



1. [About](#)

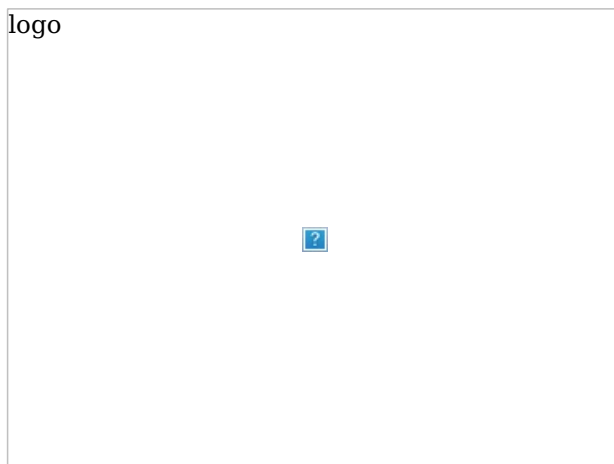


[Ampli-FIRE](#)

- 
- [Project Charta](#)
- [Data Report](#)
- [Modelling Report](#)
- [Evaluation](#)

## Other Formats

- [PDF](#)



## About

This project has led us through many ups and downs. For starters, when the topic was first suggested to us it was nothing more than a concept; an intellectually stimulating idea with no guidelines. We spent much of the first few days brainstorming the specifics of what we wanted to do and trying to map out the procedures we would need to follow in order to achieve them. Once we had established what we wanted to do, we moved onto the task of gathering data.

## Data Acquisition & Cleaning

With nearly an infinite number of songs, gathering the data we needed was a huge undertaking. Not only did we first need to find sites to scrape, but we also had to ensure that they had the information that we needed for every song possible. The parameters of the project allowed us to narrow our focus to only sites that categorized their music based on genre and artist, as these are two of the main attributes needed for our recommendation services. This is where we came across our next obstacles: website security. Many of the music streaming services that we had hoped to use had many safeguards that prevented us from using Scrapy in the way that we had been taught. Many hours were sacrificed trying to get around these obstacles. Fortunately, with the help of Professor Ira Woodring of Grand Valley State University we were able to find a site that had an extensive list of songs with all of the parameters we would need for the project. That site is [everynoise.com](http://everynoise.com).

In the end, our team was able to scrape all 584,921 songs from the 6,291 genres provided by the website, but this was not quite enough. See, the problem with Everynoise is that the songs it gave us could not be downloaded without violating copyright laws. This is a problem as we would have to rely on those audios in order to create the fingerprints needed in the deep learning section of the project. So, after a lot of trial and error, we settled for [Jamendo](#), which offered us a sizeable variety of open source songs.

## References