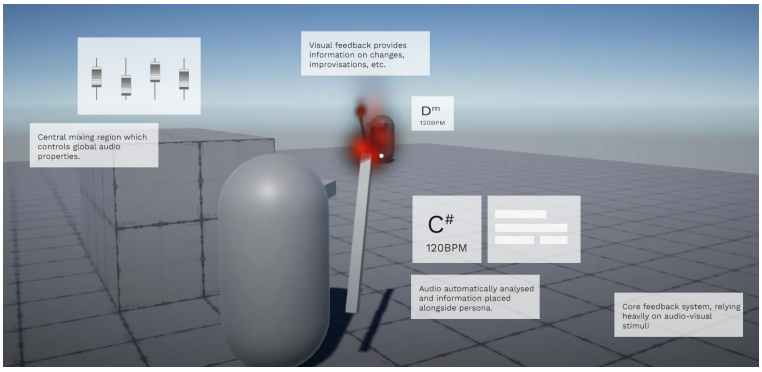

Fig. 39. Idea 1: 3D prototype-2

User Testing

The 2-Dimensional sketch was first shown to two users, and they were asked to imagine how a virtual "room" would affect their collaboration when they cannot communicate through eye-contact or real-world gestures.
Their thoughts were as follows:

- **User 1:** "How will I play well if I can't see my instrument", and "I can't imagine the following setup through a 2D representation".
- **User 2:** "Wouldn't I get disoriented because of the dearth of stimuli in this environment?"

Following this, the apt response would be to convert the sketches into a low-stimuli virtual environment so users can understand how sound and gestures would work.
To the simulated environment, they had the following response(s).

- **User 1:** "I am hesitant about this solution because I'll have to get accustomed to another interface altogether. I won't be able to play my guitar to directly jam with my friends. I also feel people who can't play instruments without looking will face challenges. The personas aren't very life-like, and it feels like they remove the "natural" element from the music."
- **User 2:** "In an ideal scenario, this will be a perfect solution, right? I can do what I did in real life but from home."

11.2 Idea 2: Performax

Type: Synchronous Collaboration
Stakeholders: Audience and performers
Interaction: One-to-many

Description

Musicians can enter the virtual environment and choose the location they want to hold the concert. A concert in the selected environment replicates one which would be experienced in person. Musicians can choose their instruments, which will correspond to a real-life object. Interactions with the object would be viewed as interactions with the instrument in the virtual world.
Viewers enter the virtual world and can choose the concert they want to attend depending on the genre of music or the

environment they want to experience. They can customize their appearance, which reflects how other viewers see them. Musicians are also viewed as their custom personas. The audience contributes to music production in concerts. Real-life objects are assigned interactions. By simply tapping them, the audience engages in activities like clapping and stomping. Such actions can also be viewed in the virtual world. Additionally, viewers can also react to the concert, which would result in the reaction symbol hovering beside their persona.

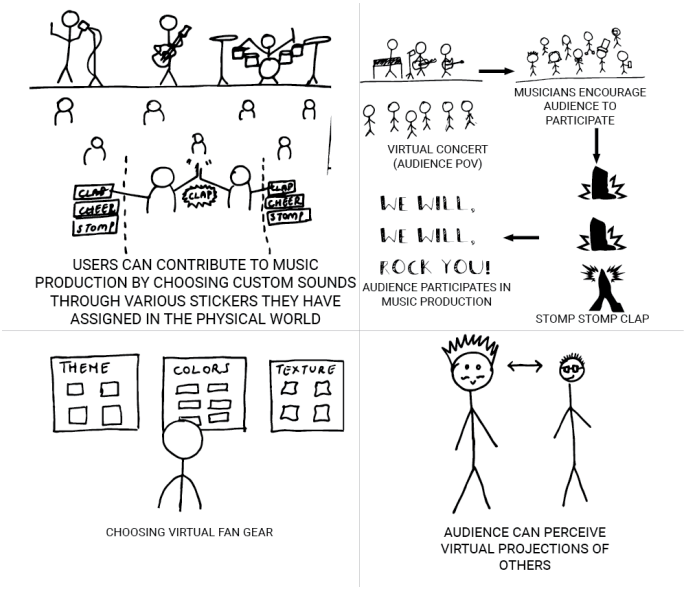

Fig. 40. Idea 2: Paper-Pen prototype

Prototype

The concept was visualized through sketches. These helped imagine what the implementation might look like in a virtual environment. This also helped plan the prototype for the idea. A 3-D low fidelity prototype using analog tools was made based on the visualizations. This was an appropriate choice for the concept as a 2-D prototype does not give the user a feel of the immersive experience they are likely to have. A 3-D prototype, on the other hand, replicates the user interactions and experiences and does a better job of communicating the solution to the user. Bottles were used to portray the personas of real users in the virtual environment. To depict customizations, each persona was given its own appearance by the use of accessories. Through interactions with these analog tools, users can visualize the concept and various features of the solution and gain a better understanding of the same.

