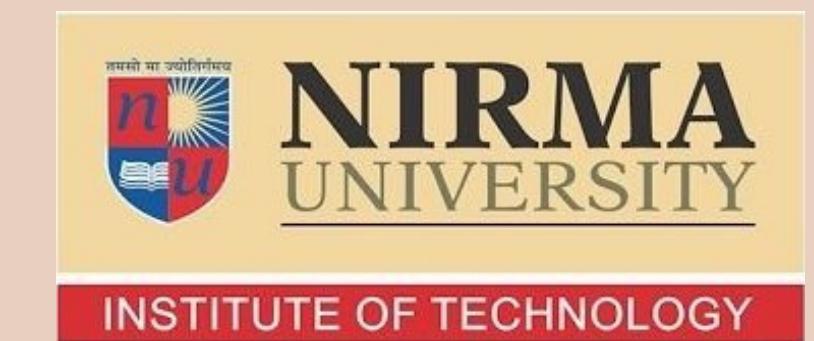


INFORMATION RETRIEVAL SYSTEM



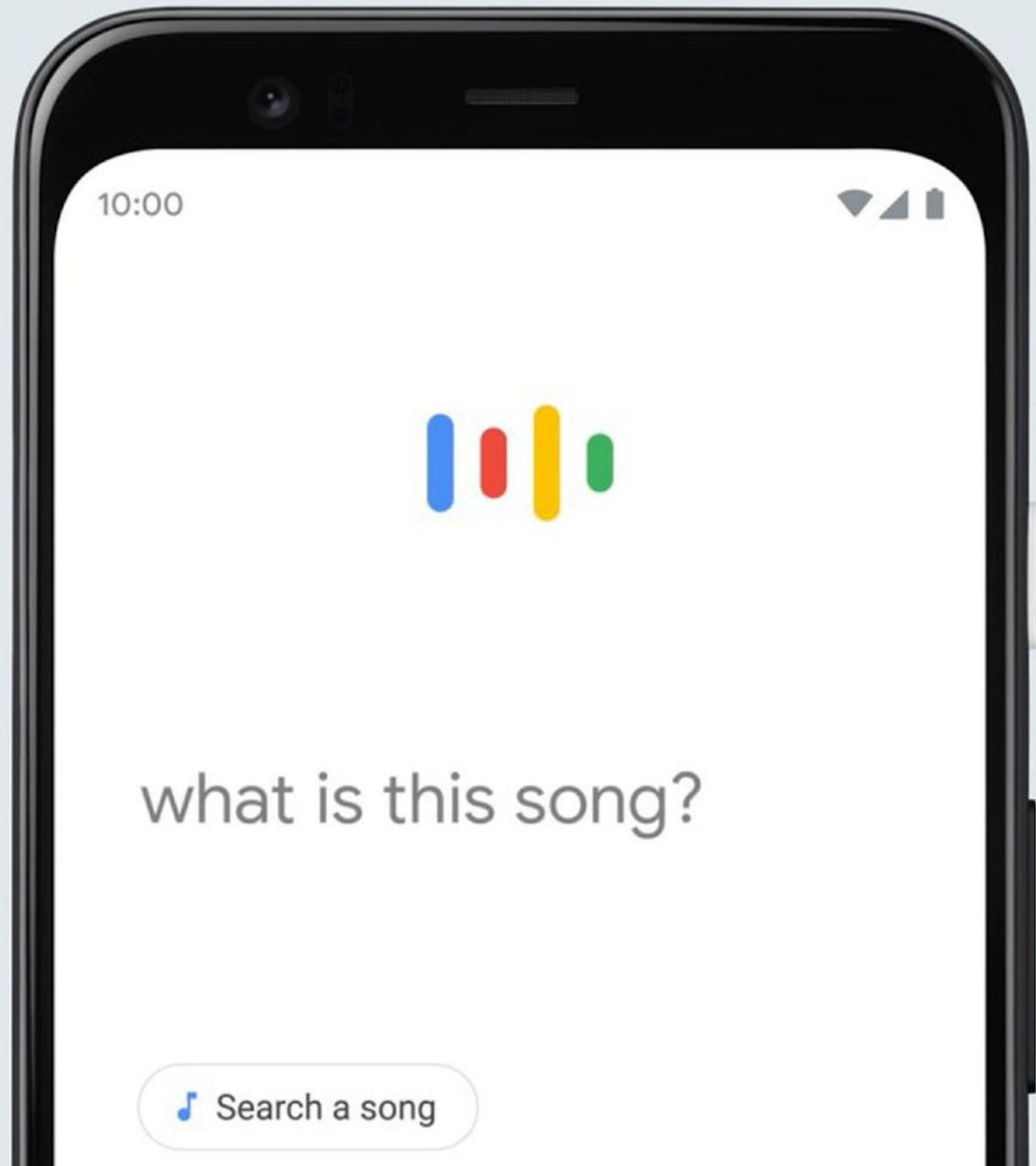
# QUERY BY HUMMING INNOVATIVE ASSIGNMENT

