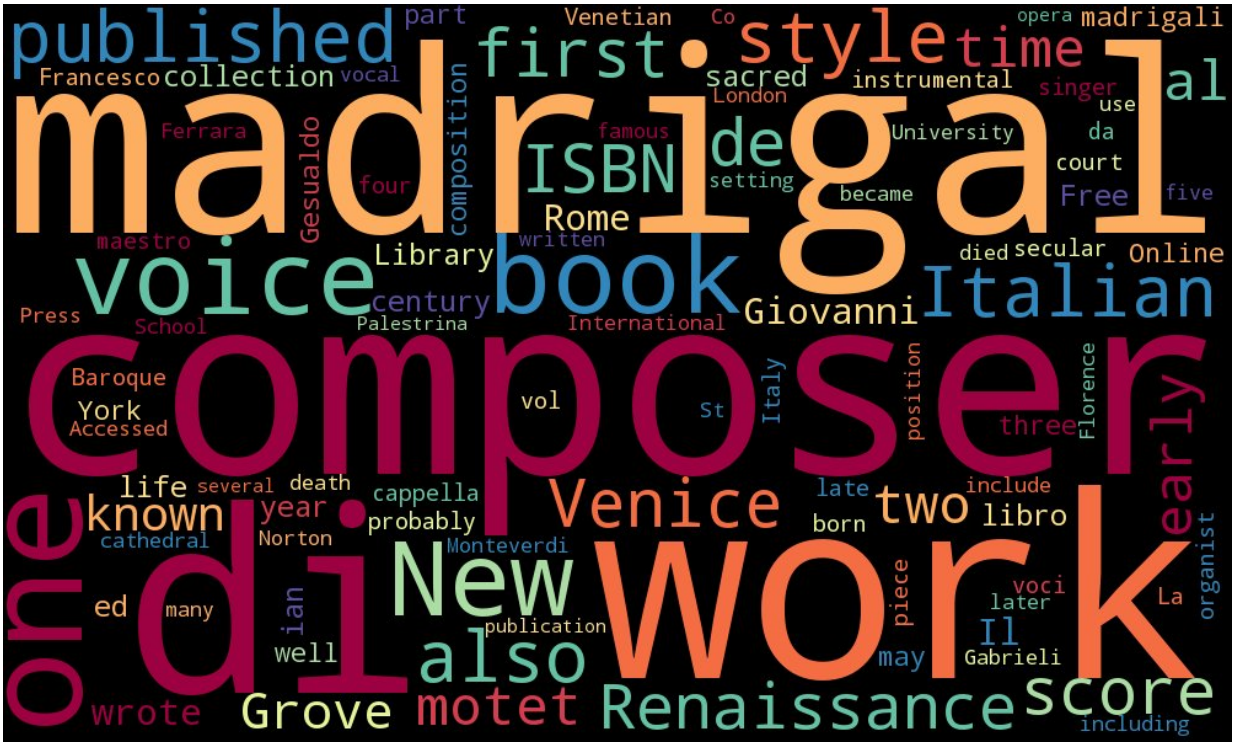
Using machine learning for music knowledge discovery