Fig. 41. Idea 2: 3D prototype

User Testing

Two users were explained the idea behind the low fidelity prototype and introduced to the setup. Their first thoughts regarding the concept were:

- **User 1:** "If I am a musician, I will not have to worry about a lot of things that hinder my concerts/ performances. I will not have to worry about traveling, security, extra expenses of promotions. I will simply reach out to a wider audience and not miss out on fans in places I can't cover in tours. At the same time, I will not be able to play my current instrument; I will have to master the usage of a completely different, new interface in order to make no mistakes in a performance. This isn't exactly what I'd want to do."
- **User 2:** "I would love to view a virtual concert that is more than simply a screen, and I can imagine myself in a location. I can attend performances from my home; it'll be like Wii games except with people and no automation. The only negative I can think of is missing the feel of proximity. Though I will be able to see others, I won't be able to feel them. That is an important aspect of attending performances for me - feeling like I am a part of a bigger crowd."

A video of the 3-D prototype was also made as a reference to create more advanced prototypes and use as a testing tool for users who would not be able to interact with the prototype physically. [Video Link here]

11.3 Idea 3: BeatMap

Type: Asynchronous Collaboration
Stakeholders: Performers
Interaction: Many-to-many

Description

This concept allows users to associate music/sounds with landmarks. It helps users capture their moods/emotions at a particular landmark. This collection of beats results in the creation of a "BeatMap," which users can post publicly or among selected people. Users can also explore other publicly available BeatMaps and even send a request to modify it by adding their own music pieces. The users can also generate beats through an intuitive drag-and-drop interface

called the “BeatBuilder.”

The users can access the BeatMaps by either scanning the landmark in-person or by using an explore world-map. They will only be able to add/edit any BeatMap during an in-person visit, while they may explore the BeatMaps from anywhere in the world. The landmark shows up in the form of a 3D model through which users can tap on and access specific BeatMaps.

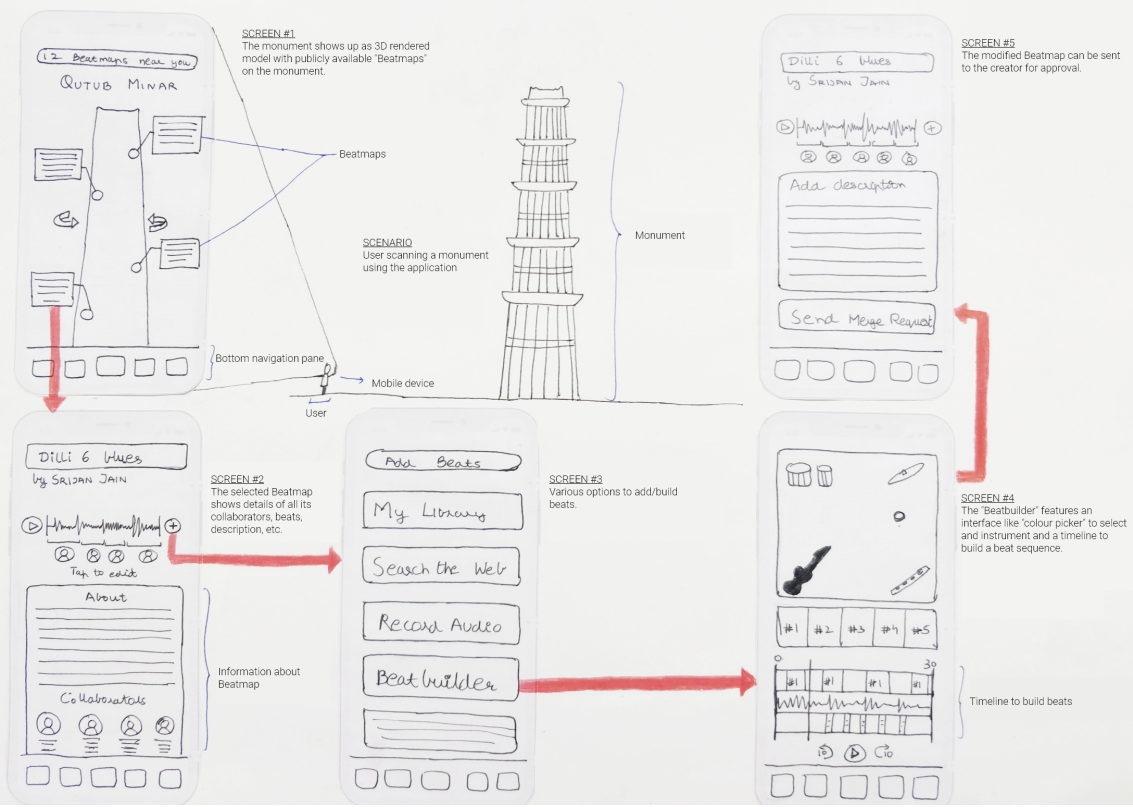

Fig. 42. Idea 3: Low Fidelity Sketches

Prototype

The concept was visualized through low-fidelity sketches and low-fidelity wireframes. It features a user scanning the landmark to view the associated BeatMaps. The user then goes on to modify an existing Beatmap using the BeatBuilder and then puts a merge request to the creator of the BeatMap.

