

Análise de Soluções para Busca por Similaridade (*Matching*) de Dados Musicais

Gisele Bernardes da Silva¹

¹Departamento de Informática e Estatística – Universidade Federal de Santa Catarina (UFSC)
Caixa Postal 476 – 88.045-108 – Florianópolis – SC – Brazil

giselebernardes@grad.ufsc.br

Resumo. *Este artigo visa apresentar e comparar soluções para recuperação de informação musical. A intenção é analisar soluções que não necessariamente buscam dados musicais apenas através do casamento direto de parâmetros de entrada para a busca, como título da música, palavras-chave ou um áudio como parte da música, mas também através do casamento aproximado (ou similar) destes parâmetros.*

Abstract. *This article aims to present and compare solutions for music information retrieval. The intention is to analyze solutions that do not necessarily search for musical data only through direct marriage of input parameters to the search, such as song title, keywords or an audio with part of the song, but also through approximate (or similar) these parameters.*

1. Introdução

A música era um meio de comunicação exclusivamente presencial. No decorrer do tempo, as técnicas e invenções aplicadas ao processo de gravação do som foram surgindo e se aperfeiçoando. Com a Internet, a música ultrapassa os limites físicos da mídia, mergulhando no universo digital. Desta forma, a organização da informação, que inclui a sua representação, tem a principal finalidade de possibilitar a recuperação dessa informação, além da sua guarda para a posteridade. A música se tornou um objeto de consumo universal e extremamente acessível [Daquino 2012, Gomes 2015].

A busca por similaridade musical está inserida dentro de um tema de estudos denominado *Music Information Retrieval*. Os pesquisadores de MIR observam que a motivação maior para essa área de pesquisa é o grande volume de música digital disponível na Internet que, quanto mais cresce, menos possibilita sua recuperação eficiente, visto que estão disponíveis em grande volume, mas sem o tratamento adequado [McLane 1996, Downie 2003, Santini and de Souza 2007].

Este trabalho tem como objetivo geral estudar o estado da arte sobre a recuperação da informação de dados musicais e realizar uma análise comparativa de algumas soluções existentes que não necessariamente buscam dados musicais apenas através do casamento direto de parâmetros de entrada para a busca, mas também através da similaridade destes parâmetros.

Para atingir este objetivo, foi necessário entender os conceitos básicos de som e como o som é transformado em áudio digital ([Muller 2007, Zuben 2004, Pacheco and Lopes 2014]); identificar os formatos de dados musicais e como é feito

o armazenamento deles em bancos de dados ou repositórios digitais ([Semidão 2013, Michels 1992]); e estudar os métodos e algoritmos utilizados para busca por similaridade de dados musicais.

Portanto, a relevância deste trabalho pretende contribuir diretamente, para agregar conhecimento, com o estudo sobre a recuperação da informação de dados musicais, que auxiliarão no desenvolvimento futuro de soluções para busca por similaridade de dados musicais.

2. Soluções Existentes

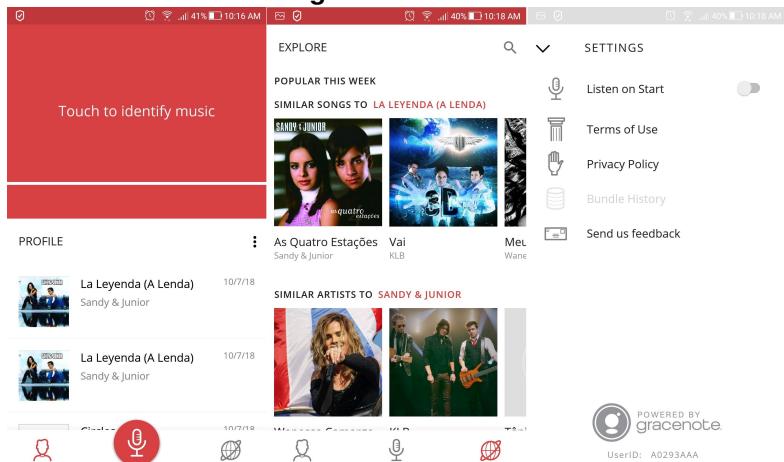
Nos últimos anos, várias plataformas digitais de *streaming* têm surgido, derivado das intensas procura por música on-line pelos usuários. Este capítulo apresenta de forma resumida as principais soluções comerciais e da academia, para busca de dados musicais.

2.1. Soluções Comerciais

2.1.1. MusicID

Gracenote Inc., fundada em 1998, é uma empresa que fornece metadados de música, vídeo, esportes e tecnologias de reconhecimento automático de conteúdo para empresas e serviços de entretenimento em todo o mundo. A solução da empresa destinada à busca de dados musicais se chama *Gracenote MusicID®* disponível para *smartphones* (ver Figura 1).