July 17 2018, by Ingrid Fadelli



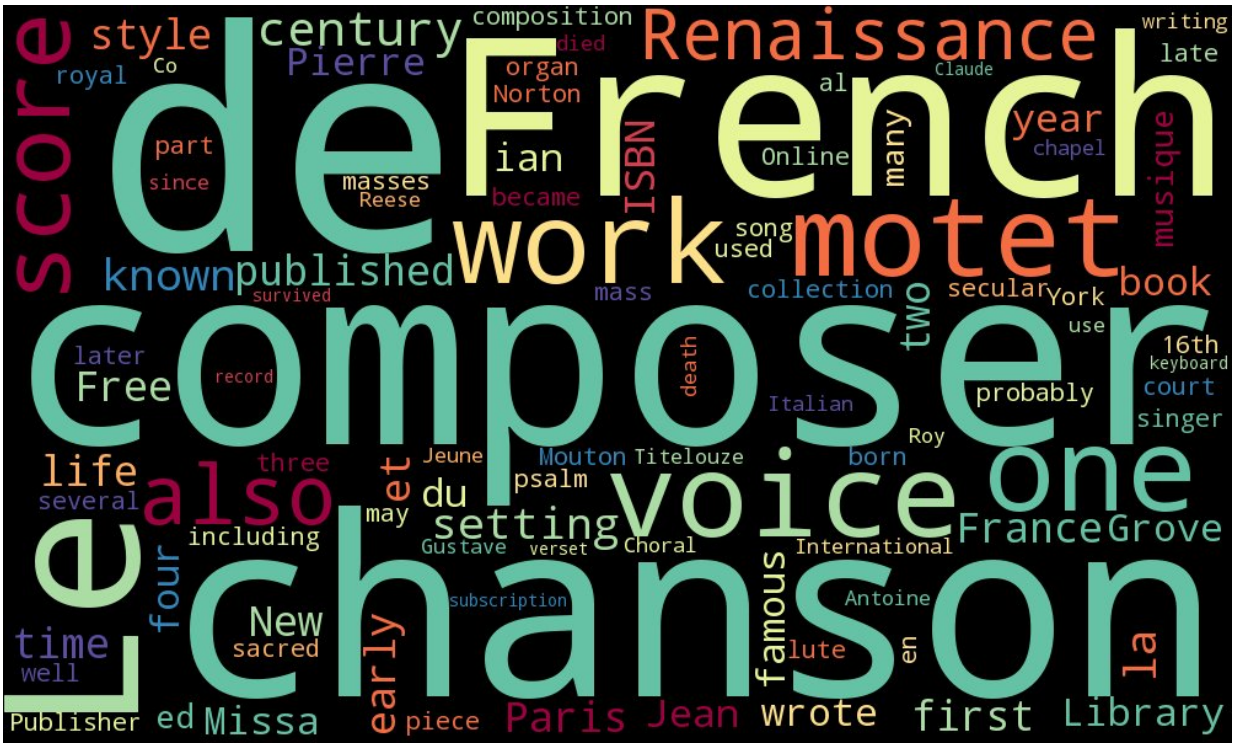
Wordcloud of the Italian Rennaissance music school. Credit: Sergio Oramas et al.

Researchers at the University of Pompeu Fabra, Cardiff University and the Technical University of Madrid used machine-learning algorithms to discover new things about the history of music.

One of the main tasks of musicology researchers is to develop and validate musical hypotheses, after studying historical documents and other available information. Many historical documents have now been digitized and can be accessed and browsed on a computer, making it easy for researchers to access them online. However, basic search engines operate at an "exact text string matching" level, and hence do not always capture the underlying meaning in the content.

In a recently published study, music data science researcher Sergio Oramas and his colleagues tested natural language processing (NLP) approaches that could make the most out of archived historical documents, helping scientists to uncover new hypotheses and identifying interesting patterns in available data.

"As a musicologist, I wanted to exploit the content of large music encyclopaedias, such as the New Grove dictionary or Wikipedia," says Oramas in an interview with *Tech Xplore*. "There is too much content to read and too little time in life, but computers can help us with this."



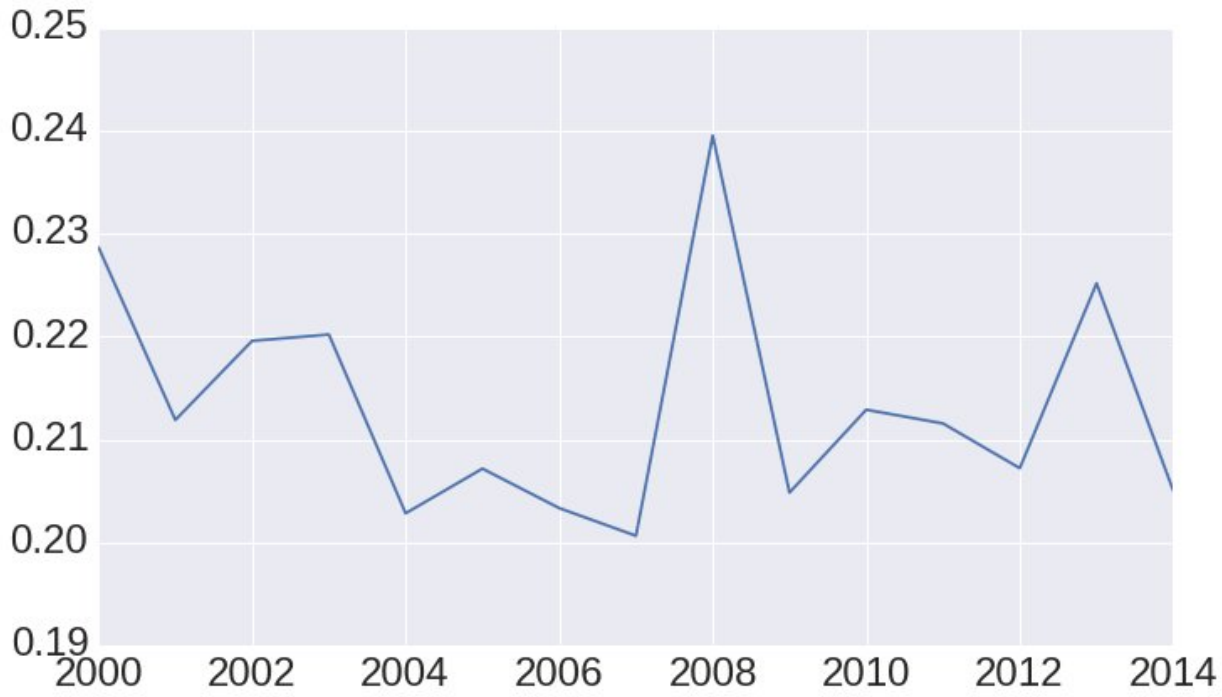
Wordcloud of the French Rennaissance music school. Credit: Sergio Oramas et al.

