Software Requirements Specification for Software Engineering: subtitle describing software

Team 8 – Rhythm Rangers

Ansel Chen
Muhammad Jawad
Mohamad-Hassan Bahsoun
Matthew Baleanu
Ahmed Al-Hayali

October 11, 2024

Contents

Revision History

Date	Version	Notes
2024-10-11	0.0	Revision 0

1 Purpose of the Project

1.1 User Business

Experimentation in music production is a process driven by intuition, i.e., lacking a core systematic structure, limited by a producer's experience and exposure to complex tools and techniques. *GenreGuru* strives to greatly reduce the effort involved in *methodically* attaining exposure to the use of tools and techniques in other songs. *GenreGuru*, by extension, democratizes access to experimentation in music production to less experienced producers, hobbyist musicians, and novices in music production.

1.2 Goals of the Project

GenreGuru shall:

- *featurize* – produce tabular features corresponding to characteristics of input songs;
- *recommend* – produce a collection of songs similar to input songs;
- *generate* – produce an audio artifact similar to input reference songs.

2 Stakeholders

2.1 Client

The project is academic in nature, hence has no formal clients beyond the supervisor, who will be consulted periodically to direct project effort.

2.2 Customer

Please refer to 2.3 and 2.4 for the current characterization of candidate customers. Section 2.5 will more succinctly specify archetypal customers after candidate customer interviews are carried out.

2.3 Other Stakeholders

- Subject Matter Experts (SMEs) – *To protect the privacy of our stakeholders, SMEs are completely deidentified with the exception of the user group they fall into, if applicable. Were this to be a commercial project, we acknowledge that we would have to, under jurisdiction of the Office of the Privacy Commissioner of Canada, abide by The Personal Information Protection and Electronic Documents Act (PIPEDA). Further, we have elected, for the interim, not to include a formal conflict resolution agreement as stakeholders' interests will be considered only under discretion of the development team.*
  - Music Producers & Sound Engineers (*subsequently "producers"*)
    - *Target subject matter knowledge:* description of current process and/or approach used to guide recording artists to explore or experiment with a new sound. For example, "*while recording, if I get an idea, I play a song with a specific cool feature (take tempo for example) and iteratively incorporate it into the current song being recorded, guiding the artist to adjust (e.g., incorporating a different cadence) across different attempts to incorporate the experimental feature.*"
    - *Extent of project involvement:* minimal, i.e., no more than three interviews per producer.
    - *Influence on project:* moderate-low – technology-keen producers may be more likely to already have a process into which *GenreGuru* can be integrated, i.e., song featurization can quickly, and in large volumes, summarize music that the producer's target audience listens to, allowing the producer to better tailor their music output. Such a producer's insights can inform and guide development, but at the discretion of the development team.
  - Musicians
    - *Target subject matter knowledge:* description of current process and/or approach used to generate novel ideas for unrecorded songs or experimenting with different ideas for already recorded songs. *We must be cautious so as to only consider the experimentation component of the musician's workflow, not the music creation in its core.*

    * *Extent of project involvement:* minimal, i.e., no more than three interviews per musician.
    * *Influence on project:* moderate – like technology-keen producers, we suspect musicians may already have a process into which *GenreGuru* can be integrated, i.e., song recommendation can quickly, and in large volumes, expose the musician to songs with desirable features as they explore how to create their own song. The musician's insights can inform and guide development, but again, at the discretion of the development team.

- Music Theorists
    * *Target subject matter knowledge:* description of current process and/or approach used to generate novel ideas for composing new songs or experimenting with different ideas for arranging existing songs. *Yet again, we must be cautious so as to only consider the experimentation component of the theorist's workflow, not the composition or arrangement process in its core.*
    * *Extent of project involvement:* minimal, i.e., no more than three interviews per theorist – *though, we currently do not have any candidate music theorists.*
    * *Influence on project:* moderate-low – music theorists may already have a process into which *GenreGuru* can be integrated, similar to producers, i.e., song featurization can quickly, and in large volumes, summarize music from a catalogue of songs of interest to identify similarities and differences in their sound properties based on their composition and arrangement. At the discretion of the development team, the music theorist's insights can inform and guide development geared for very musically literate users.

- Music Educators
    * *Target subject matter knowledge:* description of current process and/or approach used to introduce students to novel music concepts through experimentation or experimenting with different ideas for previously-learned (composite) concepts. *We must be cautious so as to only consider the experimen-*

tation component of the teacher's workflow, not the teaching practice or philosophy in its core.

  - *Extent of project involvement:* minimal, i.e., no more than three interviews per teacher – *though, we currently do not have any candidate music teachers.*
  - *Influence on project:* low – music teachers may already have a process into which *GenreGuru* can be integrated, i.e., song generation can (relatively) quickly, and in (relatively) large volumes, produce sound artifacts that introduce novel music concepts or demonstrate alternative use of one or more previously-learned concepts. Like other stakeholders, at the discretion of the development team, the music teacher's insights can inform and guide development geared for *shared* music experimentation settings.

- Affiliated corporation staff – *out of scope*
  - Label staff – *publishers, marketers, lawyers, & executives*
  - Production studio staff – *studio managers, instrument maintainers, & sound designers*
- Development team – *exclusively involves team members, so out of scope.*
- Maintenance team – *exclusively involves team members, so out of scope.*
- Music regulators – *song licensing laws to abide by when acquiring training data is the only applicable concern, otherwise out of scope. For the interim, API documentation and metadata dictionaries suffice as a resource.*

2.4 Hands-On Users of the Project

The first four stakeholders of section 2.3 are the users of concern. To maximize project reach, we do not distinguish between separate user groups with regards to some characteristics, i.e., experience level in the subject matter or technology, attitude toward technology, and physical location. A user can be any combination of: a beginner, novice, intermediate, advanced, or expert in the subject matter or technology, they may be timid to use technology or a technology fanatic, and they can be located anywhere that is within

reach of our service area. What varies between user groups are their relevant responsibilities, outlined below.

- Music Producers & Sound Engineers – *Edit, mix, and master live & recorded audio; facilitate experimentation with instruments, audio effects, and lyrics.*
- Musicians – *Play instruments and/or sing in live & recorded settings; experiment with instruments and vocals.*
- Music Theorists – *Compose new pieces of music; arrange existing music compositions.*
- Music Educators – *Conduct personal and group instruction sessions to present novel music concepts; reintroduce previously-learned music concepts used in a novel setting; present combinations of previously-learned music concepts.*

2.5 Personas

- Music Producers & Sound Engineers
  - Fictitious name – *Brianna Barboza*
  - Fictitious age – *31*
  - Relevant job – *Accountant*
  - Relevant hobbies – *Disc jockeying*
  - Relevant music genres – *pop & hip-hop*
  - Relevant likes/dislikes – TBD after interviews
  - Technology attitude – *comfortable using digital tools, but prefers analog when it comes to audio.*
- Musicians
  - Fictitious name – *Luis Braga*
  - Fictitious age – *24*
  - Relevant job – *N/A, studying for a MSc in Chemistry and Biochemistry from UWaterloo*

  - Relevant hobbies – *Breakdancing*
  - Relevant music genres – *Samba & Bossa Nova*
  - Relevant likes/dislikes – TBD after interviews
  - Technology attitude – *very proficient, he grew up spending his free time in an internet café before starting university.*

- Music Theorists
  - Fictitious name – *Goran Kodeski*
  - Fictitious age – *31*
  - Relevant job – *Consulting*
  - Relevant hobbies – *Collecting LP vinyl records*
  - Relevant music genres – *Folk & Jazz*
  - Relevant likes/dislikes – TBD after interviews
  - Technology attitude – *vehemently anti-digital, owns a flip-phone without a SIM card, and only uses VoIP.*

- Music Educators
  - Fictitious name – *Tumanako "Tui" Teka*
  - Fictitious age – *44*
  - Relevant job – *Music teacher*
  - Relevant hobbies – *Swimming in Lake Waikaremoana*
  - Relevant music genres – *Pūoro Māori*
  - Relevant likes/dislikes – TBD after interviews
  - Technology attitude – *complete beginner, and he only ever goes to the studio to record something he's performed a few times prior.*

2.6 Priorities Assigned to Users

This section builds on 2.4, appointing *music producers* & *musicians* key users, then music theorists & music educators secondary users. These priorities may change as interviews are conducted and different user groups become more concrete.

2.7 User Participation

Further extending 2.4, all users will be notified that they will be involved in no more than 3 interviews as mentioned in the extent of project involvement in 2.3. Should a user be willing to further contribute to the project after three interviews, they will be contacted as sparingly or generously as they outline. Asynchronous communication via e-mails and text are unrestricted, but expected to be within reason and not to cause a disturbance to its recipient.

2.8 Maintenance Users and Service Technicians

The maintenance team exclusively involves the team members, thus is considered out of scope and will not be explored in detail.

3 Mandated Constraints

3.1 Solution Constraints

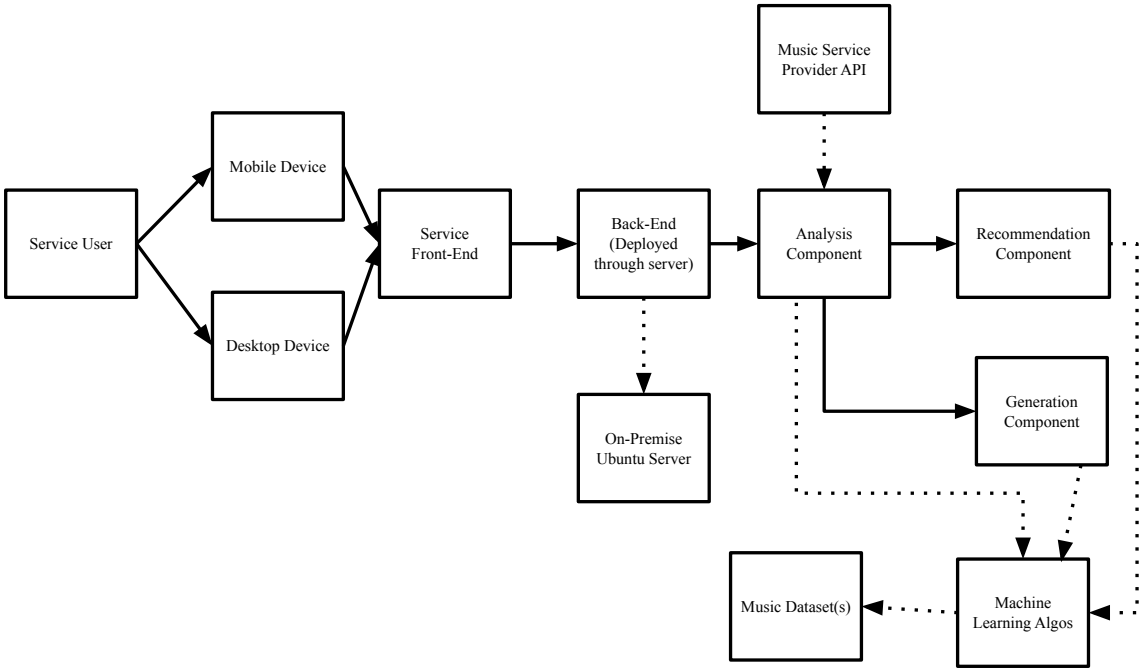
- The service uses a music dataset
  Rationale: A dataset for an AI project is necessary as some form of training data must be used in order to train the AI generative, analysis and recommendation systems.
  Fit Criterion: The machine learning algorithms use a music dataset as the training set.