Figura 1. MusicID



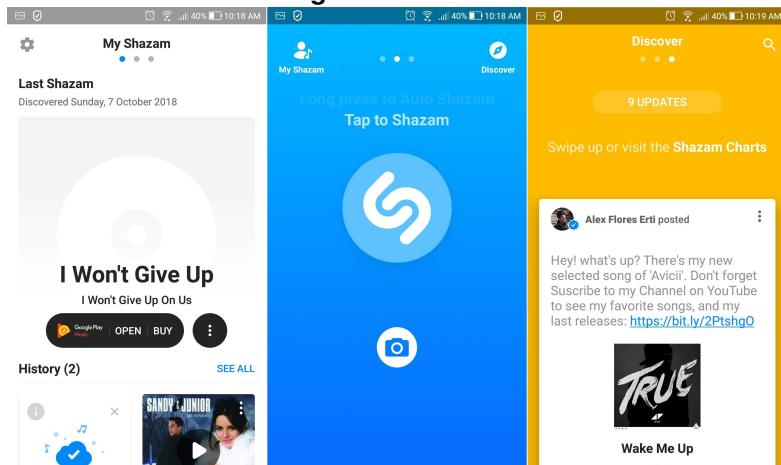
Fonte: Elaborado pela autora

O Gracenote MusicID®, faz o reconhecimento de músicas que são tocadas ao seu redor, combinado ao uso de *fingerprints* e correspondência de texto para identificar arquivos de música digital em um banco de dados mundial de informações musicais. Uma vez reconhecidos, os arquivos são organizados por nome de faixa, nome do álbum e caminhos de pastas e, então, apresentados ao usuário. Ele é um aplicativo para *smartphones* que identifica músicas ouvindo você cantar, cantarolar ou de músicas que são tocadas ao seu redor.

2.1.2. Shazam

Shazam Entertainment Ltd. foi fundada em 2000 com a idéia de prover um serviço que pudesse conectar as pessoas à música, permitindo a identificação da música através de *smartphones*. A aplicação (ver Figura 2) usa o microfone do *smartphone* ou do computador para capturar uma pequena amostra de música e, então, realiza a identificação da música em um grande banco de dados com mais de 12 bilhões de músicas, com uma alta taxa de acertos.

Figura 2. Shazam



Fonte: Elaborado pela autora

Para o trecho de música capturado pela aplicação é criado uma *fingerprint*, que é comparada com todas as outras *fingerprints* derivadas das músicas no banco de dados. Se houver uma correspondência, são enviadas informações da música para o usuário, como artista, álbum e título da música.

2.1.3. SoundHound

SoundHound Inc., fundada em 2005, é uma empresa pioneira em desenvolvimento de aplicações para reconhecimento de voz, compreensão da linguagem natural, reconhecimento de som e tecnologias de busca.

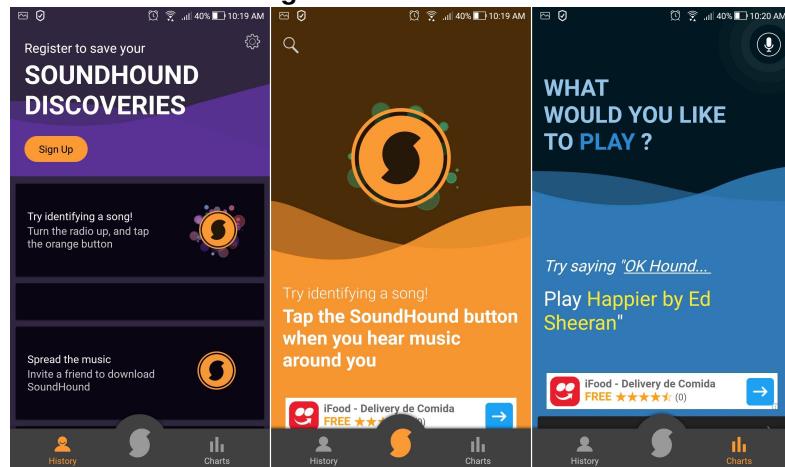
A plataforma independente de Inteligência Artificial *Houndify*, combinada ao *Automatic Speech Recognition* (ASR) e o *Natural Language Understanding* (NLU), permite ao *SoundHound* a identificação de músicas de forma rápida e eficiente. Seus dois produtos conhecidos no meio musical são:

1. *SoundHound Music Search & Play*¹: aplicativo para *smartphones* onde é possível descobrir, pesquisar e reproduzir qualquer música com controle de voz (ver Figura 3). Ele é um aplicativo para *smartphones* que identifica músicas ouvindo você cantar, cantarolar ou de músicas que são tocadas ao seu redor.

¹<https://soundhound.com/soundhound>

2. *Midomi*²: aplicação com as mesmas características do item anterior, porém possui versão para web. Sua versão mobile é destinada a modelos mais antigos de smartphones.