The work of Oramas and his colleagues applies automatic linguistic processing to large collections of music-related texts in order to discover new facts that are hidden between the lines and assess the potential of machine learning for musicology research. Their study used data from a variety of sources, including Wikipedia, DBpedia, and MusicBrainz, specifically relevant to flamenco, Renaissance music, and popular music.

Using NLP, a computational method of analysing written and spoken human language, the researchers were able to identify interesting patterns in the history of music. "We extracted directly from the data which are the most influential flamenco and Renaissance artists, and discovered migratory tendencies of composers between European cities

in the 15th and 16th century," says Oramas.

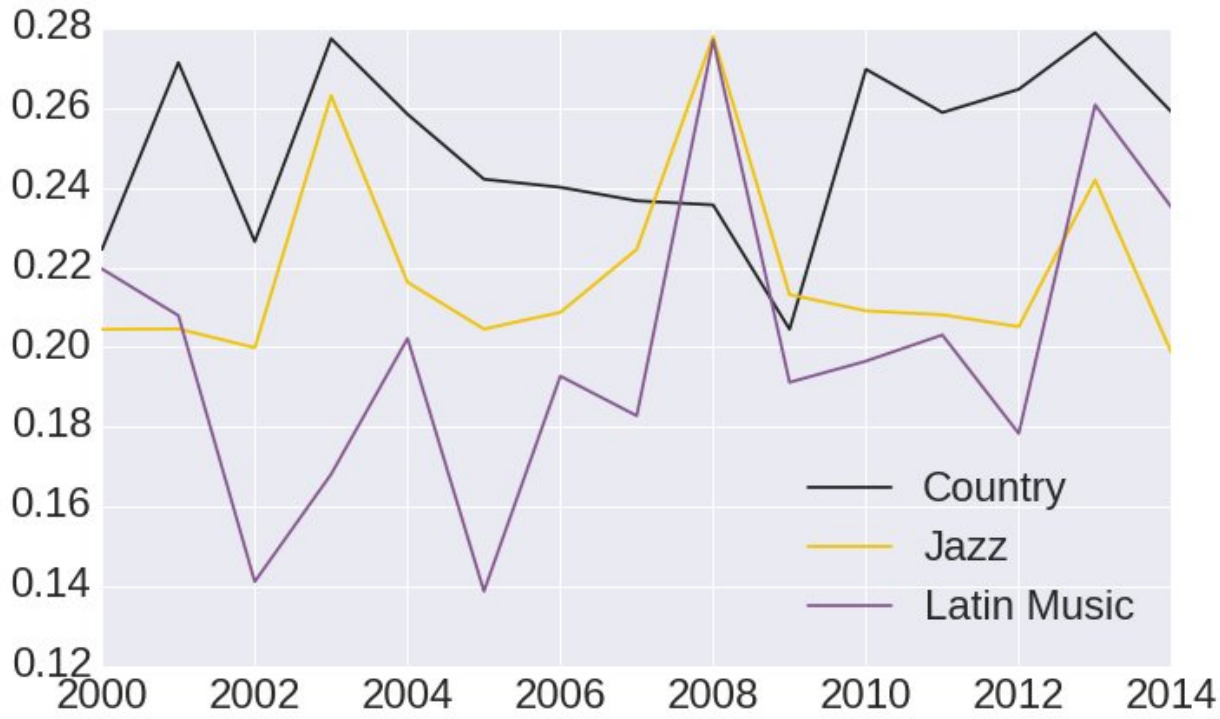
The analysis of Amazon reviews also led to interesting findings about the evolution of popular music, such as an extraordinary positivity in language use in the year 2008, which surprisingly constituted an all-time-high for almost all genres. Remarkably, genres that are traditionally associated with diverse communities, such as jazz and Latin music, had the most noteworthy improvements in the public's positive perceptions, while others (e.g., country) did not.