- The service uses a machine learning algorithm in order to generate song recommendations and snippets.
  Rationale: The general gist of this project is a leveraging of different signal processing and machine learning algorithms in order to provide an end user with a better experience for consuming music. We believe that using a machine learning algorithm for these ends would both be interesting in terms of implementing and training a model, but also practically useful for the end user to provide better recommendations by leveraging training over a very large dataset in order to produce results.
  Fit Criterion: The recommendation and generation components utilize trained machine learning algorithm.

- The service features integration with an existing music streaming provider's platform.
  Rationale: A music service provider (such as Spotify) would allow the service to bypass the need to have music inputted, instead being able to use references to a track. In addition, through API calls, providers such as spotify already have a large amount of useful labels for invidiual track(s) that can be used as features for the machine learning components of the service.
  Fit Criterion: The service has components that make uses of features (such as API calls) that belong to a music service provider's platform.

- The service uses a server to process the user requests separately from the front end.
  Rationale: Ideally, our service would use an on-premise deployed ubuntu server in order to process the user analysis, recommendation, and generation components of the project, as this would allow a more flexible front end design (such as through a web application).
  Fit Criterion: An on-premise server is deployed for the purposes of handling the analysis, generation and recommendations systems separately from the front end.

3.2 Implementation Environment of the Current System



3.3 Partner or Collaborative Applications

Interface	Partner	Integration Type	Rationale
Music Streaming Service API	Youtube, Apple Music, Spotify, Last.fm, etc	API	Some form of a music streaming service API would be hugely useful because it would allow users to interact with the service by pointing to references of a song instead of uploading the entire song file.
Music Generation AI	Jukebox (OpenAI), MusicLM (Google), MusicGen (Meta), Suno AI (Suno.inc)	ML/AI Model	Some form of existing AI model that is tweaked would allow the service to be implemented more efficiently.
Payment API	Stripe, Paypal, etc	API	If the service is to be monetized these options need to be considered.
Analytics Services	Google	API	Tracking user analytics to feedback into the system or to improve user experience.
Cloud Services	Google/AWS, etc	API	Cloud services for remote AI computation, database management, user records, etc.

3.4 Off-the-Shelf Software

There are several existing solutions that could serve as part of the music generation and recommendation system. These include:

- **Spotify API**: Provides access to a vast library of music, including song previews and metadata, which can be leveraged for generating recommendations.
- **Librosa Library**: An open-source Python package for analyzing and processing music files, suitable for extracting features from songs and facilitating generative components.
- **TensorFlow and PyTorch Pre-trained Models**: Both frameworks offer pre-trained models that could be adapted for music generation tasks. These solutions provide a basis for deep learning models without having to build and train from scratch.
- **OpenAI Jukebox**: A generative model that is capable of producing music, which could potentially be adapted and integrated into our system.

These off-the-shelf software solutions provide a foundation upon which we can build our custom features, significantly reducing the development time and leveraging existing technologies to enhance the functionality of our platform.

3.5 Anticipated Workplace Environment

- Home Usage
  Rationale: This environment is more appropriate for casual music listeners. This means users will most likely be using it in a form of a casual environment. This can include such as during a public commute on the train, where the user might try to interact with the service while there is an unreliable internet connection on a mobile device. This means the front end of the service must be easy to use and access on all types of devices, and preferrably, not a large amount of data is streamed between the user and the service. This means that preferrably, the input would have references to pieces of music instead of having say, full .MP3 files as the input. These users are also more likely interested in the recommendation system to attempt to find new pieces of music to listen to, and might be more willing to use the generation system for "fun", so they might be satifisfied even with a generation component that produces odd results.

- Studio Usage
  Rationale: This environment is more professional and what is most likely what creative professionals will be using. The expectation is that they will be using this on their computers, thus they are more likely have a reliable internet connection in addition to being willing to input a lot more music to the service. This means that the service needs to faciliate entering a large amount of music, and be able to generate results in a rapid and efficient manner. Because the users of this type of environment would likely be inviduals engaged in some form of creative endeavors, they will generally have higher standards of what the service provides, and might be more interested in the analysis component, such as being able to find new labels for their existing work to get new perspectives. This also means that they most likely expect the music generation component to give them something more "listenable", something they could directly use as a sample or inspiration.

3.6 Schedule Constraints

This project was started in the 2024 Fall Term and is expected to be completed by the 2025 winter term at McMaster university. Some of the key deadlines are:

- Proof of Concept Demonstration Plan, November 11-22
  This deadline accounts for 5% of the mark. The group should have identified the most significant risks involving the project and come up with plans on how to migitate said risks. If this is not possible, then the project can be redefined. The group also needs to predict and note other concerns about potential problems or difficulties that could arise during development, such as testing difficulties or software portability. If this deadline is not met, then it means that our group does not fully understand the potential pitfalls of our project and we need to redefine certain aspects of the project.

- Final Demonstration, March 24-30
  This deadline accounts for 20% of the mark. This deadline involves a demonstration of a finalized version of the project to supervisors before the public EXPO happening at a TBD date. By this deadline, the service should be in a completed state ready for use and demonstration. If this deadline is not met then the project can be considered a "failure".

- Final Documentation, April 2
  This deadline accounts for 30% of the mark. It involves the finalized documentation of plans pertaining to the project and the actual working program/service. Any final revisions and reflections pertaining to the project should have been completed by this deadline as no further changes to the project will be possible. Whatever documentation or code that is not completed by this deadline would be considered permanently unfinished.

3.7 Budget Constraints

The budget limit as stated by the capstone outline is 750$ CAD. Potential additional costs might include API calls, software liscensing, account fees for cloud services. For the purposes of the demonstration they should not exceed the 750$ limit.

3.8 Enterprise Constraints

There are no specific enterprise constraints as we do not have outside investors for this project.

4 Naming Conventions and Terminology

4.1 Glossary of All Terms, Including Acronyms, Used by Stakeholders involved in the Project

- Track is full song, i.e., a completed audio file published by a musician.
- A snippet is a fragment of a track, i.e., an incomplete audio file. Such audio files can be in a *lossy* compressed format, e.g., `.mp3`, a *lossless* compressed format, e.g., `.FLAC`, or an uncompressed format, e.g., `.WAV`. While the file formats of audio files are known, the source of a song or song snippet can be anything. A concrete example is a 30-second snippet from Spotify's API being an MP3 file delivered through the Spotify content delivery network, `scdn`, e.g., Jack Harlow's "First Class".
- Model: Statistical Model (Not software module)
- Module: Software Component
- AI: Artificial Intelligence
- ML: Machine Learning Model
- Music Metadata: labels obtained through streaming music API (eg spotify) that are like features (eg genre, 'danceability')

5 Relevant Facts And Assumptions

5.1 Relevant Facts

- Music contains the following core features
  - Tempo
  - Key signature

  - Time signature
  - Pitch
  - Timbre
- Song files have metadata that contains information such as:
  - Song title
  - Artist
  - Release date
- Most songs can be classified into multiple genres

5.2 Business Rules

- The user should be able to generate their own music
- The user should be able to figure out what musical features a song contains
- The user should be able to ask for similar songs
- the user should be able to interact with the system without any external installation

5.3 Assumptions

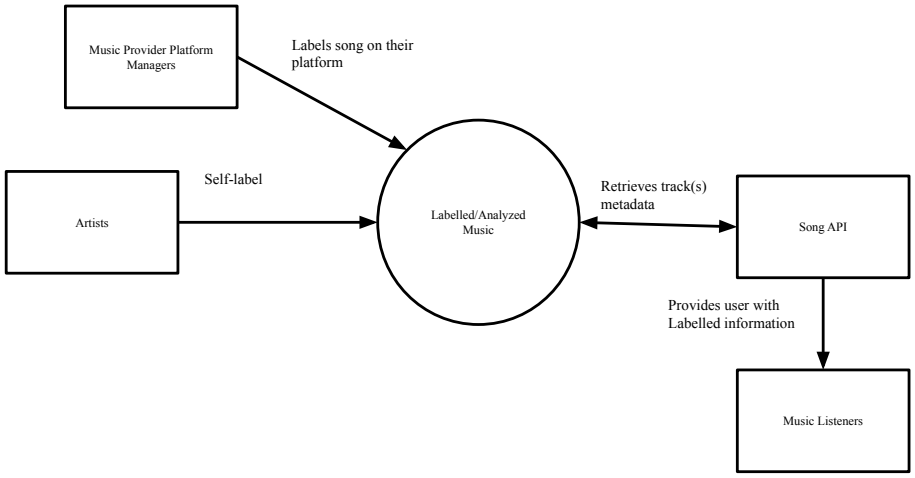
- Users will have at least some familiarity of music theory
- The analysis and recommendation systems will use as many well-established musical features as possible
- All API inputs will be easily accessible and reliable enough to support the recommendation and analysis systems
- The system will be written in a language that all developers are familiar with
- The system will use a local server to handle the processing of the machine learning model and large datasets

- Handling of niche features and cover art are designed to enhance the user experience, but these will not be a part of the core functionality of the system
- The generative system will be completed by the POC demo date
- The recommendation and analysis systems will be completed by the Revision 0 date

6 The Scope of the Work

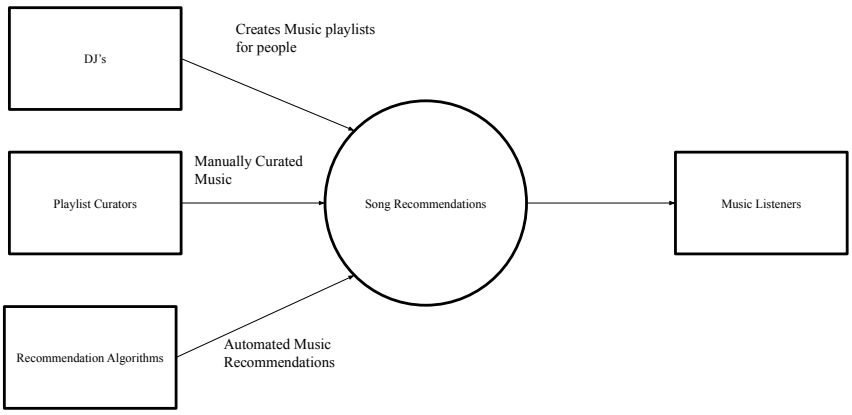
6.1 The Current Situation

- **Music Analysis Component**
  Manual Process: Currently, this is usually manually done by the artist self-labelling their work or third parties labelling the music. This can end up being time consuming.
  Automated Process: Most automation on this regard (retrieving song information from a music provider's api) is mostly useful for retreiving already labelled information, there is not a lot of automation that is be done for labelling, but a lot that is possible retreiving said labels.
  What we wish to do with this component of the service is to gather and present said information that can be gleamed from the music service provider's API to the user in a more presentable manner (such as audio visualizations) but also featurize an individual song for the other components in our service. Our service should also generate new labels using ML algorithms, which allows for faster automation of song labelling and more features for our other components.