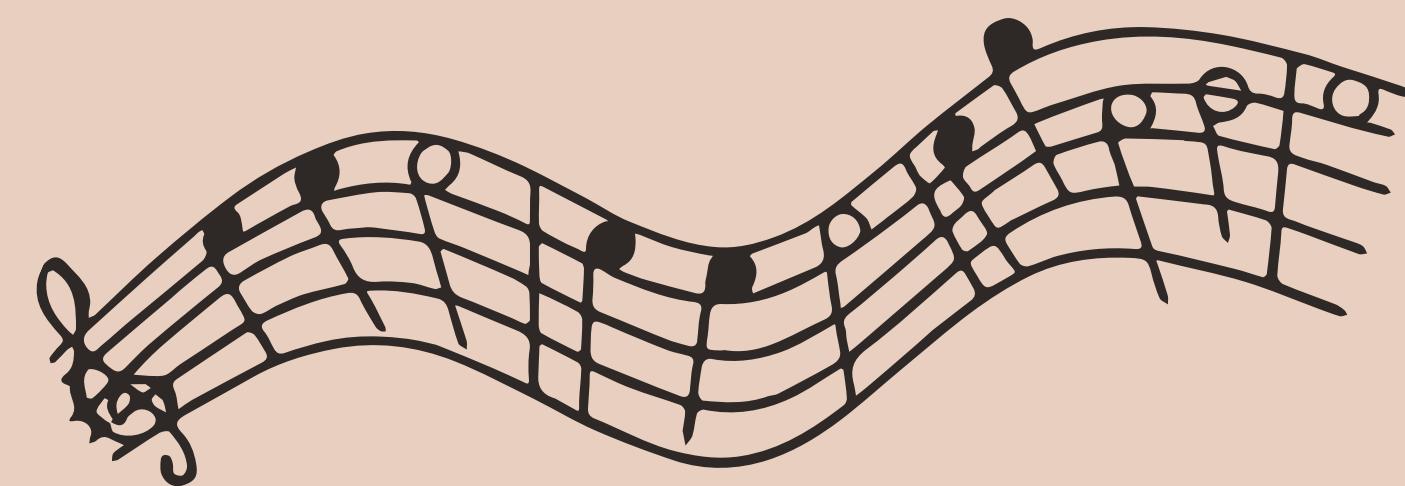
---



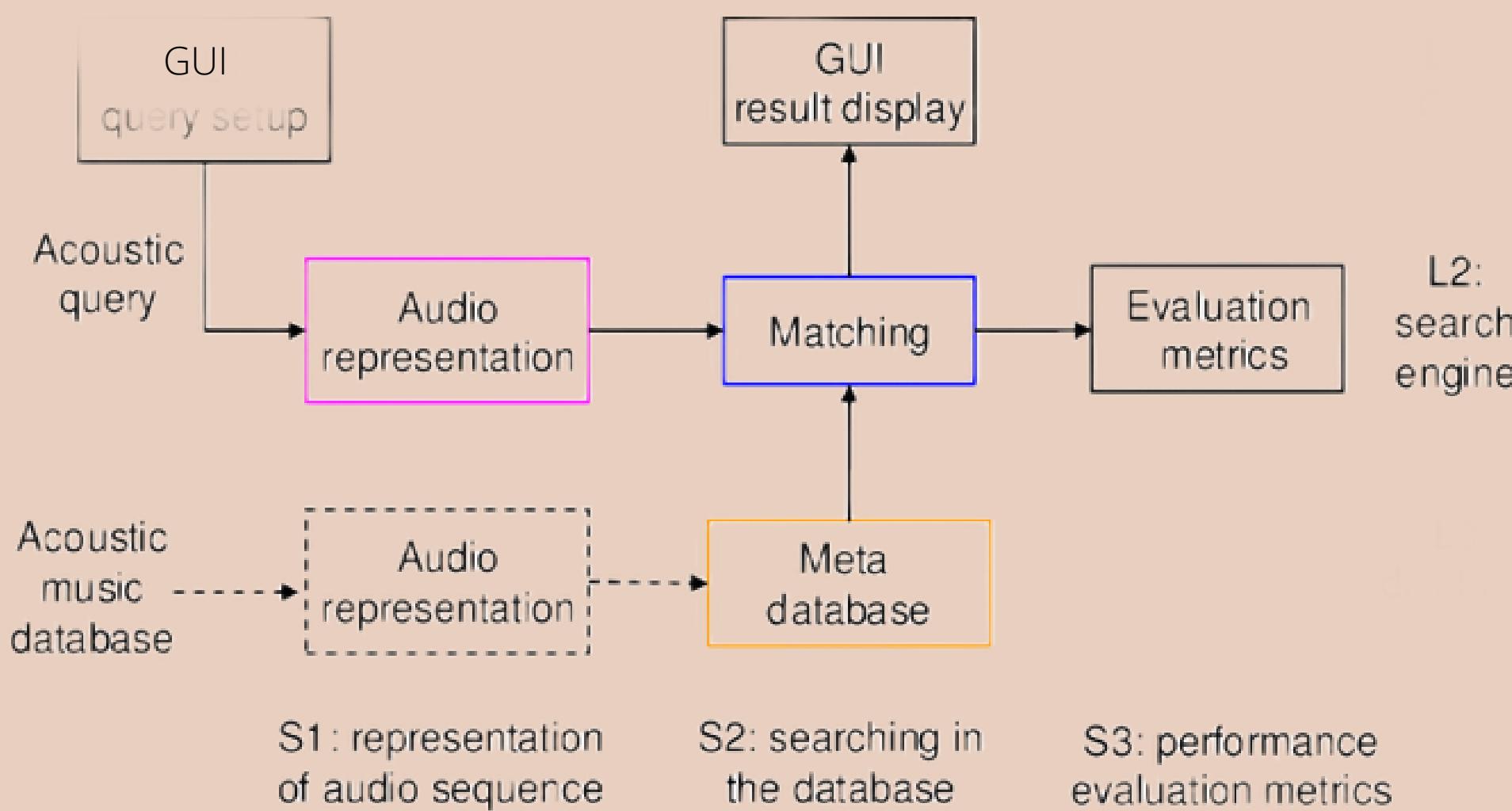
# WHAT IS QBH?

- Query by humming (QbH) is a music retrieval technique that deviates from traditional classification systems such as title, artist, composer, and genre. It usually refers to songs or other pieces of music with a single topic or tune.
- The system compares a user-hummed tune (input query) to an existing database. After that, the system delivers a ranked list of music that comes closest to the supplied query.





# MUSIC INFORMATION RETRIEVAL



- Music Information Retrieval (MIR) is a particularly interesting research subject.
- It's critical to select a match between the input query and the matching music while retrieving music.
- Query-by-Example (QBE), which takes a sample of an audio recording playing in the background and returns the result, is one of the methods that has been presented and is presently being used in the various application domain(s). In a Query-by-Humming (QBH) application, however, there is no effective way to tackle this problem. The goal of a Query-by-Humming application is to quickly obtain music that is most comparable to the hummed query.

# TECHNIQUES FOR MUSIC INFORMATION RETRIEVAL



## QUERY BY TEXT

Searches for similarities between songs using conceptual metadata such as text queries  
example : Spotify



## QUERY BY HUMMING

provide a score to reflect similarity in melody between the natural humming voice emitted by humans and the original artist's singing.



