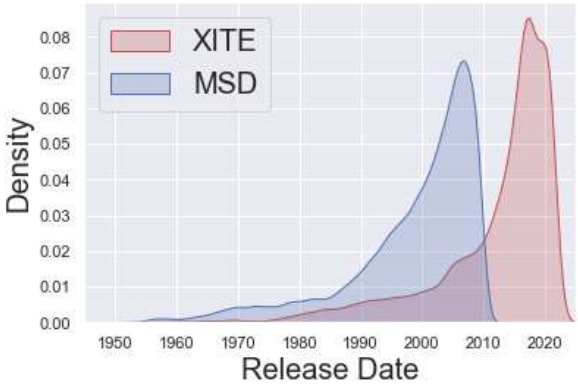
A EXPLORING THE METADATA

This appendix is meant to characterize the metadata of the private dataset. It aims to describe some of the features in the dataset, and compare it to a publicly available dataset.

MSD Comparison. There are currently 97.241 videos published to the music video platform. In addition to the audio and video data itself, the dataset contains metadata concerning the artist, genre, sub-genre, release date, origin and duration of each music video. Additionally, the music video platform keeps data on three more subjective, manually tagged features: mood, energy and urgency. Since this a private dataset, we will describe some of the characteristics of the dataset, and compare it to a public dataset that it popular in the field of music information retrieval. Specifically, we will compare the dataset to the MSD [4]. This dataset consists of roughly one million songs with metadata and audio features. Additionally, roughly 240.000 of the songs include manually labelled tags, which include features like genre ('Pop', 'Rock'), decade (70s, 90s), origin ('British', 'Dutch') and general descriptions ('Female', 'Bass', 'Acoustic'). Of course one main important difference between the datasets is that the private dataset concerns music videos, whereas the MSD just contains songs. Although there is one paper that introduces a music video dataset [46], this just contains features based on the audio and video data, as opposed to the actual signals themselves. To the best of our knowledge, there is no publicly available academic music video dataset.

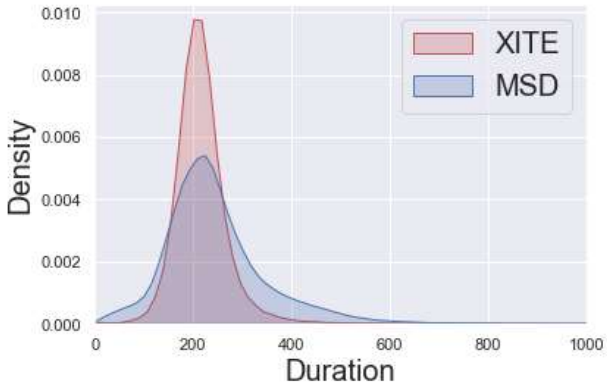
Release date. Figure 6 provides the empirical distributions of release date in both datasets. The distribution of release dates in the private dataset is similar to the distribution in the MSD, but shifted to the right. This is due to the fact that the music video app is constantly getting updated with new music videos, whereas the MSD has remained the same since its release in 2011.

Figure 6: Density plot of release dates for the MSD and the private dataset.



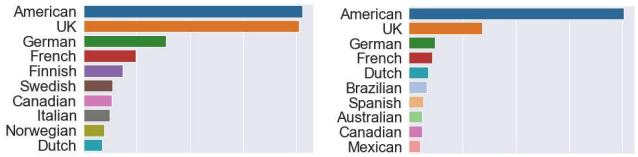
Duration. In terms of the average duration of songs, the two datasets are similar, although figure 7 shows that the MSD has more variation in song duration, whereas songs in the private dataset show less variation in their duration. This could potentially be because of the fact that the private dataset only consists of music videos, as opposed to songs. A possible explanation is that music videos generally have less variation in their length than songs, but we did not test this hypothesis.

Figure 7: Density plot of duration for the MSD and the private dataset.



Origin. We also compare the datasets in terms of the music origin. For the private dataset, getting a distribution of origin is easy, as it is one of the variables in the dataset, but for the MSD it is slightly more complicated. The MSD does contain information about origin, but it is part of the MusicBrainz tags, so there is no separate category for origin. Note that tags are on artist level, and not on song level. To get an estimate of the distribution of origin in the MSD, we selected the top 10 MusicBrainz tags that were concerned with countries. However, there were various duplicates in the tags. For example, there are separate tags for 'UK' and 'British', and for 'Canada' and 'Canadian'. In these cases, we decided to count only the most occurring tag. Note that aggregating two tags with the same meaning would not have the desired effect, as a single artist often has both tags. Therefore, this would lead to a distribution that is biased towards countries with multiple tags. Figure 8 shows that the majority of songs in both datasets comes from either the United States or Britain. The remaining countries that appear often are mostly western countries, although Brazil and Mexico do appear in the top ten countries of the private dataset.

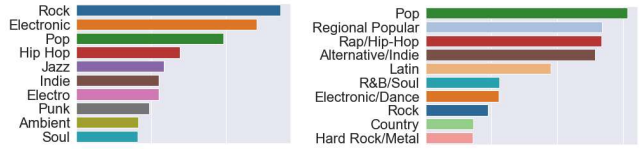
Figure 8: Empirical distribution of the top ten origin countries in the MSD and private dataset.



(a) MSD (b) Private

Genre. To get an estimate of the distribution of genre, we take a similar approach. For the XITE dataset, we simply use the genre feature, and for the MSD, we look at the Echo Nest tags, filtering out tags that are not related to genre. As we did for origin, we filter out tags that overlap in meaning, such as 'Rock' and 'Alternative Rock', because there are no 'Alternative Rock' songs without the 'Rock' tag. In these cases we kept the most common tag. Figure 9 shows that Hip Hop and Indie are more common on the music video platform, whereas Electronic and Jazz are much more frequent in the MSD.

Figure 9: Empirical distribution of the top ten genres in the MSD and private dataset.



(a) MSD (b) Private