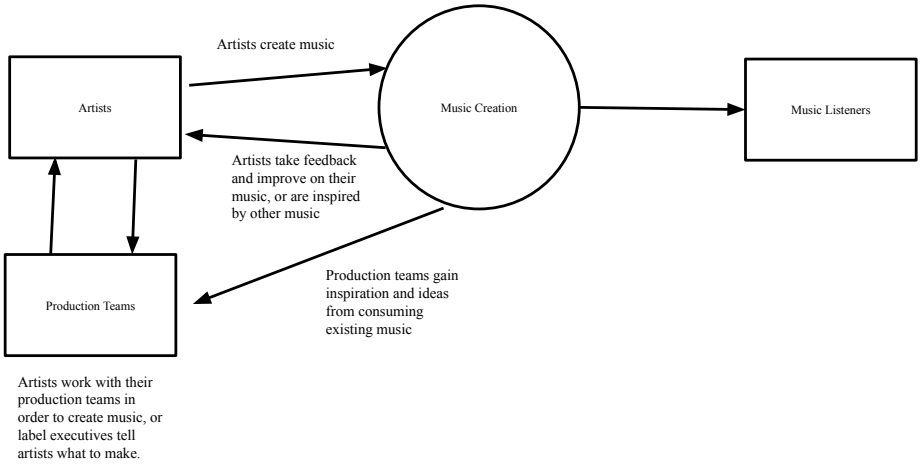
- **Music Recommendation Component**
  Manual Process: Generally, this is mostly done through music curators should as DJs, playlist makers, etc, where curators manually select songs and make recommendations based on popular trends, new artists releases, or shifts in the genre. The main issue here is that it requires a large amount of manual work, as curators need to listen to new music before adding them to some form of playlist or recommending them to other people, which greatly limits their potential recommendations.
  Automated Process: Most music service providers such as spotify, apple music or youtube already offer some form of music recommendation service use algorithms to recommend users new songs based on their listening history. These generally already work decently well, but they end up usually being heavily weighted towards more popular songs (for example, sabrina carpenter's espresso being automatically recommended through smart shuffle no matter what your prior listening history was like), as record labels often have deals with these music service platforms so it is not unusual to get recommendations that do not necessarily suit the user's music tastes.
  The main benefit our service should provide is the ability to recommend

songs that are more niche, more accurate based on the user input while still using an automated system powered by machine learning, as we would have more features due to the analysis component for the ML algorithm to process, in addition to not being heavily weighted towards more popular, trendy songs or some form of likely record label deals meddling.

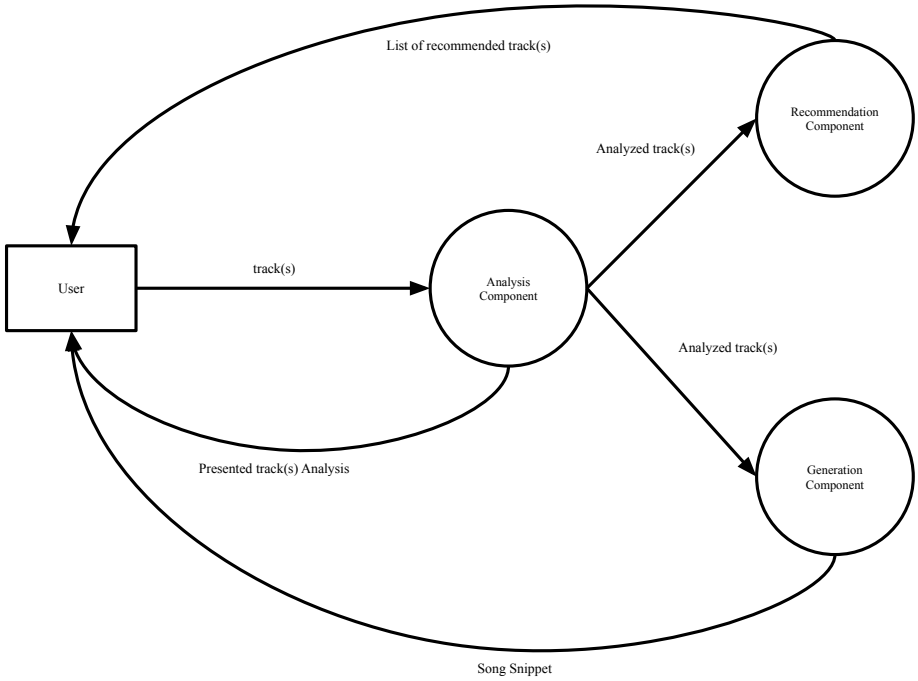


- **Music Generation Component**
  Manual Process: Currently, generating music usually requires years of dedication to the craft and involves the labour of musicians, producers, and composers. This means that there is a large amount of manpower that is necessary in order to create new unique pieces.

  We wish to have our service provide a new avenue to generate new types of music. We do not envision this service as something that replaces the traditional music creation process, but rather as something that can potentially enhance it.



6.2 The Context of the Work



6.3 Work Partitioning

Event Name	Input/Output	Summary
User Requests Song(s) Analysis	Track(s) (input)	Analysis component receives user input song(s)
User Requests Recommendations	Track(s) (input)	User inputs track(s) to get track recommendations
User Requests music generation	Track(s) (input)	User inputs track(s) to get a generated snippet
Song(s) Analyzed for generation component	Analyzed Track(s) (input)	After the user requests the song generation feature, the track(s) are analyzed for the generation component
Song(s) Analyzed for recommendation component	Analyzed Track(s) (input)	After the user requests the song recommendations feature, the track(s) are analyzed for the recommendation component
Song recommendations generated	List of recommended tracks (output)	The service returns a list of generated songs to the user
Song Snippet generated	Track Snippet (output)	The service returns the generated snippet to the user
Song(s) analysis returned to user	Song Analysis Presentation (output)	The service responds to the user request for an analysis of their inputted track(s)

6.4 Specifying a Business Use Case (BUC)

- **User Requests Track Recommendations**
  Primary Actor: App User
  Trigger: User intiates a request with the service to generate some new songs to listen to.
  Preconditions:
  - User has valid references to tracks for the service to accept
  - ML generation components are currently active

  Main Success Scenario:
  - user submits request for track recommendation
  - user inputs track(s) to the service
  - the service's analyzation component analyzes the tracks

- analyzation component passes on tracks & analysis to recommendation component
- recommendation component generates tracks
- user receives new tracks to listen to

- **User Requests Track Analysis**
  Primary Actor: App User
  Trigger: User intiates a request with the service to generate some new songs to listen to.
  Preconditions:
  - User has valid references to tracks for the service to accept
  - ML generation components are currently active

  Main Success Scenario:
  - user submits request for track recommendation
  - user inputs track(s) to the service
  - the service's analyzation component analyzes the tracks
  - analyzation component generates presentations and visualizations based on the inputted track(s)
  - the service returns generated presentations and visualizations to the user

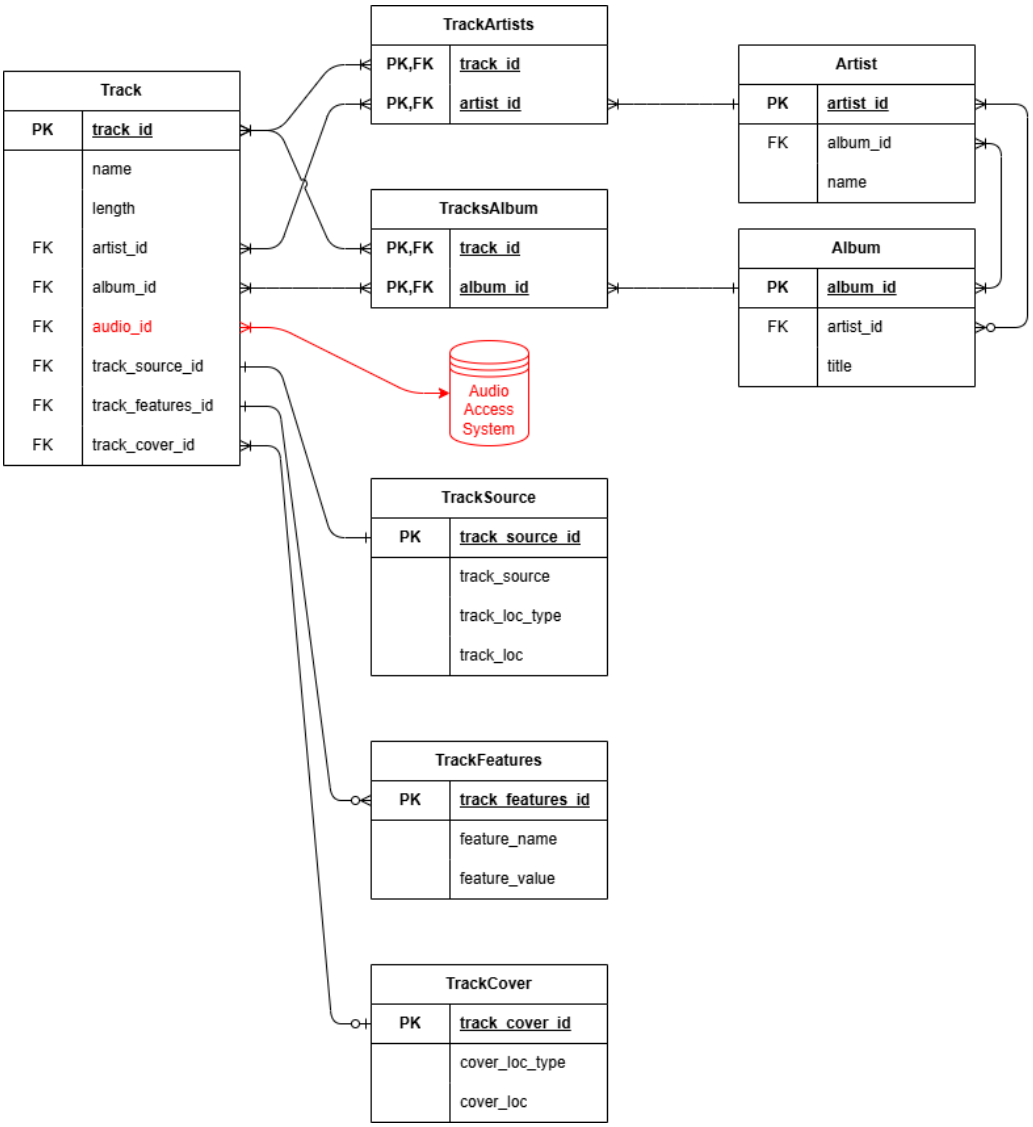
- **User Requests Snippet Generation**
  Primary Actor: App User
  Trigger: User intiates a request with the service to generate some new songs to listen to.
  Preconditions:
  - User has valid references to tracks for the service to accept
  - ML generation components are currently active

  Main Success Scenario:
  - user submits request for snippet generation
  - user inputs track(s) to the service

- the service's analyzation component analyzes the tracks
- analyzation component passes on tracks & analysis to generation component
- generation component generates tracks
- user receives a snippet to listen to

7 Business Data Model and Data Dictionary

7.1 Business Data Model



7.2 Data Dictionary

7.2.1 Relations

Please refer to table 7.2.1.