Figura 3. SoundHound

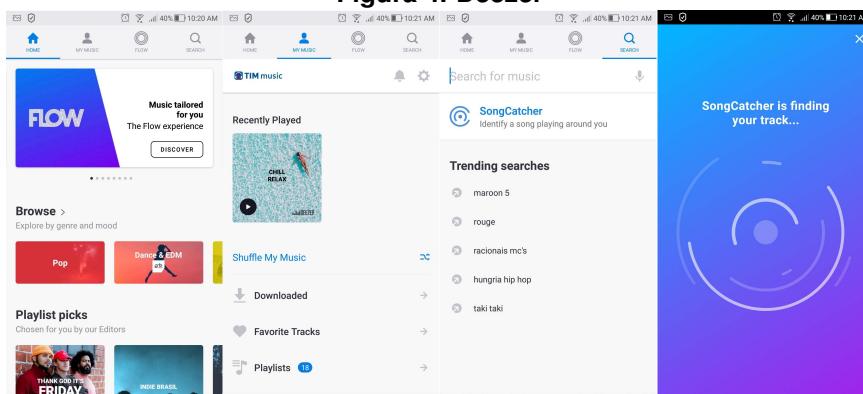


Fonte: Elaborado pela autora

2.1.4. Deezer

Esta solução nasceu da necessidade de facilitar a vida de seu fundador para ouvir e realizar o *download* de músicas. Com isso, o idealizador da plataforma desenvolveu o *Blogmusik.net* em 2006. Devido a sua popularidade, houve objeção de detentores de direitos autores, o que gerou o fechamento do site. Pouco tempo depois, um acordo foi assinado e o antigo site voltou ao ar com o nome de *Deezer* (ver Figura 4).

Figura 4. Deezer



Fonte: Elaborado pela autora

A *Deezer* (ver Figura 4), também conta com uma série de aplicativos que complementam a experiência musical do usuário. O *Stateeztics*, por exemplo, é um *in-app* exclusivo que traça o perfil musical do usuário e mostra suas estatísticas de consumo a

²<https://www.midomi.com/>

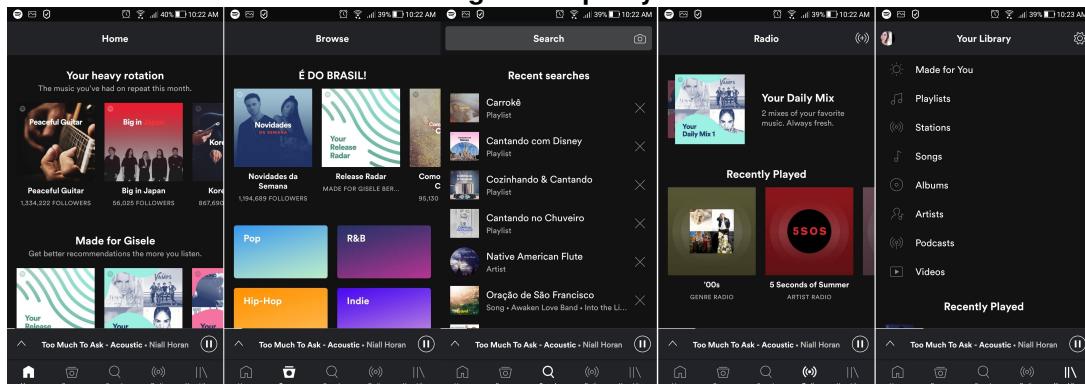
partir do seu histórico sonoro. Outra aplicação disponível é o *Edjing*, que oferece mixagem de músicas com diversas ferramentas de efeitos digitais, além de contar com uma interface bastante intuitiva. Já o usuário que está aprendendo a tocar instrumentos musicais pode contar com o *Chordify*, que reconhece o som que está tocando na Deezer e faz a transcrição automática da harmonia em cifras.

Recentemente, no final do ano de 2017, além da correspondência de texto para identificar arquivos de música digital, a Deezer lançou o seu próprio recurso de identificação de músicas que são tocadas ao seu redor combinado ao uso de *fingerprints*, o *SongCatcher*, desenvolvido pela *ACRCloud* (ver subseção 2.1.8).

2.1.5. Spotify

Spotify Ltd., fundada em 2006, é um serviço de *streaming* de música, *podcast* e vídeo, além de ser o mais usado no mundo. A plataforma fornece conteúdo protegido provido de restrição pela gestão de direitos digitais de gravadoras e empresas de mídia (ver Figura 5). O Spotify é um serviço *freemium*: ele possui recursos gratuitos com propagandas ou limitações, e recursos adicionais, como qualidade de transmissão aprimorada e *downloads* de música, que são oferecidos para assinaturas pagas.