## QUERY BY EXAMPLE

takes a segment of the original music recording and searches the database for the most similar song  
example : shazam

# RELATED WORK

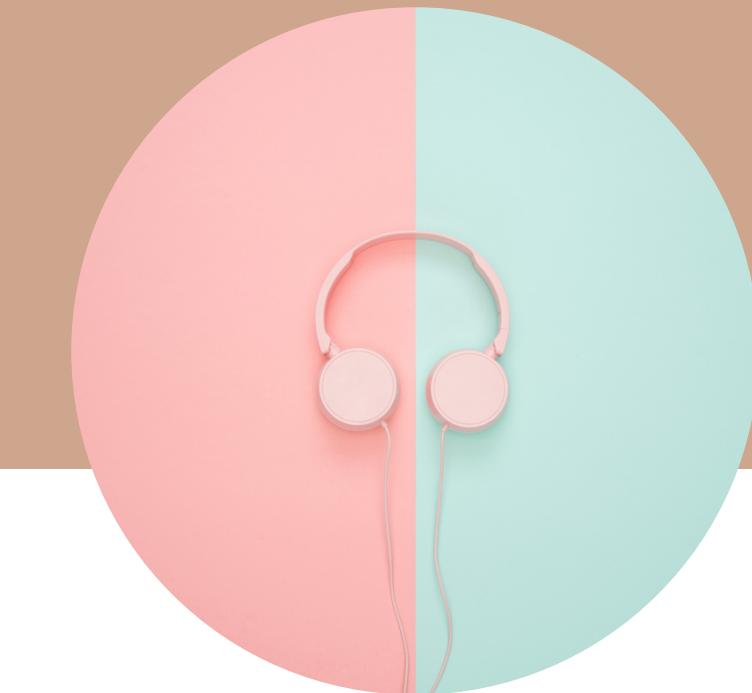
- Music Information Retrieval System Architectures
- Speech Feature Extraction
- Audio Fingerprinting
- Machine Learning-Based Music retrieval
- Noise Reduction Techniques



# MUSIC INFORMATION RETRIEVAL SYSTEM ARCHITECTURES



To retrieve music effectively through a system, one needs to follow the architectures structured of MIR system.



One can use a time-series data mapping algorithm named Dynamic Time Warping (DTW) approach.



This approach is built to be reliable and possess the potential to achieve greater retrieval accuracy.



## Speech Feature Extraction

Extraction of speech features is responsible for the conversion of speech signals into a sequence of feature vector coefficients that only entails information needed to identify a specific syllable.

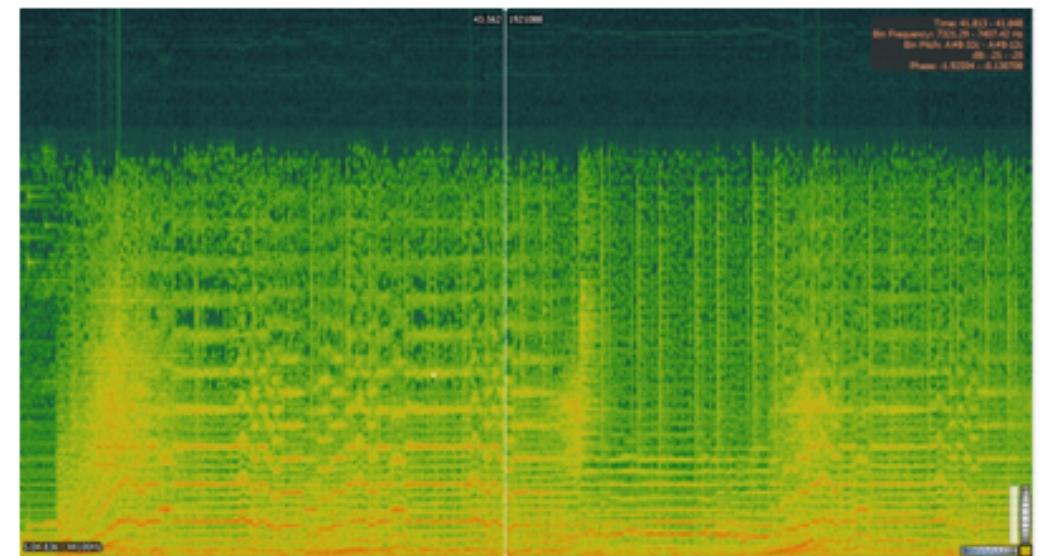
Spectral analysis techniques like Mel-frequency cepstral coefficients (MFCCs) are commonly used to extract the feature vector of voice signal.

# AUDIO FINGERPRINTING



- The technique of extracting essential information from an audio stream in order to express it in a compact manner is known as **audio fingerprinting**. It is based on the concept of human fingerprinting and records the fingerprint of ingested audio content, which may later be compared to previously recorded audio or playlists on a phone, television, or other device.
- A robust acoustic fingerprinting identifies the audio file even after compression and sound quality loss.