Relation Name	Relation Description
Track	Primary relation containing track name, track length, & references to other relations.
TrackArtists	Intermediate relation to retain information about which artists appear on which tracks.
TracksAlbum	Intermediate relation to retain information about which tracks appear on which albums, if applicable.
TrackSource	Secondary relation containing information about the source of a track, e.g., Spotify, Deezer, Amazon music, BandCamp, or local.
TrackFeatures	Secondary relation contianing information about the features of a track, following the entity-attribute-value (EAV) model.
TrackCover	Abstract relation containing information about the cover art of the track.
Artist	Secondary relation containing information about artists.
Album	Secondary relation containing information about albums.
Audio Access System	Abstract storage system containing information about the track audio.

Table 1: Relations Data Dictionary

7.2.2 Attributes

Please refer to table 7.2.2.

Attribute Name	Attribute Description
track_id	Unique track identifier.
artist_id	Unique artist identifier.
album_id	Unique album identifier.
audio_id	Unique track audio reference identifier.
track_source_id	Unique track source reference identifier.
track_features_id	Unique track features reference identifier.
track_cover_id	Unique track cover art reference identifier.
Track.name	Track name.
Track.length	Track length.
Artist.name	Artist name.
Album.title	Album title
TrackSource.track_source	Track source name.
TrackSource.track_loc_type	Track source location type, e.g., "URL" or "path".
TrackSource.track_loc	Track location, i.e., a URL or path.
TrackFeatures.feature_name	Track feature name, e.g., popularity.
TrackFeatures.feature_value	Track feature corresponding value, e.g., 88.
TrackCover.cover_loc_type	Track cover art location type, e.g., "URL" or "path".
TrackCover.cover_loc	Track cover art location, i.e., a URL or path.

Table 2: Attributes Data Dictionary

7.2.3 Relationships

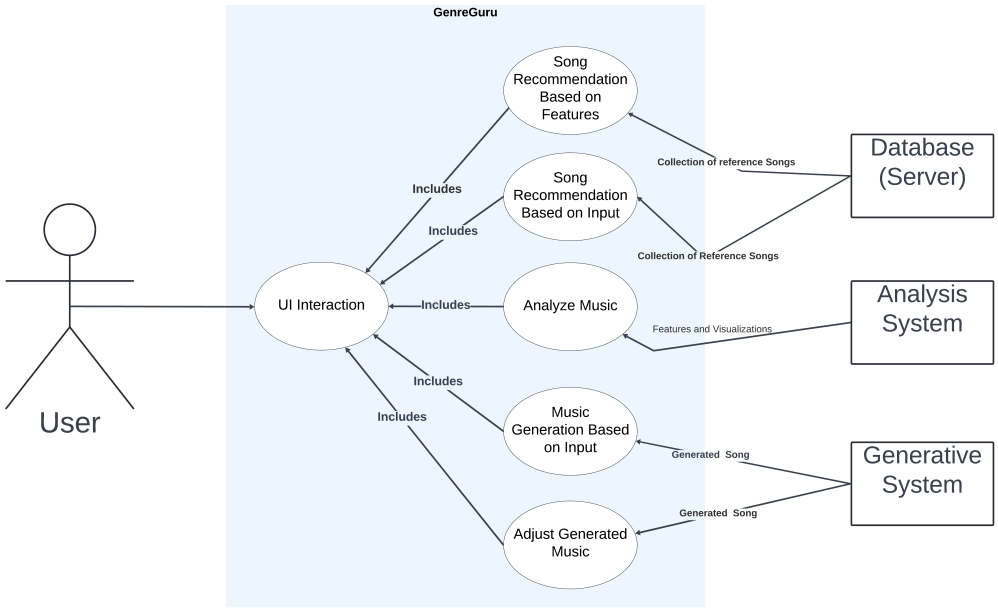
Please refer to table 7.2.3. Please consult Entity Relationship Model - Crow's Foot Notation as a notation reference if necessary - it is mostly useful for the connector multiplicities.

Relation 1	Relation 2	Relationship Description
Track	*TrackArtists*	A track can correspond to multiple artists, and vice versa.
Track	*TracksAlbum*	A track can correspond to multiple albums (a single and a long-playing (LP) album), and vice versa.
Track	Audio Access System	Multiple tracks can correspond to a single instance of an audio file, i.e., a single and an LP release of a song point to the same audio file.
Track	*TrackSource*	A track should only belong to one and only one source to avoid duplication.
Track	*TrackFeatures*	A track can correspond to 0 or more features, i.e., a song with no features has not been featurized and a song with multiple features has. This allows modular and evolvable data storage.
Track	*TrackCover*	A track can correspond to at most one cover, i.e., it can have no cover art. The cover art of an album can also apply to its constitutent tracks.
TrackArtists	*Artist*	A mapping relationship, where an artist can exist on multiple tracks.
TracksAlbum	*Album*	A mapping relationship, where an album can contain multiple tracks.
Artist	*Album*	An artist can have 0 or more albums, but an album must have at least one artist.

Table 3: Relationships Data Dictionary

8 The Scope of the Product

8.1 Product Boundary



8.2 Product Use Case Table

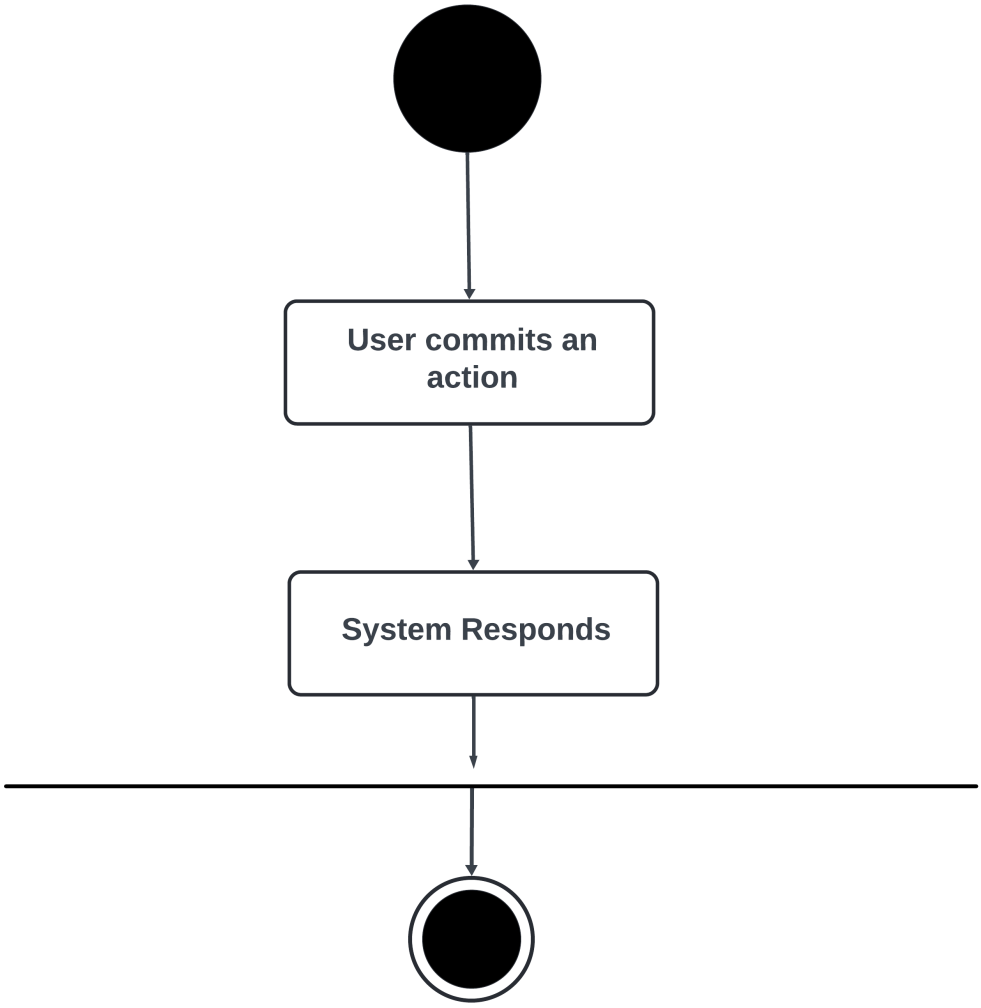
PUC No	PUC Name	Actor/s	Input & Output
1	UI Interaction	User	User Actions (click, swipe, drag) (in) System Response (out)
2	Song Recommendation Based on Features	User	User's desired features (in) Collection of reference songs (out)
3	Music Generation Based on Input	User	Reference song(s) and/or song snippet(s) (in) Generated song or song snippet (out)
4	Analyze Music	User	Reference song or song snippet (in) Collection of features and visualizations (out)
5	Song Recommendation Based on Input	User	Reference song(s) (in) Collection of reference songs (out)
6	Server Interaction for Music Generation	Server	Reference song(s) and/or song snippet(s) (in) Generated song or song snippet (out)
7	Server Interaction for Song Recommendation	Server	User's desired features or reference song(s) and/or snippet(s) (in)Collection of reference songs (out)
8	Server Interaction for Music Analysis	Server	Reference song or song snippet (in) Collection of features and visualizations (out)

Table 4: Product Use Case Table

8.3 Individual Product Use Cases (PUC's)

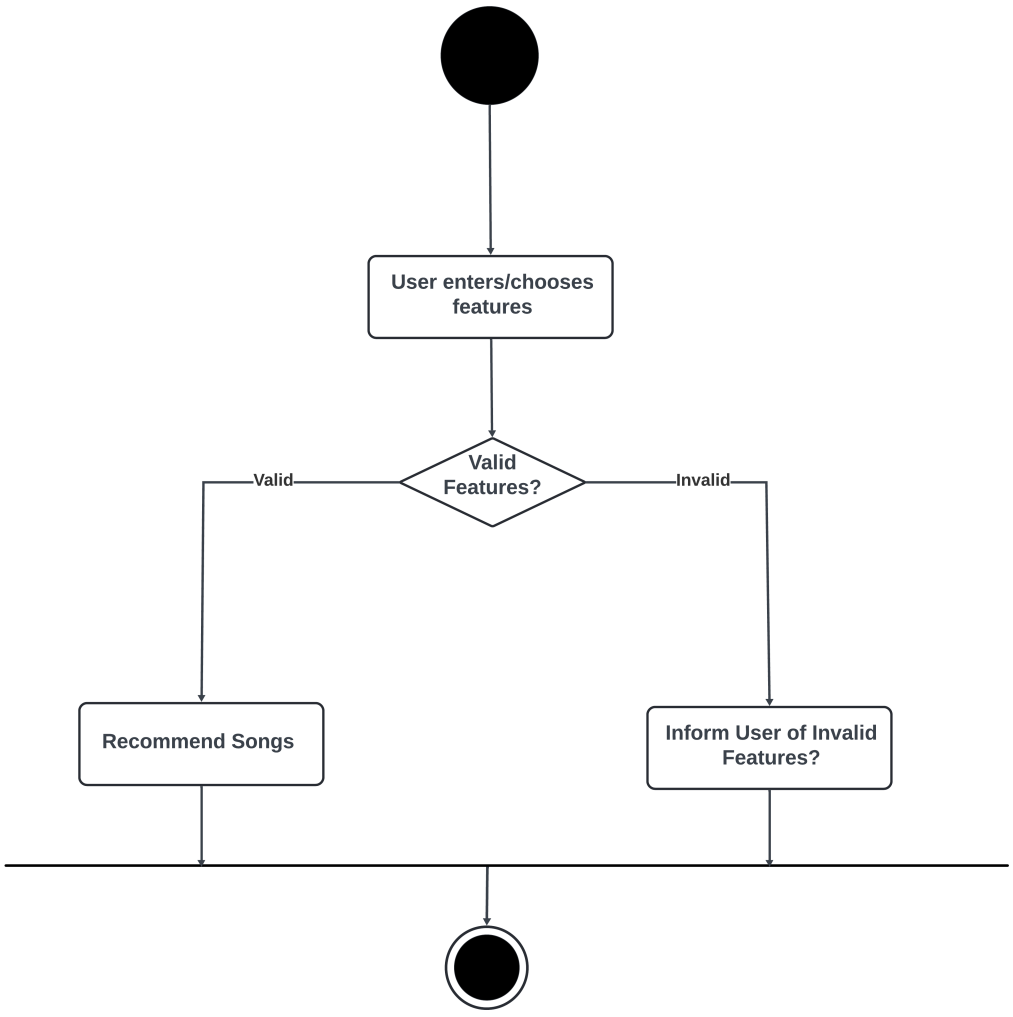
1. Product Use Case Name: UI Interaction
Trigger: User commits some action (e.g. clicking, swiping, dragging)
Preconditions: User has successfully accessed GenreGuru, or is already in