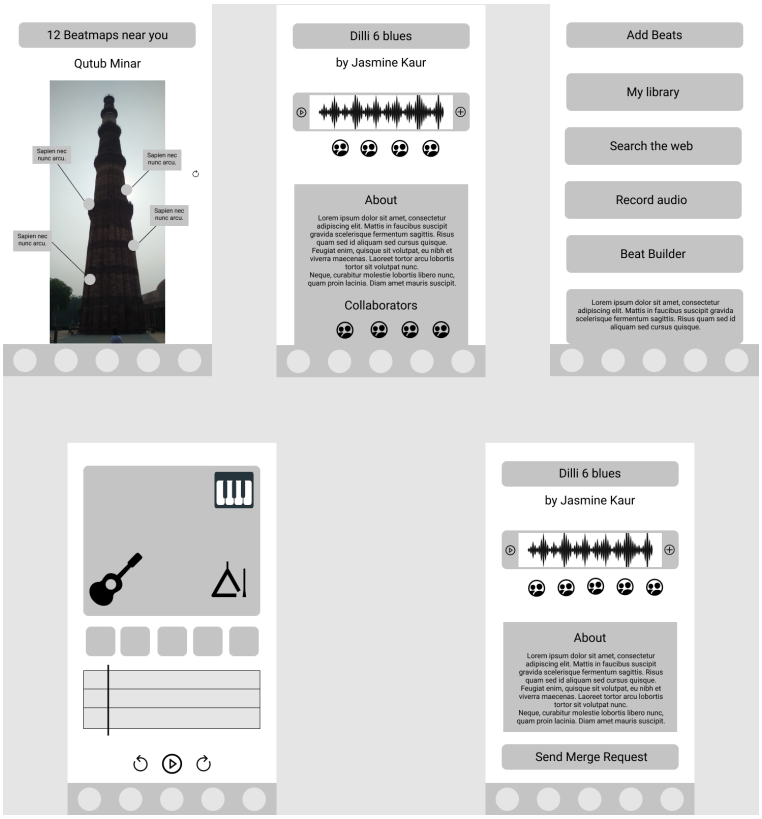

Fig. 43. Idea 3: Low Fidelity prototype

User Testing

- **User 1:** "If I am roaming about, sightseeing, I might not always have the time to make something new, especially for a particular location; I would rather assign something that I already have with me. It would also be more convenient to take a picture of the place and assign it a sound later when I have time."
- **User 2:** "I can reach out to a wide audience, more people will know about my work."
- **User 3:** "If people are manipulating my music, plagiarism is possible; I hope the app provides features to keep that in check."