Frequency  
↑



Time/ Frames ➔

# ML BASED MUSIC RETRIEVAL SYSTEM

ML can be used to train a neural network to perform tasks that would normally take a long time to complete. N. Mostafa et al. envisaged a Deep Neural Network (DNN)-based note-transcription method that uses hummed notes as features to train the neural network.

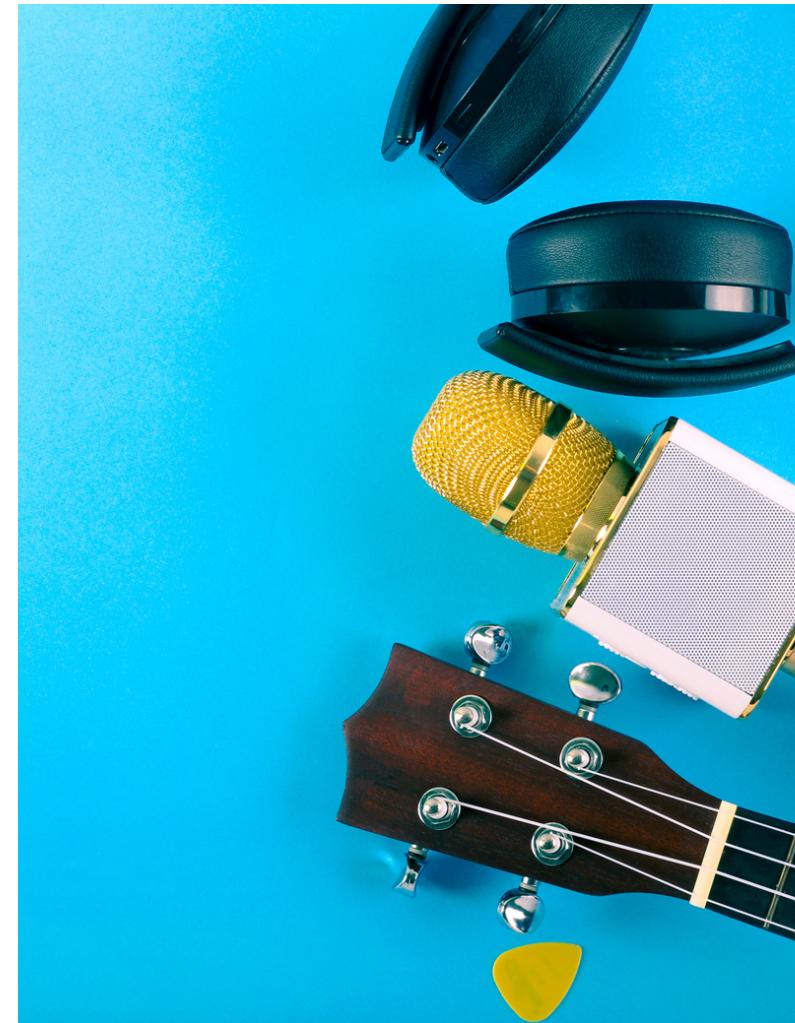
