Database Research

2/27: https://www.youtube.com/watch?v=QPLYTFgTQIg

What type of sound files are we using? .wav
We need to import wave in python in order to first take in a file.
#Import audio file as wave object
File_name = wave.open("file_name.wav", "r");
#Convert wave to bytes
File_name_soundwave = file_name.readframes(-1)
*-1 lets us take in all information from the audio file.
#To test if the convert worked
File_name_soundwave

How in depth do we want to go? Do we want to have the database store information like starting location, flight name, pilot name, etc.?

Baseline information like file name and type.

What type of database system do we want to use?

MySQL: https://dev.mysql.com/doc/refman/8.0/en/what-is-mysql.html
Yes, can be used with python.
PostgreSQL: https://www.postgresql.org/about/
Can be used with python
SQLAlchemy: https://www.sqlalchemy.org/
Can be used with python

MySQL vs. PostgreSQL: https://developer.okta.com/blog/2019/07/19/mysql-vs-postgres
MySQL has faster readability.
PostgreSQL handles large amounts of data more efficiently
PostgreSQL handle concurrency better.
PostgreSQL follows SQL standards better

MySQL vs SQLAlchemy: https://stackshare.io/stackups/mysql-vs-sqlalchemy
It doesn't really make a difference, SQLAlchemy works on top of something like MySQL.

SQLAlchemy vs PostgreSQL: https://www.reddit.com/r/learnpython/comments/slkjz/flask_sqlalchemy_or_straight_mysql/c4f1kum/
It is the same situation as MySQL vs SQLAlchemy, SQLAlchemy can work ontop of PostgreSQL.

What is concurrency?
The ability for different parts or units of a program to be executed out of order or in a set order without affecting the outcome.

What does SQL mean?
Structured Query Language; computer language designed for eliciting information from databases.

What are SQL standards?
https://www.oreilly.com/library/view/sql-for-dummies/9780470557419/ch03.html

What is a directory?
A file catalog structure that references computer files and other directories. Does not allow for files to be changed.

What is a database?
Stores data in either a rational/structured way or non-rational/unstructured way. Allows for changes to be made to files in the database.
Do we want a directory which would store all files types; can be searched by file type but not contents.

If so we then need a directory that has access to a database. A database organizes data and allows for things to be added to a file, updated, etc.

Directory Creation: https://realpython.com/working-with-files-in-python/#making-directories

- When creating a directory if the directory already exists there will be a FileExistsError.
- Can not use .glob() since part of the team uses Windows OS and it does not work with Windows.
- os.walk() generates directory tree and allows for navigation/manipulating the directory
- os.walk() returns current folder name -> list of folders within the main folder -> list of files in the main folder
- There is a way to read from multiple files
    - Might only work for text files
- There is a way to read zip files.
    - Allows for easier testability and potentially better functionality for the user
- What is a temporary file and what is it used for?
    - https://www.computerhope.com/jargon/t/tempfile.htm
    - Created to hold information while a file is being created or modified; once the program is closed the file is deleted.
- Is there a difference in using directory creation using passing of parameters or by just using the pathlib?
    - https://docs.python.org/3/library/pathlib.html

Database Creation:
*We can not create a database without the paid version of pycharm.
https://towardsdatascience.com/how-to-build-a-database-using-python-f4b62a19d190

- Redundancy: When values are repeated several times
- Normalization: Having one table of a lot of data split into separate tables

Cloud interaction:https://cloud.google.com/speech-to-text/docs/samples/speech-transcribe-async-gcs#speech_transcribe_async_gcs-python

Directory to Cloud Communication: https://support.google.com/a/answer/106368?hl=en
https://cloud.google.com/storage/docs/gsutil/commands/cp

SIDE THOUGHT

Program takes in user request -> checks directory for file type -> check database to see if that file already exists -> if so check directory for other file type(audio/text) -> check database for new file type with the same name -> if found no need to take time to run through API(s), so return correct file type.

*This could be its own function or module that interacts between the gui code and the directory/database code

Program takes set file name keywords from user -> checks directory for keywords of set file type…