Figura 5. Spotify



Fonte: Elaborado pela autora

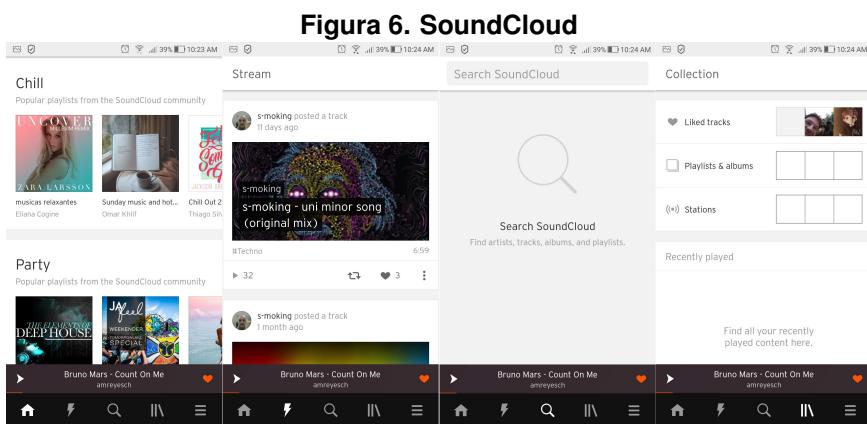
A plataforma emprega um modelo de distribuição de dados híbrido com uma combinação de compartilhamento de dados peer-to-peer³ (P2P) e uma infraestrutura de servidor. Ao pesquisar uma música através do smartphone e desejar ouvi-lá, o sistema primeiro verifica se a música já se encontra baixada na memória *cache* do *smartphone* para agilizar o processo. Em caso negativo, é feita a conexão diretamente com o servidor do Spotify ao mesmo tempo que o método busca "peers" entre milhões de usuários para que a música que se queira ouvir seja baixada o mais rápido possível. A busca da música é feita pela correspondência de texto para identificar arquivos de música digital através da técnica de recuperação por conteúdo.

³Do inglês par-a-par ou simplesmente ponto-a-ponto, é uma arquitetura de redes de computadores onde cada um dos pontos ou nós da rede funciona tanto como cliente quanto como servidor, permitindo compartilhamentos de serviços e dados sem a necessidade de um servidor central.

O Spotify disponibiliza uma *Web API*⁴ que permite que desenvolvedores integrem o conteúdo do Spotify em seus próprios aplicativos. O Spotify *Web API* é um serviço com base na arquitetura REST, que retorna em formato JSON dados sobre álbuns, artistas, faixas, playlists, entre outros. Para acessar outras informações é necessária uma autenticação *OAuth*.

2.1.6. SoundCloud

SoundCloud, criada em 2007, é uma plataforma on-line de publicação de áudio utilizada por profissionais de música (ver Figura 6). Nela os músicos podem colaborar, compartilhar, promover e distribuir suas composições. Originalmente, seu objetivo era permitir que profissionais da música trocassem ideias sobre as composições nas quais estão trabalhando, permitindo uma fácil colaboração e comunicação antes de um lançamento público. Hoje, o site também é utilizado por ouvintes e usuários da web em geral.



Fonte: Elaborado pela autora

Os usuários registrados podem ouvir o máximo de conteúdo como quiserem e podem fazer o *upload* de até 180 minutos de áudio ao seu perfil. Todos esses recursos são gratuitos e estão disponíveis para todos os usuários registrados. A plataforma possui uma API integrada a várias aplicações, que permitem fazer o *upload* ou *download* de música e arquivos de música.

O SoundCloud descreve as faixas de música graficamente como formas de onda e permite aos usuários comentar partes específicas do áudio (conhecido como comentários cronometrados). Estes comentários são exibidos ao escutar a parte do áudio que estão se referindo. Outras características incluem respostas, listas de reprodução, seguidores e *downloads* digitais de cortesia. A busca da música é feita pela correspondência de texto para identificar arquivos de música digital através da técnica de recuperação por conteúdo. Não foi encontrada documentação informando como funciona detahadamente a recuperação da música.

⁴Do termo em inglês "Application Programming Interface" que significa em tradução para o português "Interface de Programação de Aplicativos". É uma forma de integrar sistemas, possibilitando benefícios como a facilidade no intercâmbio entre informações com diferentes linguagens de programação.

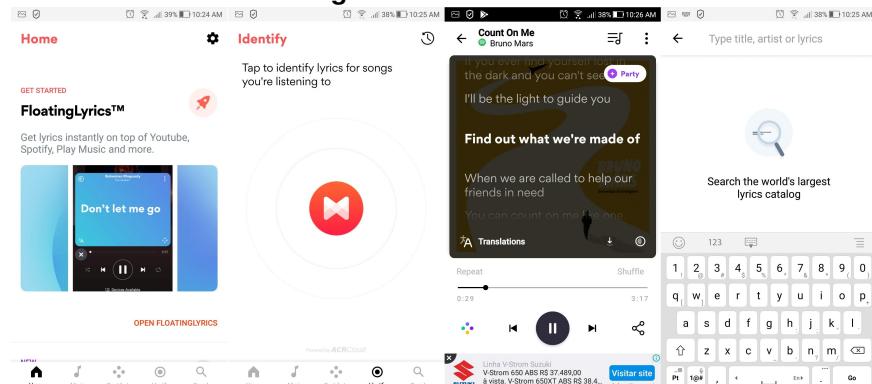
O SoundCloud também disponibiliza uma *Web API* que permite que desenvolvedores integrem o conteúdo do SoundCloud em seus próprios aplicativos. O SoundCloud *Web API* é um serviço com base na arquitetura HTTP, que retorna em formato JSON dados sobre álbuns, artistas, faixas, playlists, entre outros. Para acessar outras informações é necessária uma autenticação *OAuth*.