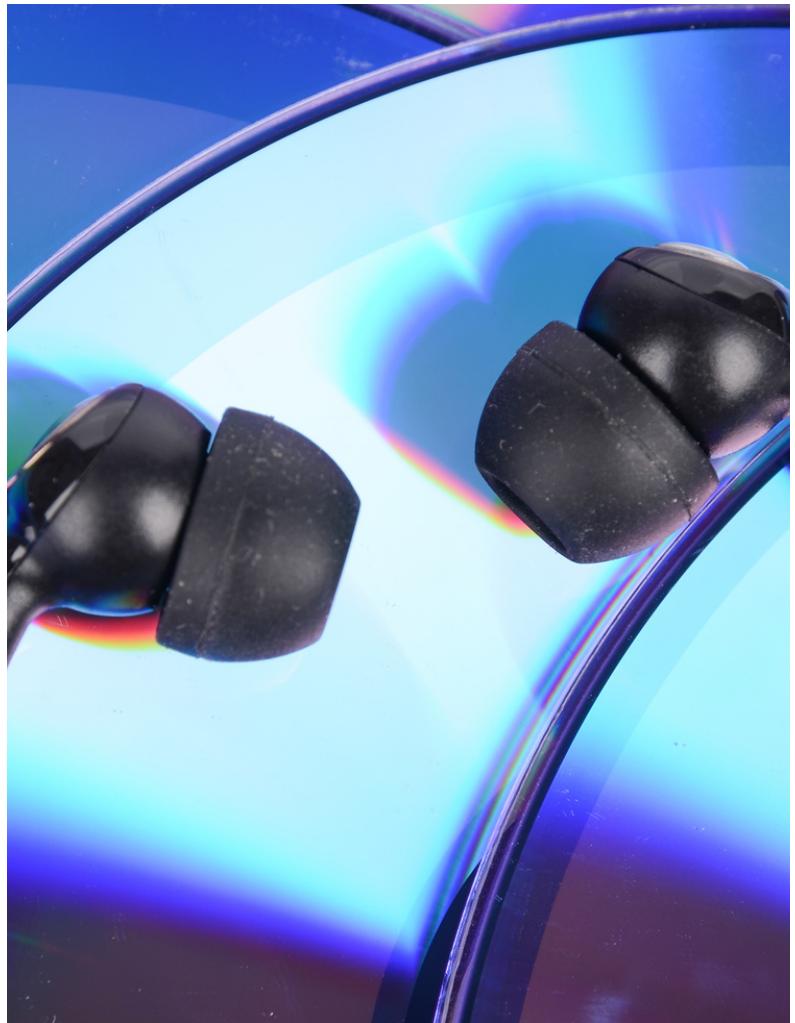
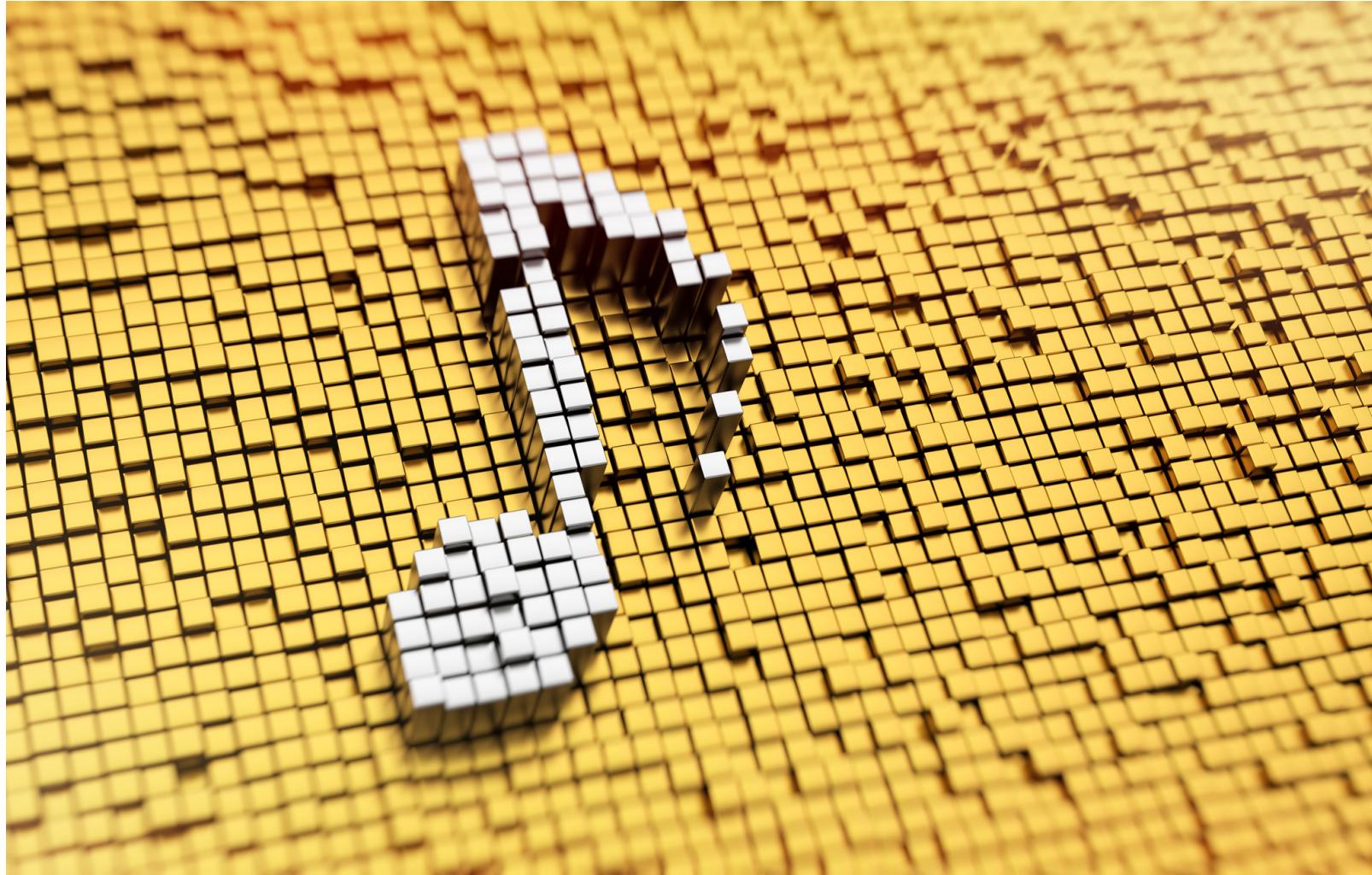
The accuracy of Mean Reciprocal Rank (MRR) reached to 0.7679 and when DNN-HMM-based acoustic model was used, the accuracy obtained was 0.8071. The authors deduced that the accuracy score of transcription and retrieval system might increase when the DNN is trained on a larger dataset.



