

Juggle – Hybrid large-scale music recommendation

José Devezas (DCC–FCUP & Laboratório SAPO/U.Porto)

Sérgio Nunes (DEI–FEUP & INESC TEC)

Projeto integrado no âmbito do MusicBox

Juggle – Hybrid large-scale music recommendation

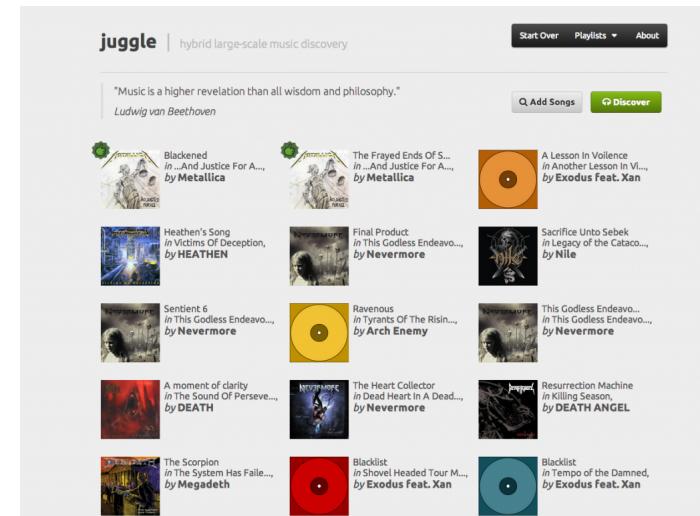
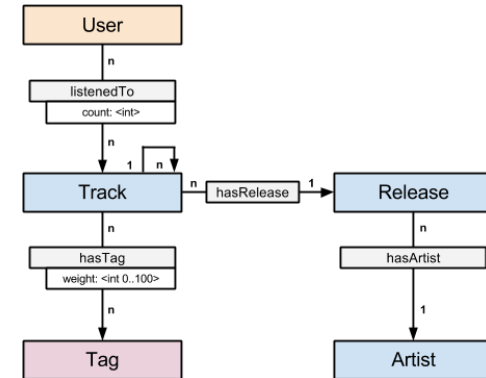
Objectivos

- Implementar e melhorar algoritmos de recomendação, de forma a facilitar a descoberta de nova música.
- Desenvolver software escalável capaz de lidar com milhões de músicas e um número crescente de características associadas a cada faixa.
- Integrar diferentes tipos de características, nomeadamente:
 - **Perfis de utilizador** (e.g. n.º de vezes que ouviu cada faixa, faixas favoritas, etc.)
 - **Conteúdo** áudio e textual (e.g. pitch, timbre, letras)
 - **Contexto** (e.g. tags, localização, estado do tempo, luminosidade, ruído do ambiente, etc.)

Juggle – Hybrid large-scale music recommendation

Resultados

- Grafo de conhecimento musical com:
 - 2,8 milhões de nós (e.g. faixas, tags, utilizadores)
 - 61,8 milhões de arestas (e.g. vizinhos mais próximos, características das faixas)
 - 71,7 milhões de propriedades, associadas aos nós e arestas (e.g. meta-dados, pesos de tags, valores pré-computados)
 - Algoritmos de recomendação baseados em Gremlin:
- A técnica foi, genericamente: (1) **tentar dar mais peso a tags mais descritivas usando o TF-IDF(tag, faixa), procurando focar na long tail da distribuição de tags;** (2) **combinar com as preferências dos utilizadores com gostos semelhantes (filtragem colaborativa);** e (3) **combinar com o conteúdo, usando a distância euclidiana para encontrar os k=10 vizinhos mais próximos segundo um vetor de características áudio.**
 - Cada característica contribuiu para o peso de cada candidato a recomendar contando o número de vezes que o algoritmo de travessia passou pelo nó correspondente.



<http://juggle.fe.up.pt/static/app.html>