2.1.7. Musixmatch

A *Musixmatch* foi criada em 2010 com o objetivo de mudar a forma como as pessoas experimentam música e letras.

A plataforma pode ser acessada através do site ou via aplicativo para *smartphones*. O Musixmatch digitaliza todas as músicas da biblioteca de música do usuário e encontra letras para todas elas, identificando a letra da música e mantendo sincronizada enquanto a música é tocada. Além da correspondência de texto para identificar arquivos de música digital através da técnica de recuperação por conteúdo, ela possui também a capacidade para capturar uma pequena amostra de música através de *fingerprints* (mesma função encontrada em soluções como o *Shazam*), desenvolvida pela ACRCLOUD.

Figura 7. Musixmatch



Fonte: Elaborado pela autora

2.1.8. ACRCLOUD

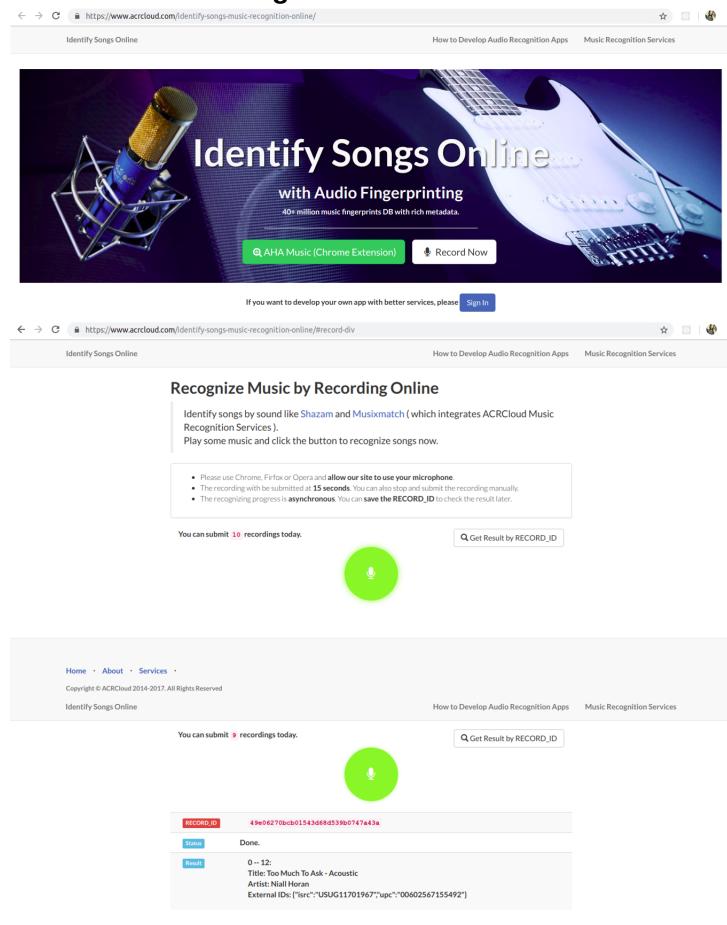
A *ACRCLOUD* foi criada em 2015, sendo a solução vitoriosa no campeonato de *Audio Fingerprinting* do MIREX2015, organizado pelo Laboratório Internacional de Avaliação de Sistemas de Recuperação de Informação Musical (IMIRSEL, sigla em inglês).

ACR (*Automatic Content Recognition*) é uma tecnologia de identificação para reconhecimento de conteúdo reproduzido em um dispositivo de mídia. Ele permite que usuários obtenham rapidamente informações detalhadas sobre o conteúdo que acabaram de experimentar sem qualquer entrada de texto ou esforço de pesquisa.

ACRCLOUD (ver Figura 8) é uma plataforma de microserviços na nuvem que possui reconhecimento de música através de *fingerprints*, onde identifica músicas ouvindo você cantar ou de músicas que são tocadas ao seu redor, além do monitoramento de transmissão

com identificação e apresentação de conteúdo, entre outros. Ele possui integração com serviços de música como o Spotify, Deezer, entre outros, que permite desenvolvedores acessarem diretamente esses serviços e oferecer links diretos para seus usuários.