Graph of the average sentiment by review publication year. Credit: Sergio Oramas et al.

Their study also found a strong correlation between the views expressed by users in their reviews and the popularity of albums released in certain

decades or of particular genres, such as pop in the '60s and reggae in the early '80s. In the case of reggae, for example, they identified a greater proportion of positive reviews between the second half of the '70s and first half of the '80s, which is often referred to as the "golden age of reggae." This increase in popularity could be related to the publication of Bob Marley's albums, which contributed to the genre's popularity at the time.

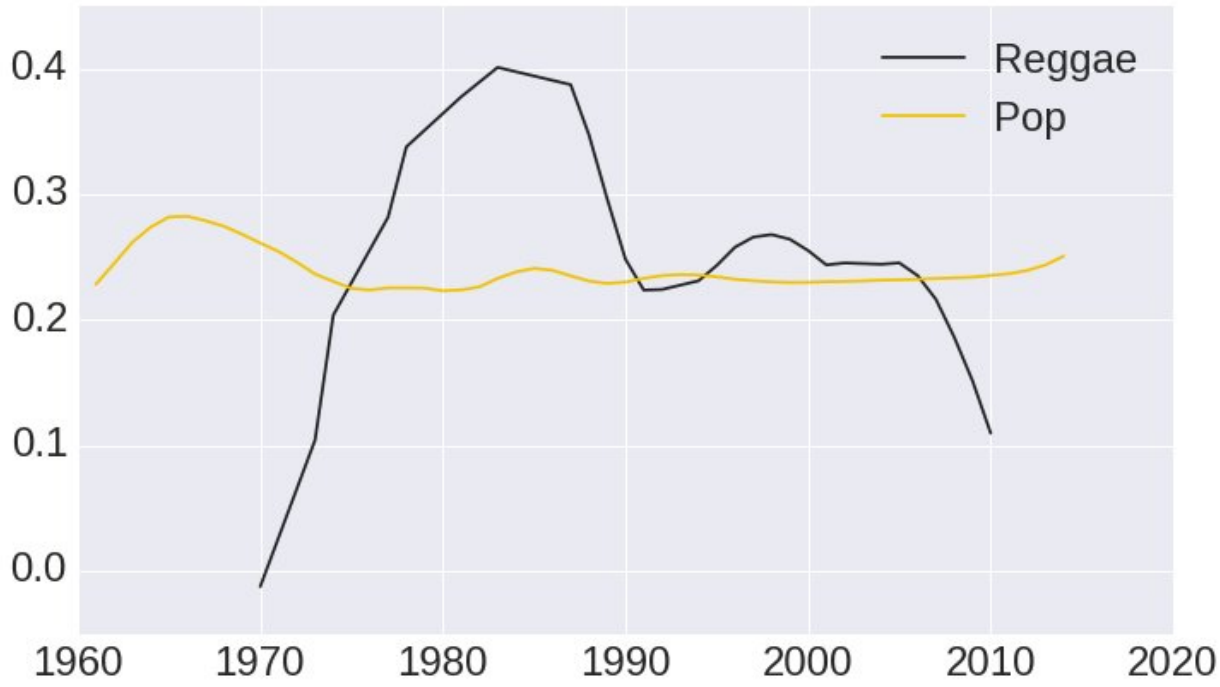


Graph of the average sentiment by genre. Credit: Sergio Oramas et al.

The work of Oramas and his colleagues proves that analysing music reviews written during particular time periods could help musicologists to discover more about the evolution of genres and identify key historical events. "Ultimately, our most meaningful finding is the

demonstration that natural language processing can help to discover new musicological hypothesis, and to gather important insights from the data that may answer these and other questions," explains Oramas.

In future, Oramas plans to expand his research by including other types of content, such as audio, images, and the data collected by the Pandora Music Genome Project, the most sophisticated taxonomy of musical information ever collected.



Graph of the average sentiment by album publication year for pop and raggea album reviews. Credit: Sergio Oramas et al.

More information: Natural Language Processing for Music Knowledge Discovery, arXiv:1807.02200 [cs.CL]

TechXplore

arxiv.org/abs/1807.02200

Sergio Oramas et al. Natural language processing for music knowledge discovery, *Journal of New Music Research* (2018). DOI: 10.1080/09298215.2018.1488878

© 2018 Tech Xplore

Citation: Using machine learning for music knowledge discovery (2018, July 17) retrieved 25 April 2024 from https://techxplore.com/news/2018-07-machine-music-knowledge-discovery.html

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.