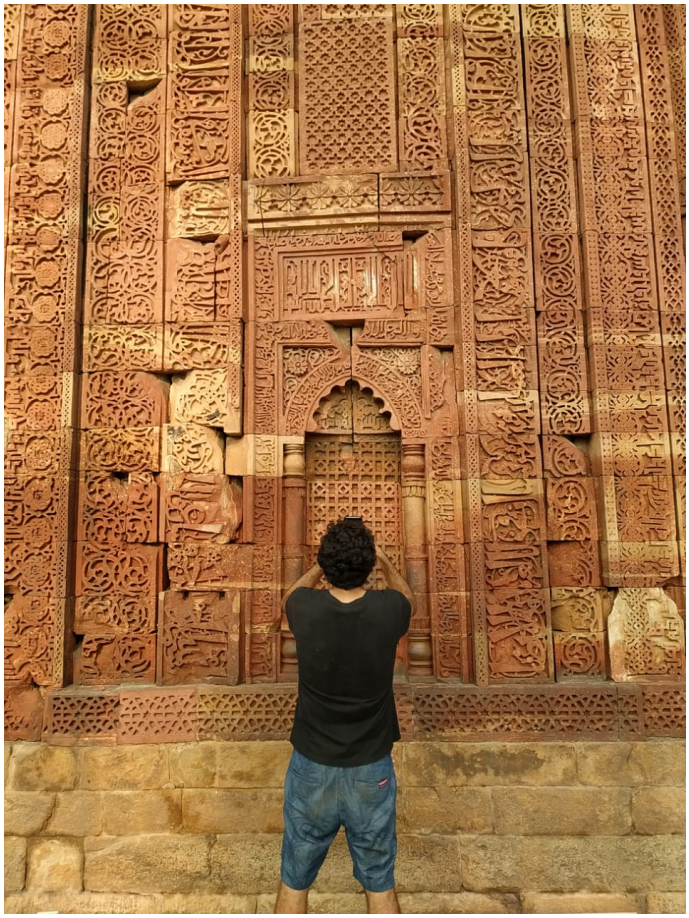

Fig. 44. User scanning monument to add music

12 ANALYSIS

Based on the individual feedback that we got for the three ideas, we realized that users were enthusiastic to test out novel methods for music collaboration. However, they were hesitant to change their medium of music production to enable collaboration. Tangible instruments play a central role for music performers and eliminating them significantly hampers the user experience.
Drawing a comparison among the ideas: the first two required the user to get familiarized with new interfaces to produce music. The third concept introduced a new way of collaborating that came with its own challenges. However, it is inclusive of people with no musical experiences and is driven by user emotions and their personal musical choices rather than theory.

Another observation made through the reviews was that users value the process of collaboration as much as the output. This leads us to conclude that any concept that we create will not be suited for our entire target audience, and

its adoption will be affected by the Diffusion of Innovation Theory and the relative experience of the user in music production.

13 REFLECTION

The last two stages of this project have added depth to our initial perception of collaboration. Initially, we were focusing on mimicking in-person interactions on a digital platform. However, through the processes of ideation, prototyping, and user-testing, we realized that collaboration could be extended to include non-conventional methods of collaboration. Since these are novel processes, they do not involve changing user behavior. Based on the feedback, we formulated our vision statement:

Expanding the boundaries of music collaboration for people who would be interested in collaborating in music in new ways.

14 VISION STATEMENT

We envision to create a shared experience of discovering, and creating music by mediating interaction among music enthusiasts.

15 IMPLEMENTATION

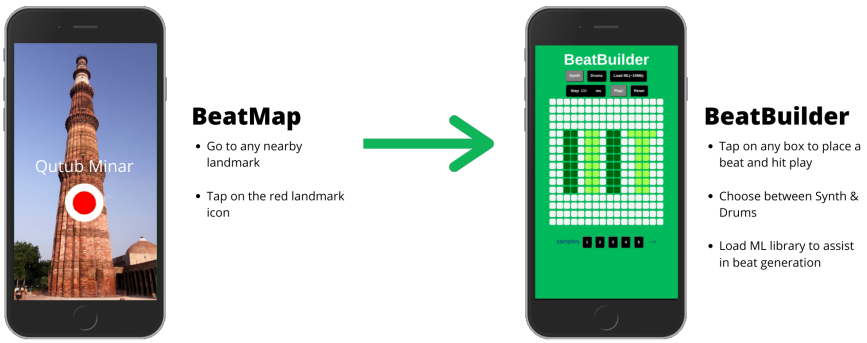

Fig. 45. Implementation: The Concept

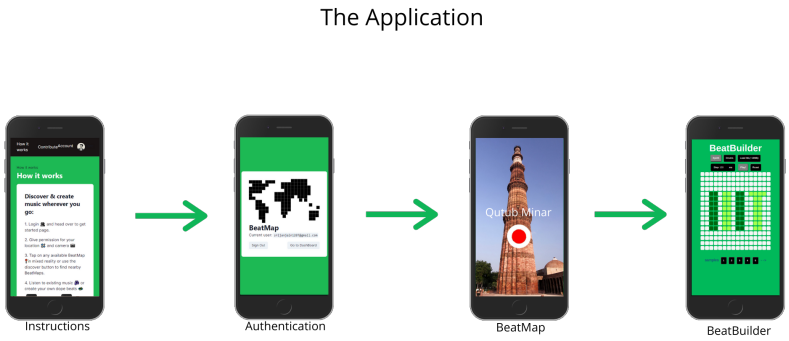

Fig. 46. Implementation: The Application

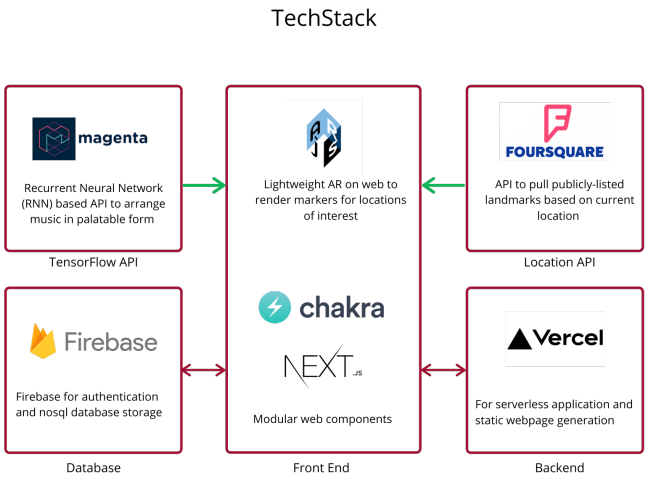

Fig. 47. Implementation: The Techstack

16 USABILITY TESTING

The method of evaluation adopted was contextual inquiry using video call. The users were asked to perform the tasks in the context of use. They were asked to use the application as they would if they were exploring different locations [12]. Keeping in mind the pandemic, the participants were asked to treat places around their respective homes as landmarks, associate music with them and edit existing music provided in the prototype.