Figura 8. ACRCLOUD



Fonte: [ACRCLOUD 2015], elaborado pela autora

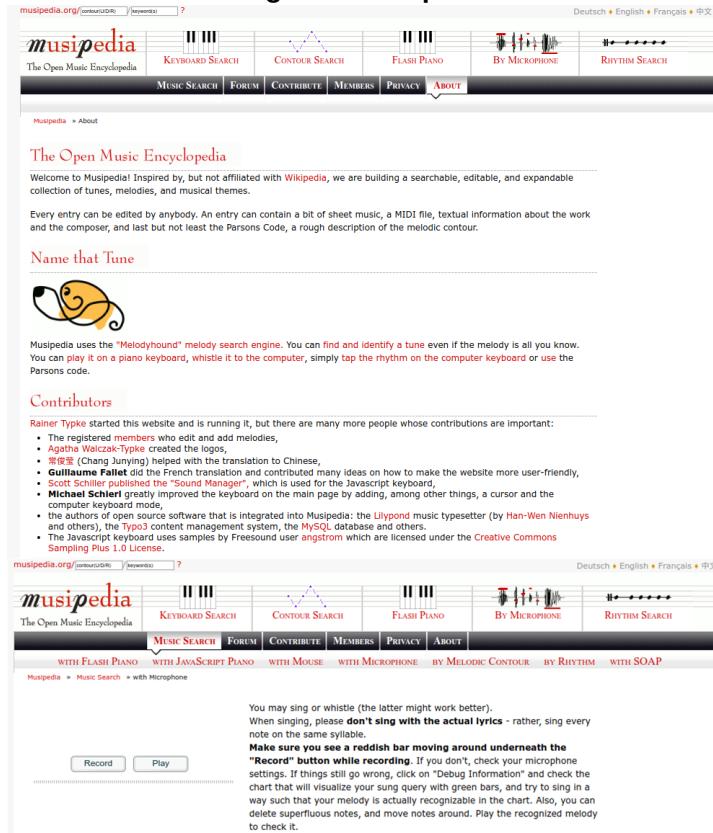
2.1.9. Musipedia

Musipedia é uma enciclopédia aberta de música, criação inspirada no Wikipedia⁵, para localização, edição e expansão de coleções de tons, melodias e temas musicais (ver Figura 9). A enciclopédia utiliza o mecanismo de pesquisa de melodias, do qual chamam de *melodyhound*, onde é possível encontrar e identificar uma música, mesmo que a melodia seja tudo o que você saiba no momento. A busca também pode ser feita através da pesquisa de contorno melódico (Código de Pearson) ou com base no ritmo. Ainda, os conteúdos podem ser alterados por qualquer usuário, podendo conter um pedaço de música, um arquivo MIDI, informações textuais sobre o trabalho e o compositor.

É possível também integrar a pesquisa do Musipedia ao seu próprio serviço web, utilizando as interfaces SOAP, que possibilitam pesquisar com base na melodia, no contorno melódico ou no ritmo.

⁵<https://www.wikipedia.org/>

Figura 9. Musipedia



Fonte: [Development 2018], elaborado pela autora

2.2. Soluções Acadêmicas

2.2.1. AMUSE

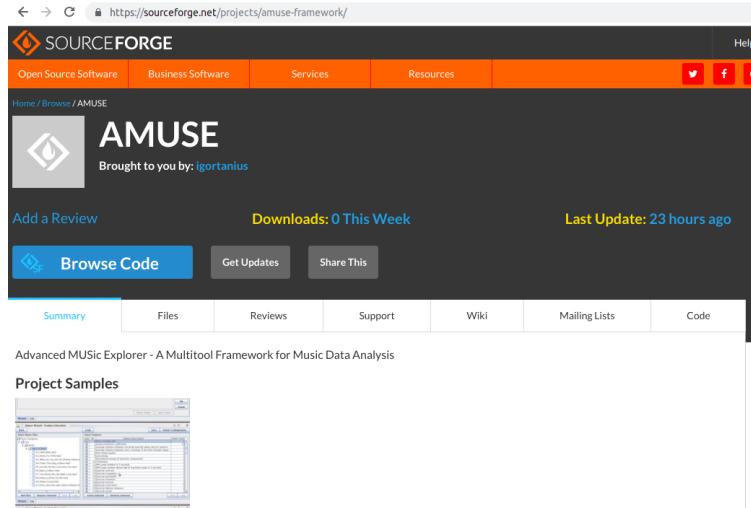
AMUSE (*Advanced Music Explorer*)⁶ é um *framework* desenvolvido pela TU Dortmund, na Alemanha, licenciado sob a GPL e implementado em JAVA. Portanto, ele pode ser executado em qualquer sistema operacional que suporte o Java Runtime Environment (ver Figura 10).

Segundo [Vatolkin et al. 2010], o *framework* fornece diferentes funcionalidades, como:

- Processamento de som, convertendo arquivos de áudio MP3 em ondas sonoras;
- *Downsampling* e estéreo para a conversão de arquivos de áudio mono;
- Divisão automática de arquivos wave;
- Escalabilidade usando multi-threading em uma máquina ou fornecendo as tarefas para sistemas de grade como Sun Grid Engine ou LSF Batch;
- Gerenciamento eficiente do conjunto de dados que suporta diretamente o formato WEKA ARFF;
- Componente logger integrado.

