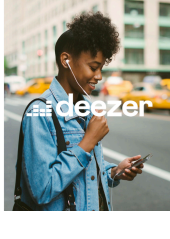# Multilingual Music Genre Embeddings for Effective Cross-lingual Music Item Annotation
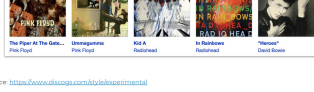
Elena V. Epure
Guillaume Salha
Romain Hennequin

research@deezer.com
https://github.com/deezer/MultilingualMusicGenreEmbedding

Context: Music Genre Annotations

Music genre annotations are essential in nowadays music-related services.

Source: https://www.discogs.com/bb/a/experimental

These are often used to:

- Create style-coherent music collections
- Enable genre-informed music recommendation and search
- Express user music profiles and preferences

Sources: http://www.deezer.com, https://www.last.fm
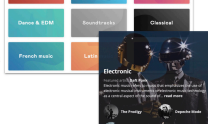
Extensive research has been conducted in classification / recognition of music genres:

- From **content**
  - Audio (e.g. Choi et al., 2016)
  - Lyrics (e.g. Mayer et al., 2008)

- From **metadata**
  - Reviews (e.g. Oramas et al., 2016)
  - Images (e.g. Oramas et al., 2017)

Limitations: Music Genre Recognition

In practice, there is **no unique** music genre representation & annotation practices (Bogdanov et al., 2019; Hennequin et al., 2018).

Music genre representations

- Could be defined by many sources and as taxonomies, vocabularies, folksonomies

- Examples of differences between them are:

Varying number of tags	500 tags in Discogs, 10K tag in English DBpedia
Varying music style relations	no relation vs. genre-subgenre vs. stylistic origin / derivative
Multiple tags for a music genre	"lo-fi", "lo-fi aesthetic", "bedroom music"
Polysemous tags	"hardcore" for "hardcore electronic" or "hardcore punk"

Music genre annotation practices

- Annotators with different backgrounds and musicological expertise such as crowdsourced, experts, users. It often results in low interrater agreement.
- Only a subset of existing music genres is use in practice, with a bias towards Western popular tags
- Often single-labeled datasets

Addressing Limitations: Music Genre Recognition

Bogdanov et al. (2019) proposed the **AcousticBrainz dataset**:

- Multi-source, multi-label, sub-genres, large-scale

Hennequin et al. (2018), Epure et al. (2019) proposed a **translation model**:

- Given music items annotated with genres from a source music genre taxonomy, the goal is to infer the equivalent annotations within a target music genre taxonomy.

Experimental, industrial rock ⇒ Experimental rock

Experimental rock ⇒ experimental last.fm

## Left Unaddressed: Multilingual Music Genres

Multilingual genre representations, especially when a parallel corpus is unavailable.

Hence, the goal is then to perform unsupervised **cross-lingual genre translation**.

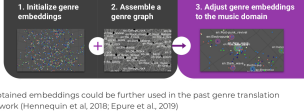However, this is not a translation in the regular sense because:

- Genre perception is culture-dependent
- Listening habits are culture-dependent (Ferwerda and Schedl, 2016; Skowron et al., 2017)

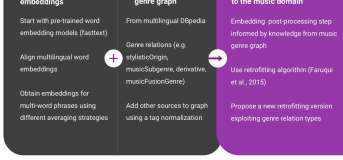Most depicted genres worldwide
Source: https://everynoise.com/
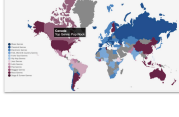
Contributions - Cross-lingual Genre Annotations

Model to learn **multilingual music genre embeddings** in 3 steps.

The obtained embeddings could be further used in the past genre translation framework (Hennequin et al, 2018, Epure et al, 2019)

In more details:

1. Initialize genre embeddings

Start with pre-trained word embedding models (fastText)

Align multilingual word embeddings

Obtain embeddings for multi-word phrases using different averaging strategies

2. Assemble a genre graph

From multilingual DBpedia

Genre relations (e.g. stylisticOrigin, musicSubgenre, derivative, musicFusionGenre)

Add other sources to graph using a tag normalization

3. Adjust genre embeddings to the music domain

Embedding post-processing step informed by knowledge from music genre graph

Use retrofitting algorithm (Faruqui et al., 2015)

Propose a new retrofitting version exploiting genre relation types

We curated a **parallel corpus of ≈50K cross-lingual annotations** from DBpedia

- Multilingual, multi-source, multi-label

Title	Type	🇬🇧 English (En)	🇫🇷 French (Fr)	🇪🇸 Spanish (Es)
Morning View	Album	Alternative_metal, Funk_rock, Alternative_rock, Post-grunge	Rock_alternatif	Metal_alternativo, Rock_experimental
Jimi Hendrix	Artist	Hard_rock, Psychedelic_rock, Blues, Rhythm_and_blues	Rock_psychédélï que Blues_rock, Hard_rock	Blues_rock, Rock_psicodélico, Hard_rock
Julio Iglesias	Artist	Dance-pop, Latin_music, Adult_contemporary_music	Pop_française	Pop_latino, Soft_rock, Balada_romántica, Adult_contemporary

We performed the **evaluation**:

- For 2 types of multi-source translation (English-language only and cross-lingual)
- With embeddings obtained with multiple model configurations

Music genre embeddings were proven **effective** in:

- Cross-lingual music genre translation
  - Macro AUC scores > 0.9
- Past English-language genre translation
  - Outperformed baseline (Epure et al, 2019) by wide margin

https://github.com/deezer/MultilingualMusicGenreEmbedding

References

- R. Mayer, R. Neumayer,and A. Rauber. Rhyme and style features for musical genre classification by song lyrics. ISMIR, 2008.
- K. Choi, G. Fazekas, and M. Sandler. Automatic tagging using deep convolutional neural networks. ISMIR, 2016.
- S. Oramas, O. Nieto, F. Barbieri, and X. Serra. Multilabel music genre classification from audio, text and images using deep features. ISMIR, 2017.
- S. Oramas, L. Espinosa-Anke, A. Lawlor, et al. Exploring customer reviews for music genre classification and evolutionary studies. ISMIR, 2016.
- D. Bogdanov, A. Porter, H. Schreiber, J. Urbano, and S. Oramas. The AcousticBrainz Genre Dataset: Multi-Source, Multi-Level, Multi-Label, and Large-Scale. ISMIR, 2019.
- R. Hennequin, J. Royo-Letelier, and M. Moussallam. Audio based disambiguation of music genre tags. ISMIR, 2018.
- E. V. Epure, A. Khlif, and R. Hennequin. Leveraging knowledge bases and parallel annotations for music genre translation". ISMIR, 2019.
- B. Ferwerda and M. Schedl. Investigating the relationship between diversity in music consumption behavior and cultural dimensions: A cross-country analysis. UMAP-SOAP, 2016.
- M. Skowron, F. Lemmerich, B. Ferwerda, and Markus Schedl. Predicting genre preferences from cultural and socio-economic factors for music retrieval. In ECIR, 2017.
- M. Faruqui, J. Dodge, S. K. Jauhar, C. Dyer, E. Hovy, and N. A. Smith. Retrofitting word vectors to semantic lexicons. In NAACL-HLT, 2015.

Thank you