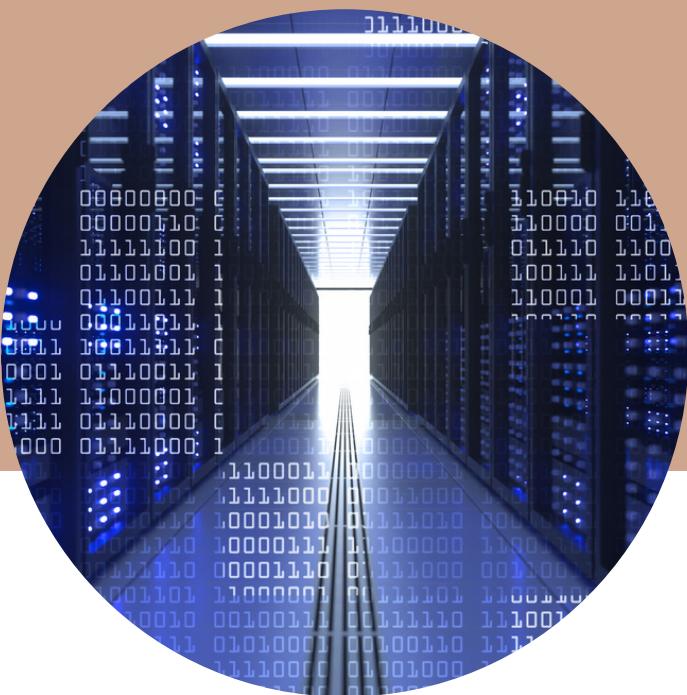
# NOISE REDUCTION

Noise reduction techniques are constantly used in the field of Natural Language Processing. Using well-known approaches, ambient noise can be reduced when audio input is coming from the microphone. For editing audio and removing noise, open-source software known as 'Audacity' can be utilised. In QBH systems, on the other hand, this can be used to 'clean' the data before training the neural network.

Active Noise Reduction (ANR) is a way for controlling or reducing noise using live input. It is a method for reducing unwanted sound by adding a sound that is specifically designed to cancel out the noise. Because the female humming frequency range is 350 Hz to 17 kHz, this feature is effective for QBH systems.



# FUTURE WORK



**ML DL model**  
models that may be trained  
on past humming data to  
recognise new/unknown  
hummed requests and the  
music associated with them.

**Active Noise Reduction**  
pass through an ANR module  
to decrease noise in live  
input. The ANR module's  
output would be a crisp  
humming sound.



**Database Indexing**  
proposed system can be improved  
by devising an efficient technique,  
such as optimising database  
indexing. Indexes are a terrific way  
to easily organise and locate data.





**THANK  
YOU**

- Aayush Shah (19BCE245)
- Shivam Panchal (19BCE150)
- Priyal Palkhiwala (19BCE214)