⁶<https://sourceforge.net/projects/amuse-framework/>

Figura 10. AMUSE



Fonte: [Dortmund 2018], elaborado pela autora

O AMUSE possui subtarefas em cadeia para recuperação da informação musical. Cada tarefa pode ser calculada em várias unidades de processamento. Inicia-se pela tarefa de extração de recursos, que fornece descritores numéricos de baixo nível ou alto nível do sinal de áudio (por exemplo, extração de melodia da música). Depois que a tarefa de extração é carregada na memória, é realizado o processamento dos recursos, em uma etapa intermediária, que serve de entrada para a técnica de classificação. Por fim, é realizada a validação dos resultados da classificação.

As ferramentas integradas não têm restrições de uso em relação aos seus códigos-fonte. Se eles não estiverem disponíveis como bibliotecas Java, as versões executáveis deverão ser fornecidas. Nesse caso, pode certamente levar à dependência do sistema operacional em execução.

O projeto é oferecido gratuitamente à comunidade de pesquisa. Informações mais detalhadas sobre o projeto podem ser consultadas em [Vatolkin et al. 2010] e [Dortmund 2018].

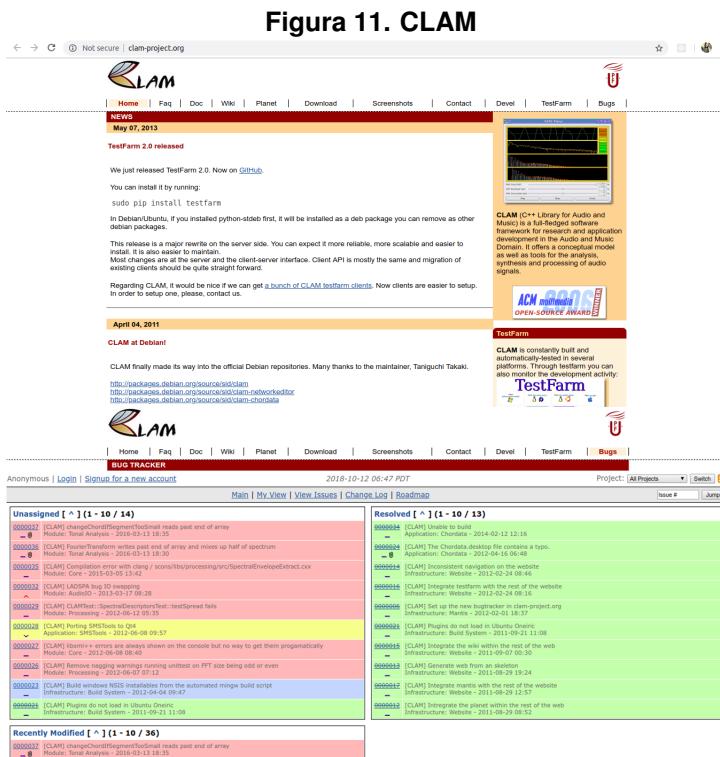
2.2.2. CLAM

CLAM (*C++ Library for Audio and Music*)⁷ é um *framework* desenvolvido em C++ no *Music Technology Group* (MTG) da Universidade Pompeu Fabra em Barcelona, Espanha. Ele oferece uma plataforma completa de desenvolvimento e pesquisa para o domínio de áudio e música baseado na técnica SMS. Além de oferecer um modelo abstrato para sistemas de áudio, ele também inclui um repositório de algoritmos de processamento e tipos de dados, bem como diversas ferramentas, como entrada/saída de áudio ou MIDI (ver Figura 11).

Segundo [Amatriain 2004], as características mais importantes do framework são:

- Verdadeiramente orientado a objetos. Extensas técnicas de engenharia de software

⁷<http://clam-project.org/>



Fonte: [Development 2004], elaborado pela autora

foram aplicadas para projetar uma estrutura que seja altamente (re)utilizável e compreensível;

- É abrangente, uma vez que não só inclui classes para processamento, mas também para entrada e saída de áudio e MIDI, serviços de serialização XML, algoritmos e visualização e interação de dados, e manipulação multi-threading;
- Lida com uma ampla variedade de tipos de dados extensíveis que vão desde sinais de baixo nível (como áudio ou espectro) até estruturas semânticas de nível superior (como frase musical ou segmento);
- É multiplataforma. Todo o código é ANSI C ++ e é regularmente compilado no Linux, Windows e Mac OSX usando os compiladores mais usados. Até mesmo o código para entrada/saída, visualização e multithreading é de plataforma cruzada até a camada mais baixa possível;
- O projeto está licenciado sob os termos e condições GPL (Licença Pública GNU). Apesar de possuir a opção de licenciamento duplo da estrutura (ou seja, oferecer uma licença comercial alternativa), tudo o que é oferecido na versão pública é GPL e o projeto é, portanto, Software Livre, código aberto e colaborativo;
- Base para todos os desenvolvimentos futuros no MTG e, portanto, mantido e atualizado regularmente;
- O framework pode ser usado como uma biblioteca C ++ regular ou como uma ferramenta de prototipagem. No primeiro modo, o usuário pode estender, adaptar ou otimizar a funcionalidade da estrutura para implementar um aplicativo específico. No segundo modo, o usuário pode facilmente construir um protótipo para testar um novo algoritmo ou aplicativo de processamento de sinais.