16.1 Evaluation Criteria

The following themes and prompts were a part of our evaluation criteria:

16.1.1 First impressions:

- What does the user think of all components (AR, Music, etc)?
- Can the user understand the information presented?
- Is the loading time of virtual objects in the scene satisfactory?

16.1.2 Navigation Labels and Information:

- General feedback
- Do the labels and text make sense to the user?
- Were the contents of the page what the users expect?

16.1.3 Visuals:

- What does the user think about the colour scheme?
- Does the user think that the order of the information is natural and logical?
- What stands out visually in the task flow?

16.1.4 Task Sequence to test usability. The users will be asked to perform 2 tasks and their ease of performing the task will be recorded:

- Generating and adding music to any location in AR
- Editing existing music

16.1.5 Conclusion (Are our goals being accomplished?):

- Is the user able to discover new music?
- Is the user able to create new music?
- What are their views about the concept? (Is the product desirable?)

16.2 Procedure of evaluation

The participants were asked to share their screen and were given the link to the prototype. Then they were asked to think out loud what they thought about the application, it's features and the task sequences they were given to perform. The duration of each evaluation was 15-20 minutes.

16.3 Tasks to be performed by a participant

The participants were asked to perform the following two tasks:

- Generating and adding music to any location in AR: The user was expected to find location tags in Augmented Reality and click on them. On reaching the Beat Map screen, they were expected to explore generating music, while experimenting with different features offered by the application.
- Editing existing music: The user was expected to open any existing music and experiment with editing it, thereby asynchronously collaborating.

16.4 Data policy

The data collected was of qualitative type. The study was performed for six participants.

16.5 Setup

The studies were conducted remotely using the prototype link hosted on the web. The participants of the evaluation were asked to share their screen and perform the tasks.

17 EXECUTION

The above mentioned evaluation plan was followed with six randomly sampled participants. To maintain user anonymity, they were assigned pseudonyms from User-01 to User-06.

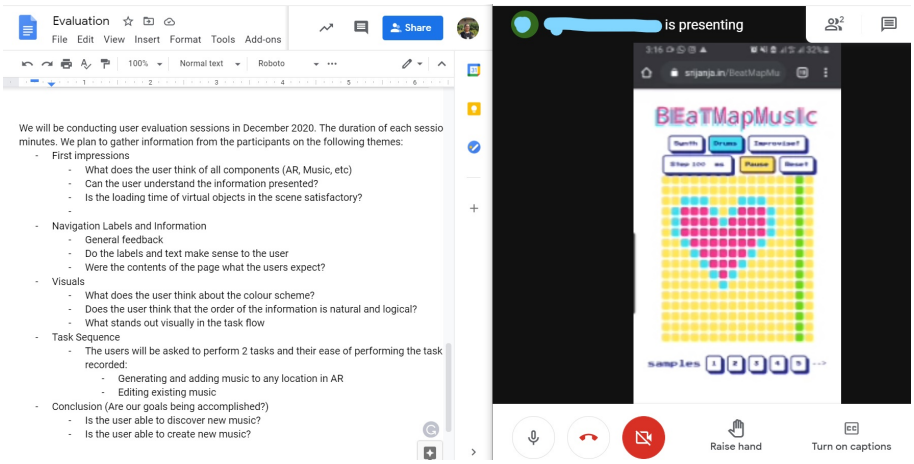

Fig. 48. User Evaluation Execution (1)

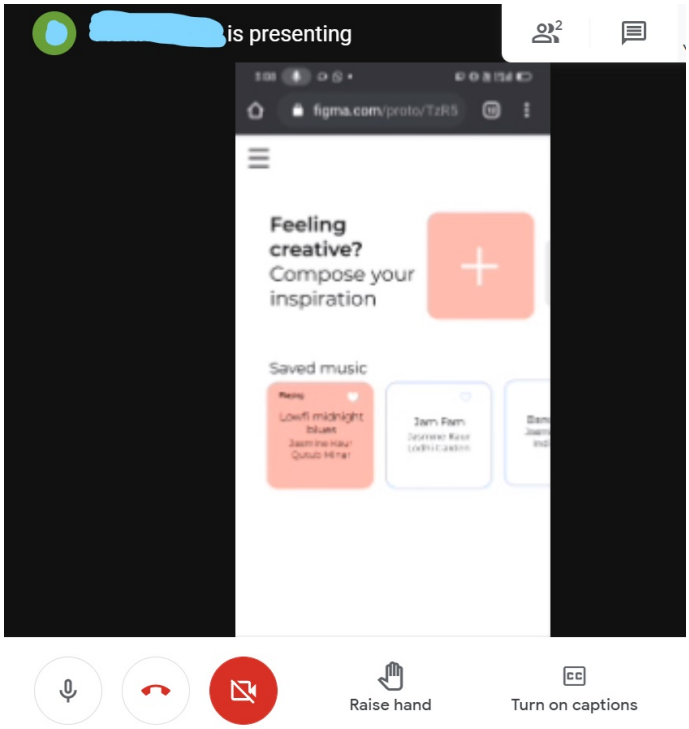

Fig. 49. User Evaluation Execution (2)

