

Spotify on the rocks

Ana Santos

ansantos3@gmail.com

Rui Pedro Lima

ruipedro.lima@gmail.com

October 8, 2014

Abstract

Spotify on the Rocks é uma aplicação desenvolvida no âmbito da disciplina de Descrição, Armazenamento e Processamento de Informação (DAPI), disciplina do 1º semestre do 5º ano do Mestrado Integrado em Engenharia Informática e Computação (MIEIC) da Faculdade de Engenharia da Universidade do Porto (FEUP). A aplicação tem como principal objetivo obter grandes dimensões de informação musical, interpretando-a ao nível do input do utilizador e representá-la de acordo com a sua componente geográfica e métrica, oferecendo ao utilizador possíveis estudos e curiosidades sobre os seus interesses musicais.

Ao longo deste documento podem-se encontrar as seguintes secções:

- Estado da arte: pequena descrição daquilo que já existe no mercado, semelhante ao que vai ser desenvolvido, e os aspectos diferenciadores desta aplicação;
- Fontes de informação: apreciação da autoridade da fonte e da qualidade dos dados;
- Estrutura da informação e datasets;
- Modelo conceptual do domínio;
- Tarefas de pesquisa: identificação de algumas tarefas de pesquisa a fazer sobre os dados (pesquisas tipo, cenários de utilização);
- Conclusões;
- Referências;

1 Introdução

A música está presente no quotidiano de milhões de pessoas, seja através de dispositivos móveis, como smartphones ou leitores de música portáteis, ou através do computador, sistema de hi-fi ou até televisão. É indiscutível a importância e valor que esta arte tem. Contudo, numa sociedade sedenta de informação, e num ecossistema como a Internet, onde a procura de informação é desde hobbie até ao modelo de negócio de várias empresas, é bastante comum, até para o casual ouvinte de música, procurar mais informação sobre o artista que está a ouvir, como por exemplo: letra da música, outros trabalhos do artista, videoclip da música, biografia da banda, etc. O desenvolvimento de uma aplicação, onde a música será o domínio no qual nos vamos focar, usando datasets disponíveis no *Spotify* e no *Song-Meanings*, através da plataforma *The Echo Nest*, que permitirá aos utilizadores pesquisas mais alargadas, de forma a satisfazer as suas necessidades e curiosidades, aumentando a experiência para lá do sentido auditivo.

2 Estado da Arte

Actualmente, o Spotify possui os seguintes campos de pesquisa:

- pesquisa por artista, faixa, álbum ou ano;
- pesquisa refinada por AND, OR e NOT, como por exemplo:
 - Zeppelin OR Floyd: lista todos os resultados com as palavras-chave “Zeppelin” ou “Floyd”;
 - Metallica NOT Anger: lista todas as faixas dos Metallica, excepto as que têm a palavra “Anger”;
- pesquisa por género musical;
- pesquisa por label;
- pesquisa por isrc: apresenta todas as faixas correspondentes ao ID, de acordo com o International Standard Recording Code;
- pesquisa por upc: apresenta todos os álbuns correspondentes ao ID, de acordo com o Universal Product Code;

- pesquisa por tag:new: lista os álbuns adicionados mais recentemente;
- Referências;

A aplicação *Spotify on the Rocks* diferencia-se relativamente à aqui descrita, na medida em que permite a combinação de vários campos de pesquisa, para além das acima mencionadas. Além disso, a aplicação integra a API do *Spotify* com o *SongMeanings* (recorrendo à plataforma *The Echo Nest*) e com a API do *Youtube*, para que o utilizador possa ter *lyrics* e videoclipes associados às músicas.

3 Fontes de informação

Spotify é, atualmente, a maior plataforma musical online e é famosa pela quantidade de informação que possui sobre o negócio, mantendo além da informação sobre artistas, álbuns e respectivas faixas, imensa informação bastante precisa sobre os géneros musicais, origem geográfica e cronológica das obras e, limitado a quem possui uma conta (gratuita), informação sobre o histórico de utilizadores, bem como as preferências e construções dos demais que constituem a comunidade virtual.

A empresa, que é sediada em Estocolmo, assinou acordos com as gravadoras Universal Music, Sony BMG, EMI, Hollywood Records e Warner Music, entre outros. O serviço tinha em 15 de setembro de 2010 quase dez milhões de utilizadores. Em março de 2012, tinha cerca de três milhões de utilizadores pagos. Ainda em 2012, o serviço foi premiado na décima sexta edição do Webby Awards, como site mais importante.

O *The Echo Nest* é uma plataforma que agrega diferentes bases de dados com informação de cerca de trinta milhões de músicas, que utiliza *web crawling*, *data mining* e técnicas de processamento de sinais digitais. A empresa fornece informação acerca destas músicas através de uma API. A empresa conta com vários parceiros, tais como *Facebook artists*, *twitter artists*, *SongMeanings*, entre outros. Em Março de 2014, o *Spotify*

anunciou a aquisição do *The Echo Nest*.

