The Demo App

Demo Slovak Music Database

Modern recommendation systems as well as the use and monitoring of local content guidelines (‘radio quotas’) require reference data about recorded music that will be played or recommended to be played, and metadata, or ‘data about the data.’ Slovakia currently does not have a database that could help the monitoring of radio quotas, or which could help the LaLa music export, or a Slovak label to find audience for its songs.

Our Demo Slovak Music Database shows how to fill this need. It is used in our Listen Slovak Demo App, and we show in our Feasibility Study what procedures and data is needed to create and constantly update a Comprehensive Slovak Music Database for these purposes.

The demo database contains information about Slovak artists, composers, and their sound recordings released in Slovakia and abroad. It can form the basis of radio quota monitoring, new type of recommendations, educations applications, or just simply improving the likelihood that Spotify or YouTube will recommend a particular Slovak sound recording or video.

The entire database can be accessed as SQLite database, or in the form of .csv, .xlsx or .rds tables. Please contact for help if you want to try it out our use it – it is free, open access, and we are happy to assist any interested reader to try it out. We also placed on this website a summary of the table, so that you can review which artists are featured in the database.

*This will be seen in the front page, and the project description phase the rest. Therefore, add this this text:*

*We added an easy to browse and download simplified version of the Demo Slovak Music Database to the website [here] – artists and their representatives can review how we see them or ask to be added or removed. You can read [there] a bit more about the database and its use, and you can use the [Contact] options to get access and help to data database in several formats that can be used with free of charge, open source software.*

## Artists

### Opt-in Process

The opt-in process was based on a voluntary questionnaire that SOZA and Reprex sent to various Slovak artists and labels. We asked age, performance maturity, language use and location specific questions about the artists. The opt-in procedure is still open – we are happy to include more artists into the demo database and demo application, and we hope that we can continue this project with the Slovak Arts Council and other partners to make the database truly comprehensive, i.e. to contain all Slovak music unless the artist wants to be featured as ‘Slovak.’ Want to be deleted? The opt-out process is very simple in the demo phase: contact us and let us know that you do not want to be part of the Demo Slovak Music Database.

### Write-in Process

The write-in process was necessarily subjective, because we could not, and did not want to use GPDR protected personal information. The current broadcasting quota defines Slovak artists by residence, which is a personal data that we do not have, and which is not applicable for deceased Slovak artists.

As an alternative, we created a long list of candidates to the database based on listening habits in Slovakia. We considered artists who appeared on radio charts, or who appeared on various, Slovakia-specific playlists on Spotify, or who described their music with the ‘Slovak’ adjective, like ‘Slovak rock’ or ‘Slovak hip hop’. We read-in all the biographies of the English language Wikipedia in the Slovak female and Slovak male music performer categories.

We narrowed down this long-list with the help of a Slovak native-speaker musicologist, Dominika Semaňáková, to identify artists who should be ‘considered Slovak’. The criteria of inclusion were publicly available city affiliation (where the artist or group works), birthplace, place of death, language of lyrics when applicable, language of the song or work titles, and publicly available information about the artist.

In the demo version we tried to limit the data written-in, and we only concentrated to Spotify (explained here why.) The following features are attached to each artist’s Spotify ID: [don’t translate labels, these are identifiers in the database files if somebody downloads them.]

* popularity: A number between 0 and 100 describing the artist’s popularity based on stream counts on the Spotify platform.
* followers: An artist’s total number of subscribers on the Spotify platform.
* genre: Up to three genres associated with each artist.
* total\_reccomendations: The number of recommendations associated with each artist after three iterations of our algorithm. Artists with zero recommendations are practically ‘invisible’ for Spotify, and likely never recommended. One of our main goals is to avoid this ‘cold start’ problem, which affects about 15% of Slovak artists in 2020. [don’t translate cold start]
* genres: Up to three genres associated with each artist.
* any\_slovak\_release: A value of 1 is assigned if the artist has released some work in Slovakia, 0 otherwise.
* slovak\_release\_pct: The percentage of sound recordings released in Slovakia.
* any\_slovak\_title: A value of 1 is assigned if the artist has some work with a title in Slovakian, 0 otherwise.
* all\_slovak\_title: A value of 1 is assigned if all the artist’s work is in the Slovak language, 0 otherwise.
* all\_english\_title: A value of 1 is assigned if all the artist’s work is in Slovak language, 0 otherwise.
* known\_slovak\_city: Slovakian cities associated with the artist.
* considered\_czech: A value of 1 is assigned if the artist is also considered to be Czech. Our greatest challenge is to define the boundaries of Slovakness and being “Czech”.
* popular\_song\_titles: A short selection of the artist’s most popular songs.
* latest\_popular\_releases: The date of the artist’s most recent popular releases.
* spotify\_url: The artist’s or bands’ official Spotify URL.
* Official\_website\_url: The artist’s or band’s official website.

## Sound Recordings

The Demo Slovak Music Database contains up to ten sound recordings associated with 0 Slovakian artists, comprising a total of 3937 unique tracks. When the Spotify API is queried for top track recommendations, 19 artists do not yield any results. For each other artist, the following features are included:[[1]](#footnote-1)

For each track, the following features are included [Footnote 2: A full list of the data Spotify keeps for each track can be consulted at <https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/>]:

* danceability: An approximation to how suitable a track is for dancing, considering musical elements such as tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable.
* energy: A perceptual measure of intensity and activity aggregating features including dynamic range, perceived loudness, timbre, onset rate, and general entropy.
* key: An approximate overall key for the track. Spotify expresses this in pitch notation, and here this is converted to the notation’s tonal counterparts.
* loudness: Average loudness of a track in decibels (dB).
* mode: An approximation to a major-minor modality: 0 is minor, 1 is major.
* speechiness: Detects the presence of spoken words in a track on a scale between 0 and 1.0.
* acousticness: A confidence measure from 0.0 to 1.0 of whether the track is acoustic. Values above 0.5 are intended to represent instrumental tracks, but confidence is higher as the value approaches 1.0.
* liveness: A confidence measure from 0.0 to 1.0 of whether the track is live depending on the detection of an audience in the recording.
* valence: An approximate measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry).
* tempo: The overall estimated tempo of a track in beats per minute (BPM).
* duration (ms): Track length in milli-seconds.
* time\_signature: An overall approximation to the track’s beats per measure.

### Genres

We all know that genres are very important for music discovery and music selection, but music cannot be very well described with natural language words. Any taxonomy of genres is dependent on the language, for example, ‘jazz’ may mean a bit different for music fans in Slovakia, Japan, and the United States.

Digital streaming services and media platforms usually use AI applications that understand English language taxonomies. Therefore, it is critically important to have a correct understanding and use of the Slovak and English versions of genre descriptions. A song that is accidentally labelled as ‘rock’ instead of ‘pop’ may become invisible, because it does not fit the ‘rock’ audience’s taste. It will be skipped by many listeners, and the AI algorithm will ‘learn’ that this is not a good song. Even though ‘pop’ audiences may like it.

Our databases genre table map abstract distances of ‘Slovak’ genres, such as “Slovak indie” or “Slovak jazz” to “Belgian indie” or “Japanese jazz”. This is a very important part of our database, and it requires a broader consensus among radio editors, music export promoters, labels, artist, and computational musicologists to get this right. However, or [opt-in process]( reference to earlier part of the text) show that the lack of knowledge about the way genres work in global platforms is an important reason of low visibility for Slovak music.

1. A full list of the data Spotify keeps for artists can be consulted at <https://developer.spotify.com/documentation/web-api/reference/artists/get-artist/> [↑](#footnote-ref-1)