18 COLLECTED DATA

First Impressions

User-01	User-02	User-03	User-04	User-05	User-06
'Where are we going today' on the landing screen wasn't clear to the user. They want to see a different Call to Action.	The user initially did not understand how to assign music to the red markers.	They found the app to be a bit tricky, didn't know what to expect out of 'nearby', and how to edit music or scan.	The user expected the prototype to actually play music on hitting the play button rather than just have a dummy button.	They found the 'cool factor' of the app on the Beat Map page. They felt that the rest of the app interface did not create the kind of hype that it should have.	The user found the AR mode to be sudden and without warning- they expected a prompt to come up first.
The user did not find the red dots in AR intriguing. They expected music symbols.	The user feels that they don't know what to do next after they have generated music.	Found the information presentation to be fine.	They found the opening of different tabs for exploring locations confusing.	The user was astonished and found the entire concept lovely.	They were excited about the idea.
The user found the loading time for the red dots in AR satisfactory.	The user feels that they don't know what to do next after they have generated music.	They did not find the red dots to be intuitive, they expected a tutorial for AR.	They found AR intuitive and easy to understand.	They found the interface of the Beat Map to be eye-catching and appealing.	
They found the app to be aesthetically pleasing, plain and simple.	They were unsure of the improvise feature.	The user found the loading time in AR satisfactory.	The user wanted more instruments and possible chords labelled.	They found the loading time in AR to be quick.	They found the app to be intuitive.
	The user felt that AR markers were cluttered, had unnecessary variations in size and were randomly positioned.	They found the information present intuitive.			
	The user can easily understand information present				

Table 3. First Impressions

Navigation Labels and Information

User-01	User-02	User-03	User-04	User-05	User-06
The user thought that the labels made sense.	The user said that the navigation labels did not do justice to the social aspect of the application. They expected keywords like trending, popular music etc.	The user found the labels to make sense once they tried out different buttons and screens. However, they were skeptical about them before.	The user wanted a more detailed description of the labels on the Beat Map screen- what each dot on the numpad meant. They felt that it would make it easier to edit music and create new music without having to play it continuously to test.	The user found the application to be intuitive and easy to use and navigate through.	The user found the existing labels to be good but they expected more labels and prompts. They wanted pop ups indicating what was going to happen - at least for the first time usage of the application.
They said that pages were what they had expected.	The user suggested the keyword explore with nearby. Other than that, the labels were self-explanatory.	The user wanted some better or different indicator labels for the application.	They found the in app labels to be intuitive.		
	They said that nearby felt like a satellite map with music at places.	They could navigate well The user felt the labels were misleading. They felt they could edit recently played music.			

Table 4. Navigation Labels and Information

Visuals

User-01	User-02	User-03	User-04	User-05	User-06
The user said that the app looked good but simple. The expected recently played like other music apps rather than a list.	The user did not expect to see recently played on the home screen. The rest of the visual information seemed fine.	The user felt that the app did not do justice visually to the name Beat Map.They expected a better colour scheme for the interface.	The user's first reaction on the Beat Map screen was that it looked like a toy for kids-it was fun and colourful.	The user did not find the interface to be visually engaging. They said that it did not give away the vibe of a music app like spotify.	The user found the app to be visually pleasing but expected a more colourful interface that might "hype me up". They want the colours of a music app to "pop in one's face".
They added that the visual flow of information is natural and logical. Nothing stood out visually in the task flow for them.	The user said that the colour scheme was fine.	After navigating to the Beat Map screen, they felt that the interface was visually a let down after the application.	They found the visual hierarchy to have a good flow.	They appreciated the visuals of the Beat Map screen.	
	The user said that the colour scheme was fine.	They caught a prototype error: The screens were scrollable due to the difference in the screen size of their phone and the application.	They liked the visuals and appreciated the subtleness of the colour scheme.		
	The users did not like having samples on the add music screen of Beat Map.	They felt that the information was logically presented.			
	They found the UI to be boring.				