GenreGuru
Interested Stakeholders: All
Actor/s: User
Activity Diagram:



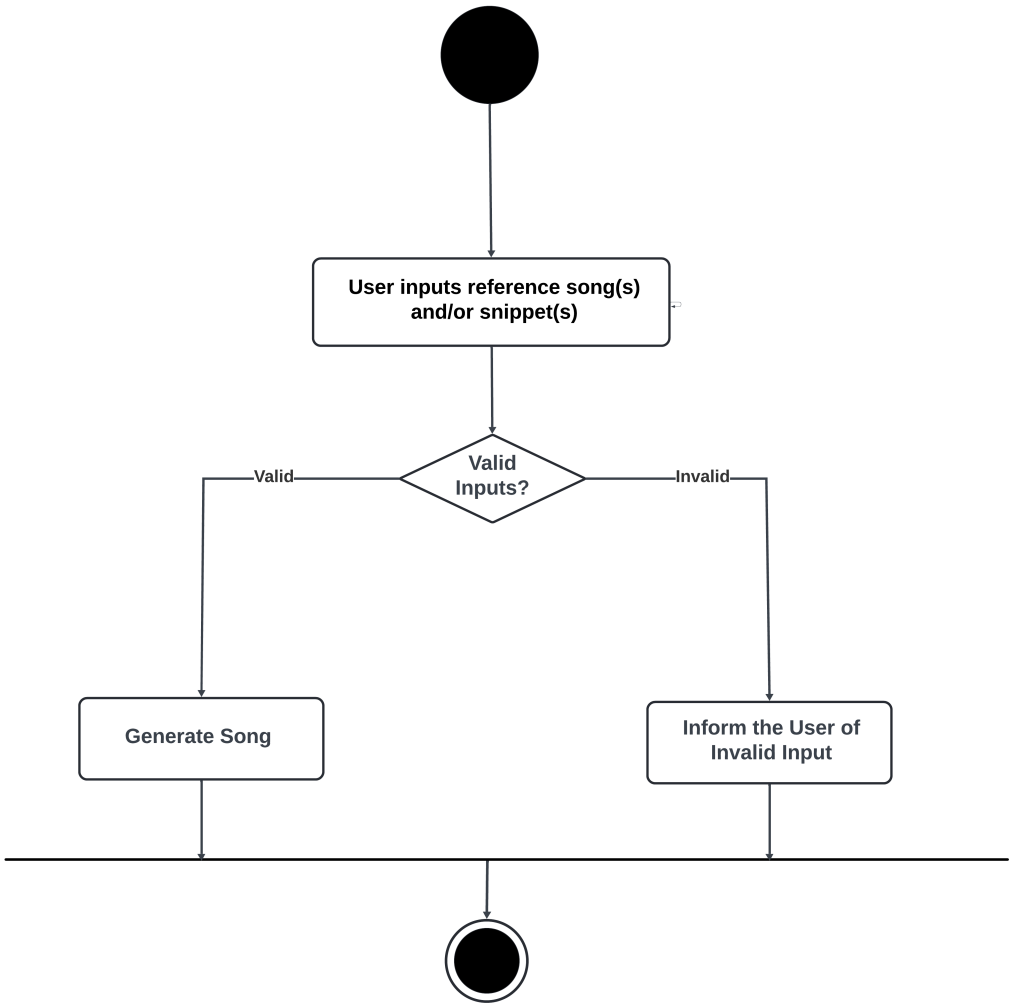
Outcome: The user will commit an action like swiping or pressing and the system will react depending on the action.

2. Product Use Case Name: Song Recommendation Based on Features
Trigger: User picks features, and indicates they want to search for recommendations
Preconditions: User must have GenreGuru open, the user has selected features to search for
Interested Stakeholders: Casual Music Listeners, Hobbyist Musicians
Actor/s: User
Activity Diagram:



Outcome: The user will select or manually enter features they are looking for in a song, and the system will first check to see if the features they selected/entered are valid, and the system will return a collection of reference songs that match those features.

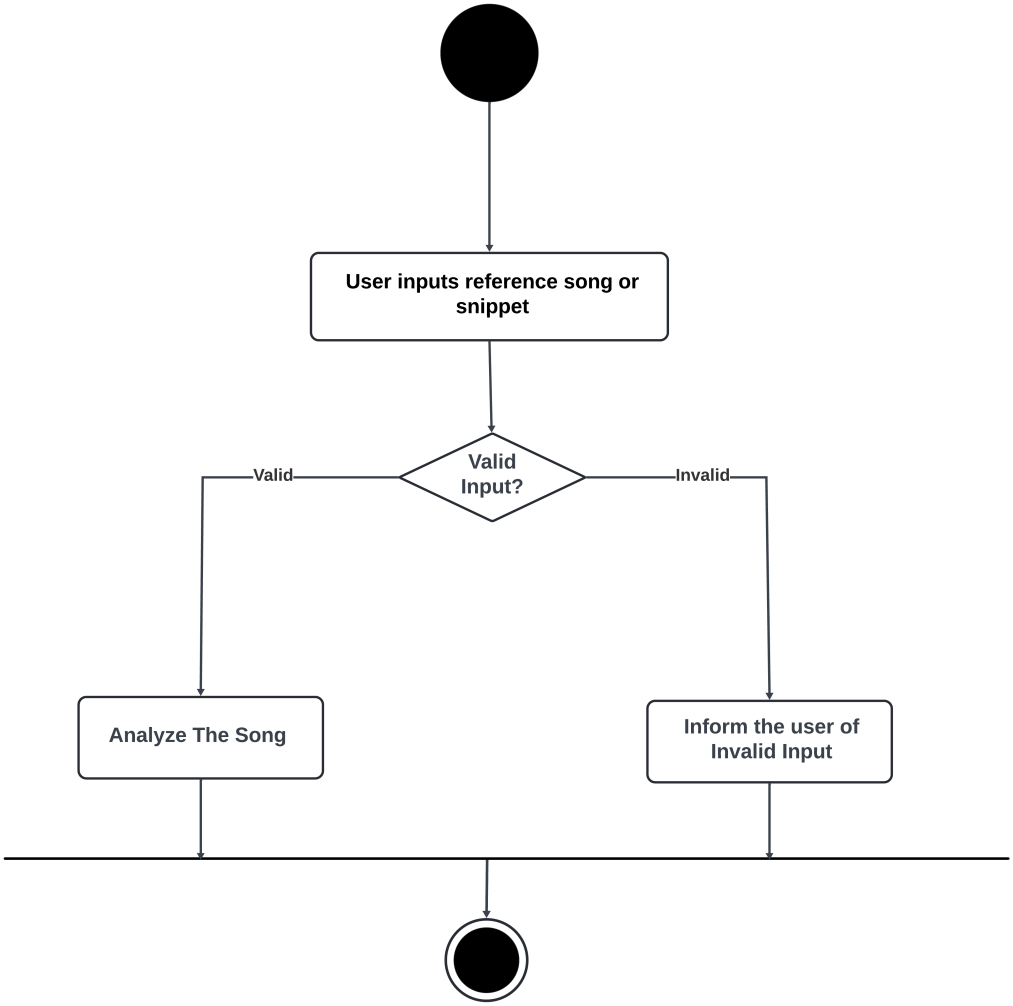
3. Product Use Case Name: Music Generation Based on Input
Trigger: User inputs reference song(s) and/or song snippet(s), and indicates they want to generate a song
Preconditions: User must have GenreGuru open, and the user has provided a valid input(s)
Interested Stakeholders: Music producers, Hobbyist Musicians
Actor/s: User
Activity Diagram:



Outcome: The user will enter song(s) and/or song snippet(s) and indicate to the system that they want to generate music, the system will check that these inputs are valid (correct format) and then will generate a song and return it to the user.

4. Product Use Case Name: Analyze Music
Trigger: User inputs a reference song or song snippet and indicates they want to analyze the music

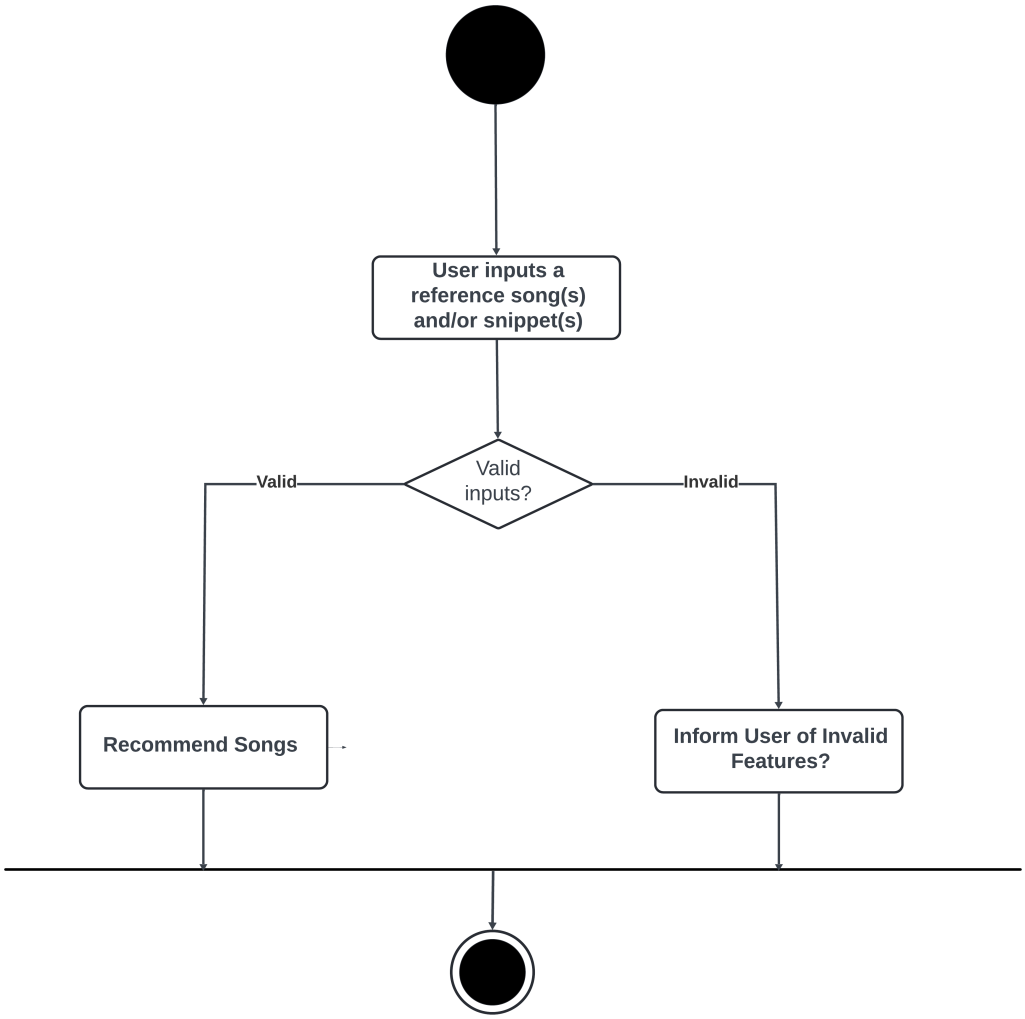
Preconditions: User must have GenreGuru open, and the user has provided a valid input
Interested Stakeholders: Music Producers, Audio Engineers, Music Educators
Actor/s: User
Activity Diagram:



Outcome: The user will input a reference song or song snippet and indicate they want to analyze the song, the system will validate the input and return

a set of features and visualizations.