De entre dos vários parceiros do *The Echo Nest*, foi seleccionado para este trabalho o *SongMeanings*, uma vez que este não é um site de *lyrics* como os outros: é uma comunidade de milhares de amantes de música, que, além de contribuírem com *lyrics*, discutem e comentam sobre os significados e mensagens subjacentes de determinadas canções. Em setembro de 2011, a *SongMeanings* concordou com os termos da *LyricFind*, licenciando mais de um milhão de *lyrics*. Este acordo faz da *SongMeanings* uma entidade legal, entre as centenas de sites de letras de músicas ilegais, para além de permitir ter letras exactas.

SongMeanings não é um site de *lyrics* como os outros: é uma comunidade de milhares de amantes de música, que, além de contribuírem com *lyrics*, discutem e comentam sobre os significados e mensagens subjacentes de determinadas canções. Em setembro de 2011, a *SongMeanings* concordou com os termos da *LyricFind*, licenciando mais de um milhão de *lyrics*. Este acordo faz da *SongMeanings* uma entidade legal, entre as centenas de sites de letras de músicas ilegais, para além de permitir ter letras exactas.

Uma vez que se pretende que a aplicação disponibilize videoclips, vai também ser utilizada a API do *Youtube*. O *Youtube* é o site mais utilizado para visualizar vídeos. Embora qualquer utilizador possa colocar conteúdos, existem guidelines destinadas a reduzir o abuso de recursos do site. Geralmente, o material proibido inclui conteúdo sexual explícito, vídeos de maus tratos a animais, vídeos chocantes, conteúdo enviado sem o consentimento do titular dos direitos de autor, discursos de ódio, spam, e comportamentos predatórios.

Posto isto, consideramos que todas estas fontes são de elevada reputação, daí as termos escolhido para este trabalho.

4 Estrutura de informação e datasets

Toda a informação do *Spotify* está disponível sob forma de uma *Application Programming Interface* (**API**) online, seguindo arquitetura **REST** (*Representational State Transfer*), oferecendo uma colossal fonte de informação sobre artistas, álbuns, faixas e géneros musicais, bem como a recursos cronológicos sobre os trabalhos, informações geográficas dos artistas, estilos associados e bandas relacionadas, bem como preferências e listas de reprodução construídas pelos utilizadores. A informação proveniente das pesquisas serão guardadas pela aplicação de forma a facilitar a combinação com outros serviços, produzindo análises com teor analítico, facilitando o estudo ou a descoberta de curiosidades fruto do cruzamento de dados. A aplicação é responsável por recolher uma amostra de dados baseadas na informação introduzida pelo utilizador ou, por omissão, descrever as últimas pesquisas efetuadas. Embora a riqueza do serviço, apenas parte da **API** é usada para a aplicação.

5 Modelo conceptual de domínio

6 Tarefas de pesquisa

Com a aplicação anteriormente descrita, pretendemos que seja possível efectuar vários tipos de pesquisa/operações sobre os dados que vamos utilizar, tais como:

- Procurar músicas por título, artista, álbum, género, país, ano ou uma combinação de qualquer destes campos, por exemplos:
 - procurar uma música do género Blues, cujo artista seja da Inglaterra;
 - procurar uma música que seja um trabalho conjunto de 2 artistas;
- Obter uma lista de músicas através de combinações de pesquisas, podendo com isto criar uma playlist, se o utilizador estiver autenticado, como por exemplo:

- uma playlist com 10 músicas dos Metallica e dos Muse;
- uma playlist com músicas portuguesas, ou com músicas portuguesas e brasileiras
- Utilizar a informação de uma música para obter a letra (*lyrics*) a esta associada, através do *SongMeanings*. Para além disso, podemos apresentar as discussões feitas acerca do significado dessa mesma música, uma vez que é esse o principal objectivo do *SongMeanings*;
- Tirar partido das playlists dos utilizadores, para ver quantos *followers* o *Spotify* tem, e representá-los geograficamente (por exemplo, através de um mapa ou de um gráfico)
- Obter, através de toda esta informação, uma representação gráfica de dados estatísticos, como por exemplo:
 - número de utilizadores portugueses registados no *spotify*
 - músicas mais ouvidas ou que são mais vezes avançadas (*skipped*)
 - género musical mais apreciado pelos utilizadores
 - *word counter* em pesquisas conjuntas, seja ao nível de faixa, álbum, descrição, etc

Consideramos que estes cenários de utilização vão ser bastante úteis, uma vez que o *Spotify* não possibilita pesquisas num âmbito tão alargado.

7 Conclusões