Table 5. Visuals

Task Sequence and Features to test Usability

User-01	User-02	User-03	User-04	User-05	User-06
The user was used to location automatically turning on after a prompt, so they were confused about why the AR markers did not show up.	The user faced considerable issues while trying to add music. They did not find beatmap easy to use while adding music.	The user did not understand how to add music to a place at first and clicked on nearby to add. They later realized they could click on scan.	The user enjoyed navigating through the application to perform the two tasks. They compared sharing music to instant messages on iMessage.	The user found the task sequence flows to be smooth, "anybody can do this!"	The user felt that the interface was subtle and it partially acted as a hindrance in motivation to navigate through the app to perform the task sequence.
The user found a prototype error: the existing music "gulu gulu" did not work, the page was blocked.	They had to tinker around to learn how to make music, but still found it uncomfortable to use, due to lack of labels.	The user did not know what to do once Beat Map was open and felt that play button should've been more visually prominent.	They expected an in app tutorial, especially for a first time user since the concept was unique.	The user added, "the cool factor of the app was visible the minute I landed on the music generation page".	The user was able to perform the task sequence at a slow pace.
		They did not understand Beatmap even after experimenting and felt the need of instructions. 'Load ML' was unclear.	About improvise they said, "it'll be amazing for enthusiasts without formal training".	They were quick to perform the task given.	They wanted more pop ups or prompts as indicators that they were performing the task in a correct fashion.
		To editing music did not feel intuitive and they needed options to make copies.	They found editing music to be easy.	They appreciated being able to access and work on other people's existing music.	The user was worried about their music getting lost-they wanted an explicit save button.
		Prototype error: the existing music "gulu gulu" did not work, the page was blocked.	They expected more instruments and features for music generation.	The user found the load ML-improvise feature to be excellent.	They wanted the application to urge them to try out the improvise feature.

Table 6. Task Sequence and Features to test Usability

Conclusion (Are our goals being accomplished?)

User-01	User-02	User-03	User-04	User-05	User-06
The user appreciated the application. They felt the implementation is good but it looks very simple.	The user found the idea to be interesting, however felt the implementation was weak, and did not let the idea make an impact.	The application was very confusing to use for the user and felt that there were a lot of improvements to be made in the implementation.	The user appreciated the concept behind the application. They believe it can get their creative juices flowing, even if they are not expressing their emotions but are just passing time waiting in some queue.	The user said, "this is one of those innovations that you can see everyone talk about". They believe that a product like this in the market could create a great hype for a bunch of different industries and can be the future of marketing.	The user said that the thought behind the prototype was excellent and worth working on to release in the market.

Table 7. Conclusion (Are our goals being accomplished?)

19 INSIGHTS

Overall, the users really liked the concept of geolocation based social music discovery and creation platform. On analysis, we concluded that many issues that the users pointed out in the application, like the different sizes and placement issues of the AR markers were a limitation of the prototype rather than that of the application. The major problems that we noted were:

- Users struggled with instructions and didn't know how to proceed
- The UI for the application, AR map and music builder were not coherent
- The application didn't feel connected, too many breaks
- The UI felt a bit dull
- Some users would prefer an application on their phone rather than a web app
- Users wanted more methods to product music

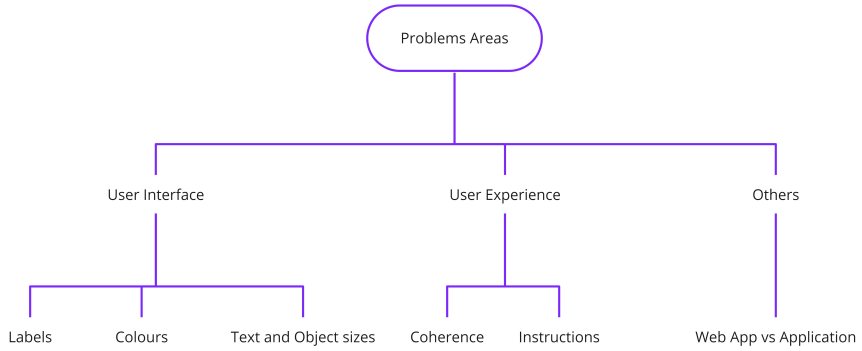

Fig. 50. Problems identified through user evaluation

20 DISCUSSION

User feedback is valuable in any user centred design process. Beat Map too was created as an innovation for users to asynchronously collaborate and make music, with users involved at multiple stages of the project. The prototype will be improved based on the feedback. The changes that we propose to do in future iterations:

- First time onboarding tutorials
- Coherent labels, colours and text.
- Improve UX by building a packaged application
- Improve UI by taking inspiration from social apps such as Instagram, Snapchat and music apps such as Spotify, Apple Music and Shazam.
- A mobile application based on the web app
- More options to create music

The future plan of this project, post the scope of the Design of Interactive Systems course is to conduct more user evaluations with IRB approval. We plan to take it up as an entrepreneurial venture and release the application on playstore and iOS in a few months. Post launch, we wish to start pilot testing on select landmarks of Delhi-NCR.