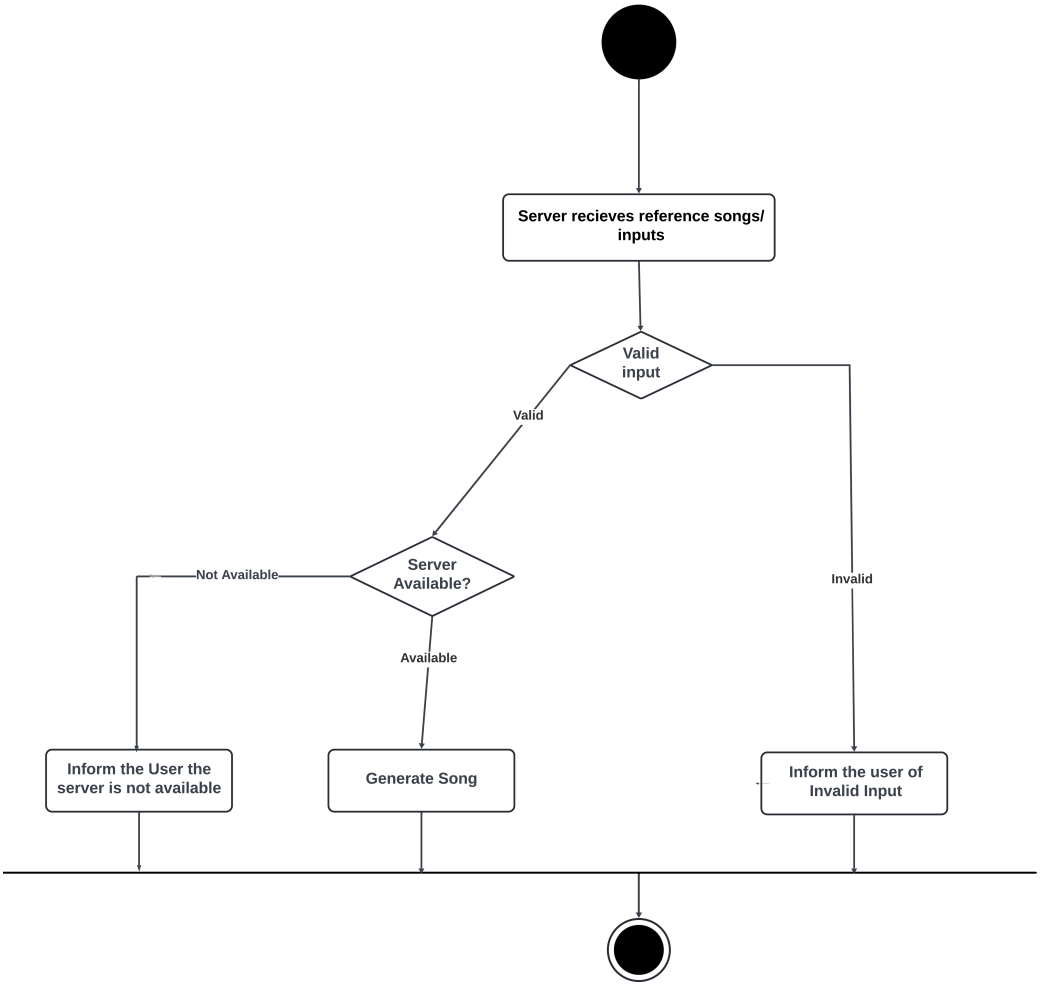
5. Product Use Case Name: Song Recommendation Based on Input
Trigger: User inputs reference song(s) and/or snippet(s), and indicates they want to search for recommendations
Preconditions: User must have GenreGuru open, and the user has provided a valid input(s)
Interested Stakeholders: Casual Music Listeners, Hobbyist Musicians
Actor/s: User
Activity Diagram:



Outcome: The users will input reference song(s) and/or snippet(s), the system will first check to see if the inputs are valid. Then the system will return a collection of reference songs.

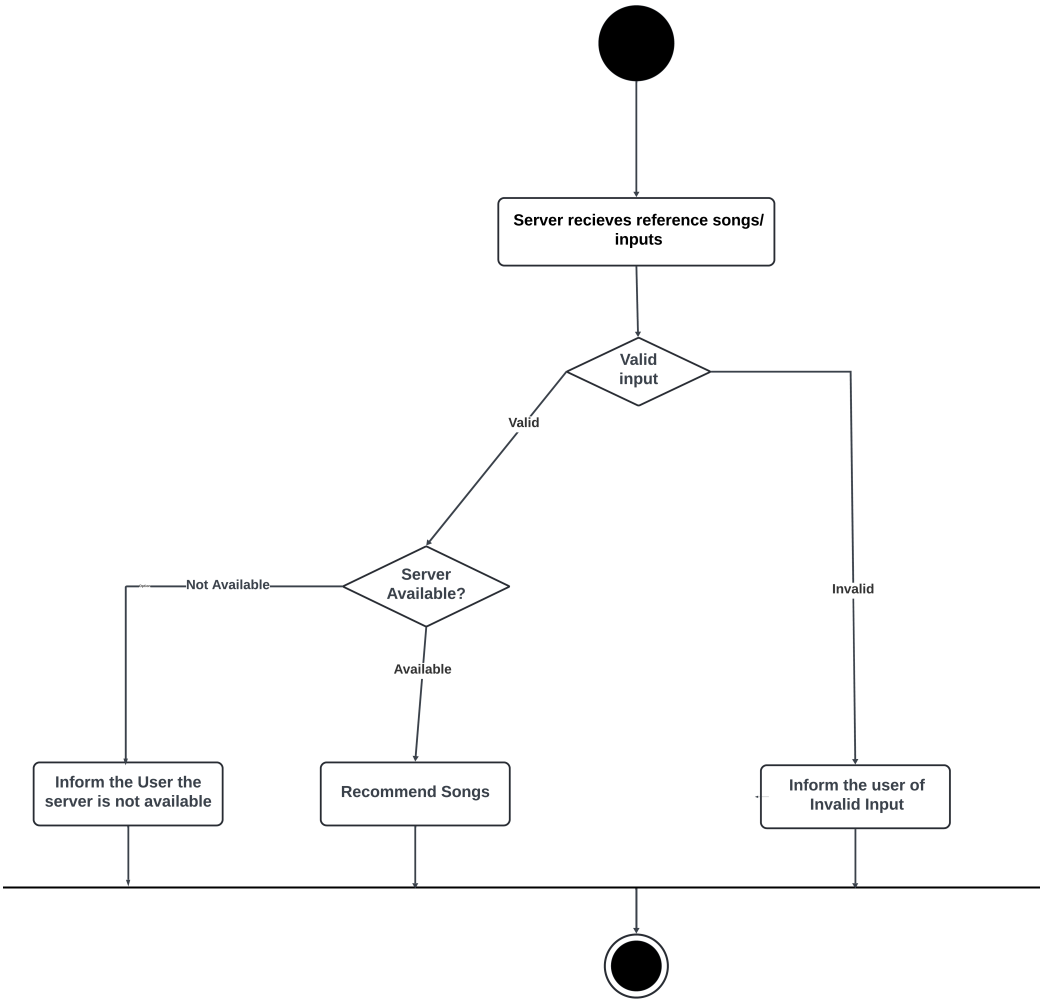
6. Product Use Case Name: Server Interaction for Music Generation
Trigger: User submits a reference song and/or snippet and requests music

generation
Preconditions: User has provided a valid input through, and the server is operational
Interested Stakeholders: Music Producers, Hobbyist Musicians
Actor/s: Server
Activity Diagram:



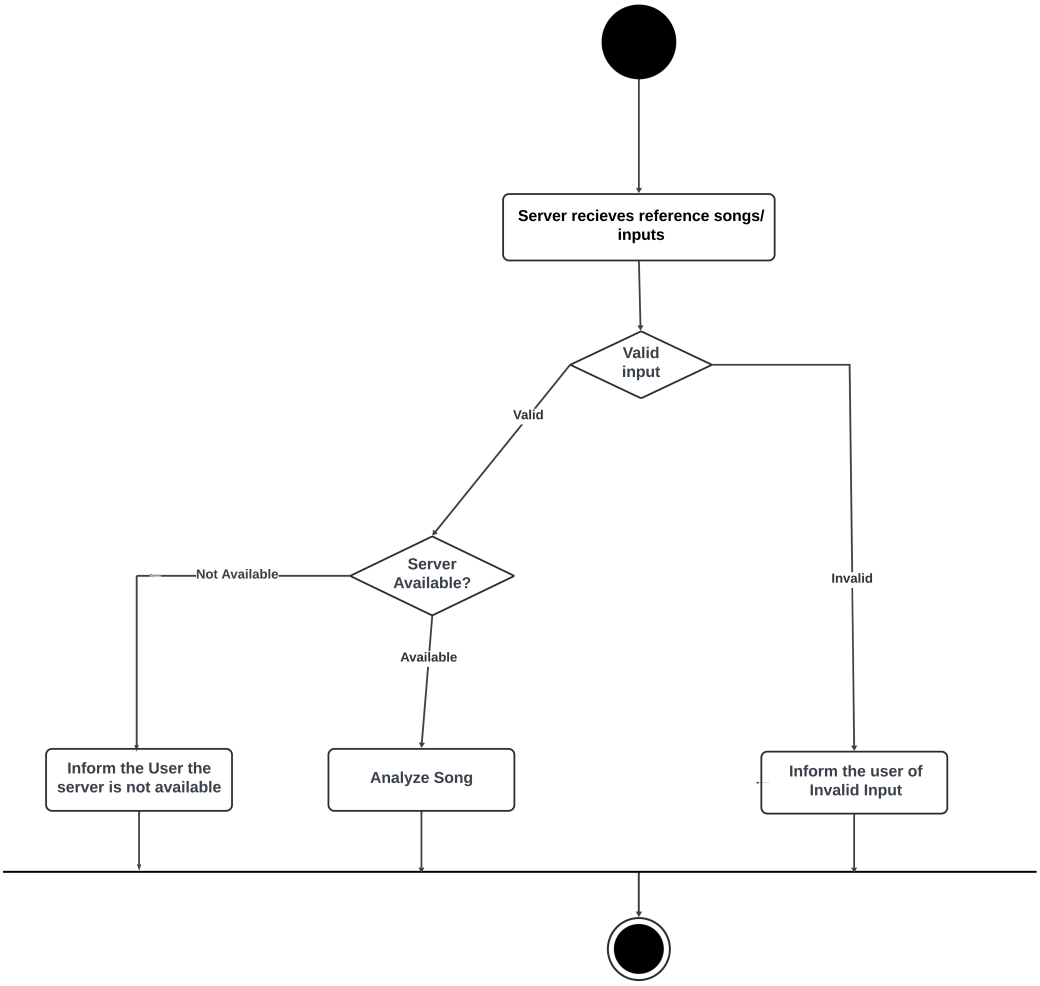
Outcome: The server processes the input, generates music, and returns the generated song to the user

7. Product Use Case Name: Server Interaction for Song Recommendation
Trigger: User submits desired features or reference songs/snippets and requests song recommendations
Preconditions: User has provided valid input, and the server is available
Interested Stakeholders: Casual Music Listeners, Hobbyist Musicians
Actor/s: Server
Activity Diagram:



Outcome: The server processes the input and returns a collection of recommended songs based on the input features or reference songs/snippets.