21 TIMELINE

Task	Start Date	End Date	Day
Project proposal	1 September 2020	September 14, 2020	0
Ideation	15 September 2020	6 October 2020	14
Prototyping	20 October 2020	2 November 2020	49
Implementation	3 November 2020	22 November 2020	63
Evaluation	23 November 2020	7 December 2020	83
Iteration and Updates	9 December 2020	15 December 2020	99

Table 8. Project Timeline

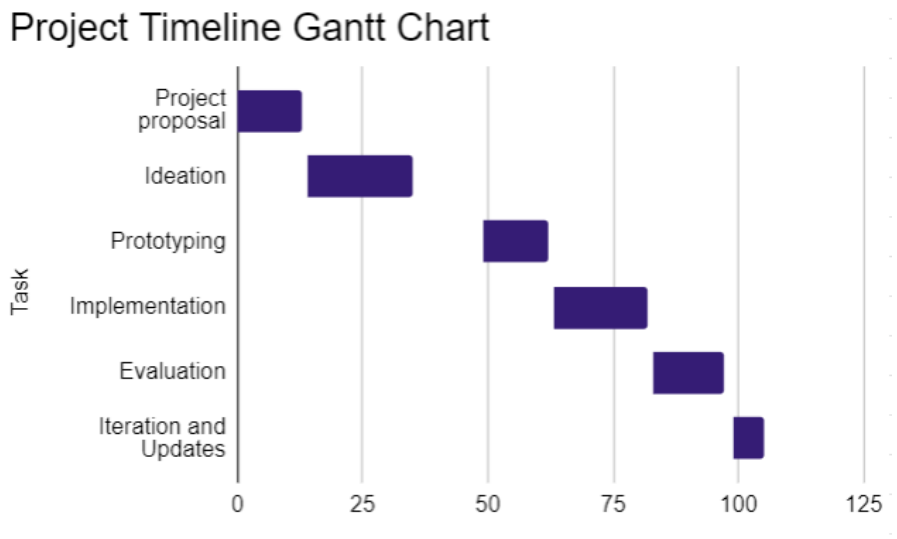

Fig. 51. Project Timeline: Gannt Chart

22 OTHER DELIVERABLES

- Web App
- Website
- Video
- Source Code(1)
- Source Code(2)

REFERENCES

[1] S. Brave, H. Ishii, and A. Dahley, "Tangible interfaces for remote collaboration and communication," 1998, pp. 169–178.
[2] H. Zou and H. Tang, "Remote Collaboration in a Complex Environment," Springer, 2020, pp. 808–818.
[3] N. Bryan-Kinns and P. G. Healey, "Daisyphone: Support for remote music collaboration," 2004, pp. 27–30.
[4] A. Carôt, T. Hohn, and C. Werner, *Netjack-Remote music collaboration with electronic sequencers on the Internet*. na, 2009.
[5] N. Nayar and D. Lohani, "Networked Remote Music Collaboration using Machine Learning," IEEE, 2020, pp. 708–713, ISBN: 1-72813-187-1.
[6] X. Xiao and H. Ishii, "MirrorFugue: Communicating hand gesture in remote piano collaboration," 2010, pp. 13–20.
[7] B. Loveridge, "Networked Music Performance in Virtual Reality: Current Perspectives," *Journal of Network Music and Arts*, vol. 2, no. 1, p. 2, 2020, ISSN: 2690-2044.
[8] A. Lakra, *Persona Scenarios*, Oct. 2020. [Online]. Available: https://docs.google.com/presentation/d/1lV3ysMNUHHa29I0iaq4m-mKRboYcyYzeRuUT7LCwWNE (visited on 10/06/2020).
[9] *Ideation, Online Whiteboard for Visual Collaboration*. [Online]. Available: https://miro.com/app/board/o9J_kjp4Y3M=/ (visited on 10/06/2020).
[10] *Visualization - Google Drive*. [Online]. Available: https://drive.google.com/drive/u/0/folders/1ISunpqL4h3YCVTvKlhIliYEmpdeQBg6R (visited on 10/06/2020).
[11] J. Arora, A. Saini, N. Mehra, V. Jain, S. Shrey, and A. Parnami, "VirtualBricks: Exploring a Scalable, Modular Toolkit for Enabling Physical Manipulation in VR," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19, New York, NY, USA: Association for Computing Machinery, May 2019, pp. 1–12, ISBN: 978-1-4503-5970-2. DOI: 10.1145/3290605.3300286. [Online]. Available: https://doi.org/10.1145/3290605.3300286 (visited on 10/06/2020).
[12] T. Singh, *Enterprise UX: Contextual inquiry in the discovery timeframe*, en, Jul. 2018. [Online]. Available: https://uxdesign.cc/enterprise-ux-contextual-inquiry-in-the-discovery-timeframe-713bc2714869 (visited on 12/07/2020).