8. Product Use Case Name: Server Interaction for Music Analysis
Trigger: User submits a reference song or snippet and requests music analysis
Preconditions: User has provided a valid input, and the server is ready to analyze
Interested Stakeholders: Music Producers, Audio Engineers, Music Educators
Actor/s: Server **Activity Diagram:**



Outcome: The server analyzes the song or snippet and returns a collection of features and visualizations to the user.

9 Functional Requirements

9.1 Functional Requirements

Requirement # 1
Requirement Type: 9
Event/Use Case #: (need to figure this out later)
Description: The system must allow users to input song queries via text or upload audio files
Rationale: Provides the primary interface for user interaction
Originator: Requirement Analyst
Fit Criterion: The GUI should display input fields for text and file upload
Customer Satisfaction: 5
Customer Dissatisfaction: 0
Priority: High
Conflicts: None
Supporting Material: None
History: Created January 18, 2025

Requirement # 2
Requirement Type: 9
Event/Use Case #: (need to figure this out later)
Description: The system GUI must display search results, including song title, artist, and album cover
Rationale: Enables users to identify the desired track.
Originator: Requirement Analyst
Fit Criterion: Results should be displayed within 10 seconds after a query is submitted
Customer Satisfaction: 5
Customer Dissatisfaction: 0
Priority: High
Conflicts: None
Supporting Material: None
History: Created January 18, 2025

Requirement # 3
Requirement Type: 9
Event/Use Case #: (need to figure this out later)
Description: The system must allow users to play a 30-second preview of each search result
Rationale: Helps users confirm their song selection
Originator: Requirement Analyst
Fit Criterion: Clicking the "Preview" button should play the audio clip
Customer Satisfaction: 5
Customer Dissatisfaction: 0
Priority: High
Conflicts: None
Supporting Material: None
History: Created January 18, 2025

Requirement # 4
Requirement Type: 9
Event/Use Case #: (need to figure this out later)
Description: The system must validate uploaded audio files to ensure they are in supported formats
Rationale: Prevents processing errors from unsupported files
Originator: Requirement Analyst
Fit Criterion: The system should reject unsupported formats with an error message
Customer Satisfaction: 5
Customer Dissatisfaction: 0
Priority: High
Conflicts: None
Supporting Material: None
History: Created January 18, 2025

Requirement # 5
Requirement Type: 9
Event/Use Case #: (need to figure this out later)
Description: The system must convert non-WAV files to WAV format before processing
Rationale: Standardizes inputs
Originator: Requirement Analyst
Fit Criterion: An uploaded file should be converted to WAV
Customer Satisfaction: 5
Customer Dissatisfaction: 0
Priority: High
Conflicts: None
Supporting Material: None
History: Created January 18, 2025

Requirement # 6
Requirement Type: 9
Event/Use Case #: (need to figure this out later)
Description: The system must construct a Spotify query based on user input and retrieve a list of matching songs
Rationale: Enables the discovery of songs via text based search
Originator: Requirement Analyst
Fit Criterion: A query string should be generated and sent to the Spotify API
Customer Satisfaction: 5
Customer Dissatisfaction: 0
Priority: High
Conflicts: None
Supporting Material: None
History: Created January 18, 2025

Requirement # 7
Requirement Type: 9
Event/Use Case #: (need to figure this out later)
Description: The system must extract features
Rationale: Provides the necessary data for recommendations
Originator: Requirement Analyst
Fit Criterion: The extracted features should match ground truth with an accuracy of at least 90%
Customer Satisfaction: 5
Customer Dissatisfaction: 0
Priority: High
Conflicts: None
Supporting Material: None
History: Created January 18, 2025

Requirement # 8
Requirement Type: 9
Event/Use Case #: (need to figure this out later)
Description: The system must support feature extraction for all 8 features listed in the project scope
Rationale: Ensures compatibility with all defined use cases.
Originator: Requirement Analyst
Fit Criterion: Each feature module must output valid data for a supported audio file
Customer Satisfaction: 5
Customer Dissatisfaction: 0
Priority: High
Conflicts: None
Supporting Material: None
History: Created January 18, 2025

Requirement # 9
Requirement Type: 9
Event/Use Case #: (need to figure this out later)
Description: The system must generate song recommendations based on extracted features
Rationale: Helps users discover new songs aligned with their input
Originator: Requirement Analyst
Fit Criterion: Recommendations should be displayed
Customer Satisfaction: 5
Customer Dissatisfaction: 0
Priority: High
Conflicts: None
Supporting Material: None
History: Created January 18, 2025

Requirement # 10
Requirement Type: 9
Event/Use Case #: (need to figure this out later)
Description: The system must transmit user-selected inputs (text queries or audio files) to the server for processing
Rationale: Facilitates communication between the client and server
Originator: Requirement Analyst
Fit Criterion: Data packets should reach the server without loss or corruption
Customer Satisfaction: 5
Customer Dissatisfaction: 0
Priority: High
Conflicts: None
Supporting Material: None
History: Created January 18, 2025

Requirement # 11
Requirement Type: 9
Event/Use Case #: (need to figure this out later)
Description: The system must receive processed results (recommendations) from the server
Rationale: Displays processed information to the user
Originator: Requirement Analyst
Fit Criterion: GUI updates with the processed results (recommendations)
Customer Satisfaction: 5
Customer Dissatisfaction: 0
Priority: High
Conflicts: None
Supporting Material: None
History: Created January 18, 2025

10 Look and Feel Requirements

10.1 Appearance Requirements

APR1. The system shall present usable options without clutter.
APR2. The system shall be visually accessible by using highly contrasting elements.
APR3. The system shall use an easily navigable, professional design.

10.2 Style Requirements

STR1. The system shall be responsive to various screen sizes.

11 Usability and Humanity Requirements

11.1 Ease of Use Requirements

EUR1. The system shall show tooltips pop-ups upon users hovering on an interactable element.

11.2 Personalization and Internationalization Requirements

PIR1. The system shall support different display colour themes.

11.3 Learning Requirements

LR1. The system shall provide a first-time user guide to use features upon first launch.
LR2. The user shall be able to complete the tutorial within *10* minutes.

11.4 Understandability and Politeness Requirements

UPR1. The system shall not generate vulgar content. **UPR2.** The system shall only use standardized, inoffensive iconography wherever icons are needed.

11.5 Accessibility Requirements

ACR1. The system shall only use accessible fonts.
ACR2. The system shall offer a color-blind mode.

12 Performance Requirements

12.1 Speed and Latency Requirements

N/A – while there is a server and networking in the system, such requirements are considered out of scope.

12.2 Safety-Critical Requirements

N/A – there are no safety-affiliated components affiliated with the project.

12.3 Precision or Accuracy Requirements

PAR1. Query request times are displayed to within 1 minute of precision.
PAR2. Displayed percentages will be rounded to the nearest two decimals.

12.4 Robustness or Fault-Tolerance Requirements

RAR1. The system shall maintain 97.5% uptime, i.e., an average downtime of 18 hours every month.

12.5 Capacity Requirements

CR1. The system shall handle concurrent usage by up to four users.
CR2. The system shall store a cached record of songs and corresponding information.

12.6 Scalability or Extensibility Requirements

N/A – while there is a server in the system, scaling is not in scope.

12.7 Longevity Requirements

N/A – maintenance and longevity of the project is not in scope, at least not until after the POC demonstration.

13 Operational and Environmental Requirements

13.1 Expected Physical Environment

EPER1. The system server shall run on a modified Dell OptiPlex 3050.

13.2 Wider Environment Requirements

N/A

13.3 Requirements for Interfacing with Adjacent Systems

N/A

13.4 Productization Requirements

PRR1. The system front-end shall be accessible from a browser across different devices.

13.5 Release Requirements

NA – the project is academic in nature, so a release cycle is unnecessary.

14 Maintainability and Support Requirements

14.1 Maintenance Requirements

MR1. The system shall facilitate switching between music providers overnight.

14.2 Supportability Requirements

NA – the project is academic in nature, and its lifespan is unlikely to extend such that it needs support.

14.3 Adaptability Requirements

N/A – a web-app is the most portable format for the specified target users.

15 Security Requirements

15.1 Access Requirements

ACR1. The user shall only be able to access songs they upload or ones that are licensed for their use.

15.2 Integrity Requirements

IR1. *The server shall perform a weekly backup of its database to prevent data loss in the event of a catastrophic failure.*
IR2. *The server shall implement deduplication measures to guarantee no data redundancy.*

15.3 Privacy Requirements

PR1. The system shall only expose query requests to the user that made them.

15.4 Audit Requirements

AUR1. The system shall maintain a history of user query requests.

15.5 Immunity Requirements

N/A – again, the project is academic in nature, and its purpose is not to be robust for attacks on a commercial software.

16 Cultural Requirements

16.1 Cultural Requirements

CUR1. The system shall not contain any (potentially) culturally offensive iconography or language.

17 Compliance Requirements

17.1 Legal Requirements

LGR1. The system shall not make use of any copyrighted material without express permission.
LGR2. The data collection process must obey all potential API developer rules.

17.2 Standards Compliance Requirements

N/A – out of project scope.

18 Open Issues

- Generated output use regulation, #1
  Summary: The service EULA agreement might need to consider how and where the generated snippet by the service is allowed to be used.

  This is important because the training data or inputted track(s) into the service might have copyright laws & rules regulating commerical use associated with them, thus we need to consider whether those rules & regulations also apply with the generated snippet.

  For example, the EULA might need to state that generated snippets cannot be used for commercial purposes if the input contains song from a specific artist belonging to some specific record label. This means that a creative professional's ability to actually use this portion of the service could be greatly limited depending on what tracks they are inputting, and/or what the machine learning algorithm was trained on.

- Dataset Bias, #2
  Summary: We are currently not exactly certain about what dataset we will be using for the algorithm training. For example, if we are using some form of non-copyrighted music, there's a likely chance it is heavily biased towards electronic dance music as a genre. This could skew the algorithm's ability to interact with music that is not of that genre, thus the service might not be able to properly process requests users who perhaps listen to more niche genres of music. This means we might need to train the machine learning algorithms on multiple different datasets, or we need to manually modify an existing dataset to integrate more niche genres of music as the training source.

19 Off-the-Shelf Solutions

19.1 Ready-Made Products

Insert your content here.

19.2 Reusable Components

Insert your content here.

19.3 Products That Can Be Copied

Insert your content here.

20 New Problems

20.1 Effects on the Current Environment

The platform's high computational requirements, especially during the training and usage of generative models, may put a strain on existing computational resources. This could lead to slower performance for other applications that share these resources. Additionally, increased data transmission due to the integration with APIs such as Spotify and Deezer could lead to higher network bandwidth utilization, potentially affecting the performance of other network-dependent systems. Mitigating these effects may require infrastructure scaling, such as increasing server capacity or optimizing data processing, to ensure that the existing systems are not disrupted.

20.2 Effects on the Installed Systems

The new platform will interact with various installed systems, leading to the introduction of new dependencies. For example, using third-party APIs like Spotify or integrating machine learning frameworks such as TensorFlow may require changes to the current software stack. These dependencies could introduce compatibility issues, particularly if third-party providers update their APIs or if new versions of required libraries are released. Such updates might necessitate timely system changes to maintain functionality and avoid disruptions. Ensuring that installed systems remain compatible requires diligent version management, comprehensive testing, and a structured process for managing library and API updates.

20.3 Potential User Problems

The platform is designed to offer advanced features, which may present usability challenges for some users. For example, users with limited technical skills may struggle to understand the process of modifying musical features like key, rhythm, or tempo, which are essential for customizing music generation. Furthermore, the quality of generated outputs is inherently subjective,

and users may feel frustrated if the system-generated music does not meet their expectations. Another potential problem is system latency, especially when resource-intensive operations like model inference are being performed. Long waiting times could negatively impact user experience, making the platform feel less responsive, which may deter regular use.

20.4 Limitations in the Anticipated Implementation Environment That May Inhibit the New Product

The platform is expected to be implemented on a local server, which presents certain limitations. The local server environment will need sufficient processing power to handle computationally intensive tasks, such as model training and music generation. If the server infrastructure is under-resourced, performance issues such as latency and reduced processing speed could arise. Additionally, reliance on a local server means that the platform's effectiveness is tied directly to server availability and maintenance. Any server downtime could render the platform inaccessible to users. If we decide to extend the platform's implementation onto personal devices, new limitations arise. Personal devices, especially those with limited hardware capabilities, may struggle with the computational requirements of generative music models, leading to significant performance bottlenecks. This could make the platform inaccessible or difficult to use for some users. Moreover, the reliance on internet connectivity for accessing external APIs is another major limitation. Users in areas with unreliable internet or low bandwidth may face difficulties, particularly with real-time features like music recommendations or generating new songs based on streaming data. These limitations could restrict the user base to only those with high-performance devices and stable internet connections.

20.5 Follow-Up Problems

Once the platform is deployed, several follow-up issues will need to be addressed to maintain and improve the product. One of the primary challenges will be keeping up with third-party API changes, such as modifications to Spotify or Deezer APIs, which could lead to broken features if not handled promptly. The generative models used by the platform will also require reg-

ular updates and retraining to stay relevant to emerging musical styles and user preferences. Another follow-up problem is related to usability: feedback from users might reveal unforeseen pain points or desired improvements. Addressing these concerns will require an iterative development approach, incorporating user feedback into future versions of the platform to enhance usability and performance.

Insert your content here.

20.6 Follow-Up Problems

Insert your content here.

21 Tasks

- Revise requirements document
- Design the systems structure
- Build prototype and get feedback
- Integrate Spotify API for song data
- Develop music recommendation module
- Work on music generation module
- Create song analysis moudle
- Test the platform and gather feedback
- Revise based on feedback
- Write user guide and documentation
- Finalize revisions to documentation

22 Migration to the New Product

22.1 Requirements for Migration to the New Product

There are no migration requirements as this project is not a replacement or upgrade of a previous project

22.2 Data That Has to be Modified or Translated for the New System

Similarly, there currently is no data that needs to be modified

23 Costs

The monetary cost estimate of the project is $0 CAD. All of the necessary equipment is owned by at least one group member.

The total time cost estimate of the project is 8 months (September 2024 - April 2025).

The function point cost estimate is 12. This is derived from the above sections, mainly the functional requirements, non-functional requirements, and business rules

24 User Documentation and Training

24.1 User Documentation Requirements

The music featurization feature is heavily API-driven, and as such, detailed documentation will primarily be covered within the API reference section. This approach ensures that developers and advanced users can understand the feature's capabilities without needing additional user guides.

To ensure that users can effectively interact with the music generation and recommendation platform, the following user documentation will be provided:

- **Quick Start Guide**: A concise guide aimed at helping new users get started with the basics of generating and recommending music.

- **API Reference and Technical Specifications**: Detailed documentation of the platform's API, including available endpoints, request/response formats, and example queries. This reference is crucial for developers and advanced users who want to integrate the platform with other applications or automate tasks.
- **Installation Guide**: A step-by-step guide for installing the platform on local servers, including system requirements, installation commands, and troubleshooting common setup issues.
- **FAQs and Troubleshooting**: A list of frequently asked questions and troubleshooting tips to help users solve common issues independently.
- **Video Tutorials**: Step-by-step video guides that visually demonstrate key features and workflows, including setting up the platform, using the API, and generating music.

These documents will be designed for users of varying technical backgrounds to ensure they can fully utilize the platform's capabilities. The documentation will be created and maintained primarily by the development team, ensuring accuracy and alignment with the latest platform features. However, feedback from user groups will be actively sought to improve clarity and address any documentation gaps. Updates to the API reference and technical specifications will be managed as part of the regular software update cycle.

24.2 Training Requirements

To provide users with sufficient knowledge to operate the platform effectively, the following training resources will be developed:

- **Video Tutorials**: Developed by the development team, these tutorials will cover various aspects of the platform, including API usage, generating music, and using advanced features.
- **Online Training Modules**: If additional resources become available, online training modules could be developed to provide users with a structured learning path. However, due to current resource constraints, we do not plan to offer live training sessions.

These training requirements aim to encourage users to explore the full potential of the platform, regardless of their prior experience in music production or technology.

25 Waiting Room

1. The recommendation system will be able to recommend songs from less popular music genres (jazz, blues, etc.)
2. The analysis system will be able to extract musical features from less popular music genres (jazz, blues, etc.)
3. The generative system will be able to generate songs from less popular music genres (jazz, blues, etc.)
4. The recommendation system will be able to recommend songs from unpopular music genres (gnawa, libyan funk, etc.)
5. The analysis system will be able to extract musical features from less popular music genres (gnawa, libyan funk, etc.)
6. The generative system will be able to generate songs from less popular music genres (gnawa, libyan funk, etc.)
7. The recommendation system will be able to search for songs with cover art similar to an input song
8. The generative system will be able to generate new cover art for a newly generated song, based on the user's input criteria
9. The generative system will be able to generate new covert art for an existing song

26 Ideas for Solution

- **Hybrid Recommendation System**: A hybrid recommendation system combines content-based filtering and collaborative filtering techniques to provide a more personalized experience for users. Content-based filtering analyzes song features, such as genre, key, and rhythm,

to suggest similar tracks. Collaborative filtering uses user preferences and historical listening patterns to suggest music. By combining these approaches, the system can offer users personalized suggestions while also helping them discover new genres and music styles.

- **Generative Music Model**: To enable the creation of new music, a generative model will be used. This model could be based on techniques such as a Generative Adversarial Network (GAN) or Recurrent Neural Network (RNN). A GAN would allow for the generation of realistic music by having the generator and discriminator work together to produce convincing compositions. An RNN, on the other hand, would be well-suited for learning the sequential nature of music, generating new melodies based on learned patterns. This solution provides users with an innovative way to create new music based on their inputs and preferences.

- **Feature Manipulation Interface**: This interface will allow users to interact directly with song features, such as tempo, key, and rhythm, enabling them to create customized versions of existing tracks or generate entirely new compositions. By adjusting different musical parameters, users can personalize their musical experience and experiment with creative variations, providing a high level of control over the output.

- **Integration with Existing Platforms**: Integrating the system with existing music platforms, such as Spotify, will allow users to easily access and analyze a large library of songs. Users will be able to input their favorite tracks from these platforms and generate variations or receive recommendations. This integration ensures a smooth user experience, allowing seamless interaction between existing music libraries and the platform's generative capabilities.

Appendix — Reflection

The information in this section will be used to evaluate the team members on the graduate attribute of Lifelong Learning. Please answer the following questions:

1. What knowledge and skills will the team collectively need to acquire to successfully complete this capstone project? Examples of possible knowledge to acquire include domain specific knowledge from the domain of your application, or software engineering knowledge, mechatronics knowledge or computer science knowledge. Skills may be related to technology, or writing, or presentation, or team management, etc. You should look to identify at least one item for each team member.

   As a team, the Rhythm Rangers, we need to acquire a diverse range of knowledge and skills from various domains, including software development, music generation, and collaborative teamwork, to successfully complete our capstone project. Given the scope of this task, it is essential for each team member to focus on specific areas of expertise that align with their skills, passions, roles, and responsibilities, as well as learn new skills and gain new knowledge. Outlined below is the knowledge and skills the team will collectively need to acquire to successfully complete this capstone project:

   Music Analysis and Signal Processing: This capstone project involves developing expertise in audio signal processing to analyze sound data and extract valuable insights for music recommendation and generation systems. The team will learn to implement machine learning models for tasks such as genre classification and feature extraction. Proficiency in Python libraries for audio analysis and model training is essential. This will deepen the teams understanding of music theory and the connections between song features and genres.

   Frontend or Backend Development: The team will need to understand backend frameworks for building and managing the recommendation system's infrastructure. The will be integrating external APIs to access song previews and features. They will also gain knowledge in database management for storing and organizing song data and

user preferences. Furthermore, this involves learning how to scale and efficiently handle data for a local server-based system.

UI/UX and Design: The team will need to design user-friendly interfaces that ensure smooth interaction with the music recommendation and generation systems. UI/UX design skills will need refinement and utilization of frontend development frameworks will be needed to craft the systems user interface. They will also learn to connect frontend components with backend APIs for real-time updates, such as delivering song recommendations.

Music Generation: An understanding of generative models to create music snippets from input tracks or references will be a huge component for this project. The team will delve into music feature engineering, transforming audio data into usable features for machine learning applications. Familiarity with music data will assist in generating new content and making recommendations.

Team Management and Infrastructure: For this project to be a success improving project management and team coordination skills will foster effective communication, sprint planning, and task assignment. We will need to learn how to establish and maintain local server infrastructure for efficient hosting and operation of the platform. Understanding security best practices to safeguard user data and ensure the system's resilience against vulnerabilities is critical. Mastery of version control and Git management will promote seamless collaboration and code review among the team.

2. For each of the knowledge areas and skills identified in the previous question, what are at least two approaches to acquiring the knowledge or mastering the skill? Of the identified approaches, which will each team member pursue, and why did they make this choice?