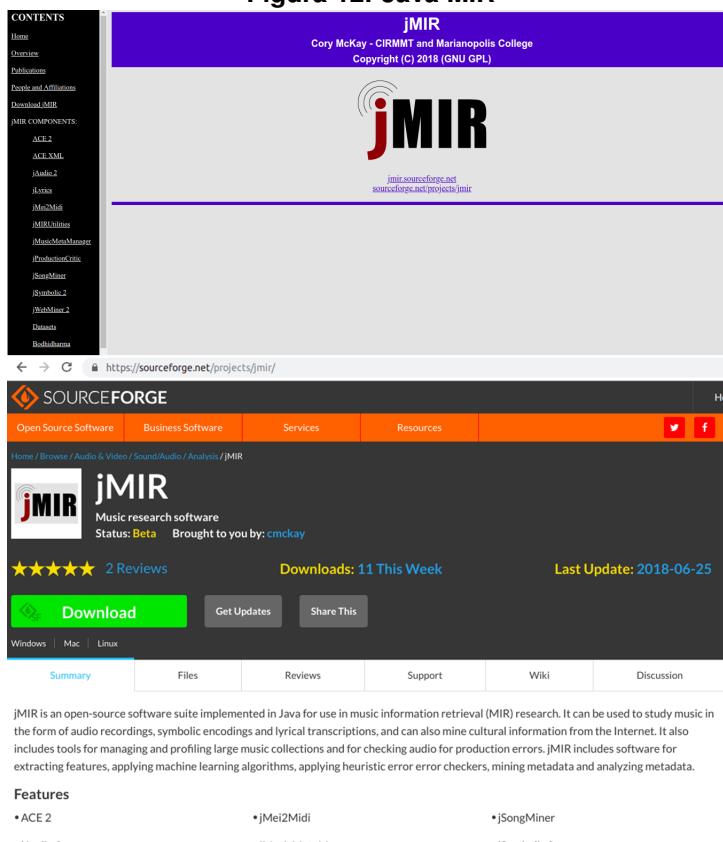
Informações mais detalhadas sobre este *framework* podem ser consultadas em

[Amatriain 2007, Amatriain 2004].

2.2.3. Java MIR

jMIR (Java MIR)⁸ é um *software* que possui um conjunto de componentes desenvolvido na CIRMMT e Marianopolis College, ambos localizados no Canadá. Cada um dos componentes pode ser utilizado separadamente ou como um todo (ver Figura 12).

Figura 12. Java MIR



Fonte: [McKay 2018], elaborado pela autora

O software é de código livre implementado em Java para uso nas pesquisas de Recuperação de Informação Musical (MIR) e se baseia em uma técnica de mineração de dados, a classificação. Ele pode ser usado para estudar música na forma de gravações de áudio, codificações simbólicas e transcrições líricas, e também pode extrair informações culturais da Internet. Ainda, ele inclui ferramentas para gerenciar e criar perfis de grandes coleções de músicas e para verificar o áudio quanto a erros de produção. É bem documentado e inclui GUIs para aumentar a usabilidade geral.

O objetivo principal do *software* é auxiliar nas pesquisas em classificação automática de música e a análise de similaridade, proporcionando as seguintes características:

⁸<http://jmri.sourceforge.net/>

- Tornar tecnologias sofisticadas de reconhecimento de padrões acessíveis a pesquisadores de música com históricos técnicos e não técnicos;
- Eliminar duplicação redundante de esforço;
- Aumentar a cooperação e a comunicação entre os grupos de pesquisa;
 - Facilitar o desenvolvimento iterativo e o compartilhamento de novas tecnologias MIR;
 - Facilitar comparações objetivas de algoritmos.
- Facilitar a pesquisa combinando características musicais de alto nível, baixo nível e culturais (ou seja, características simbólicas, áudio e web-minadas).

Informações mais detalhadas sobre o projeto estão disponíveis nas publicações acadêmicas⁹. Manuais e documentação para cada componente também podem ser consultados em [McKay 2018] e [McKay 2010].

2.2.4. MIRtoolbox

*MIRtoolbox*¹⁰ é um pacote de ferramentas escritas em Matlab para a extração de recursos musicais, como tonalidade e ritmo, tanto para especialistas quanto para não especialistas do Matlab (ver Figura 13). Ele foi desenvolvido dentro do contexto de um Projeto Europeu chamado “Tuning the Brain for Music”, dedicado ao estudo da música e da emoção, com colaboração entre neurociências, psicologia cognitiva e ciência da computação. Os grupos e instituições envolvidos são a Music Cognition Team da University of Jyväskylä na Finlândia e o Music Acoustics Group do KTH em Estocolmo.

Segundo [Lartillot 2013], foi elaborado um manual onde são descritas as seguintes especificações da solução:

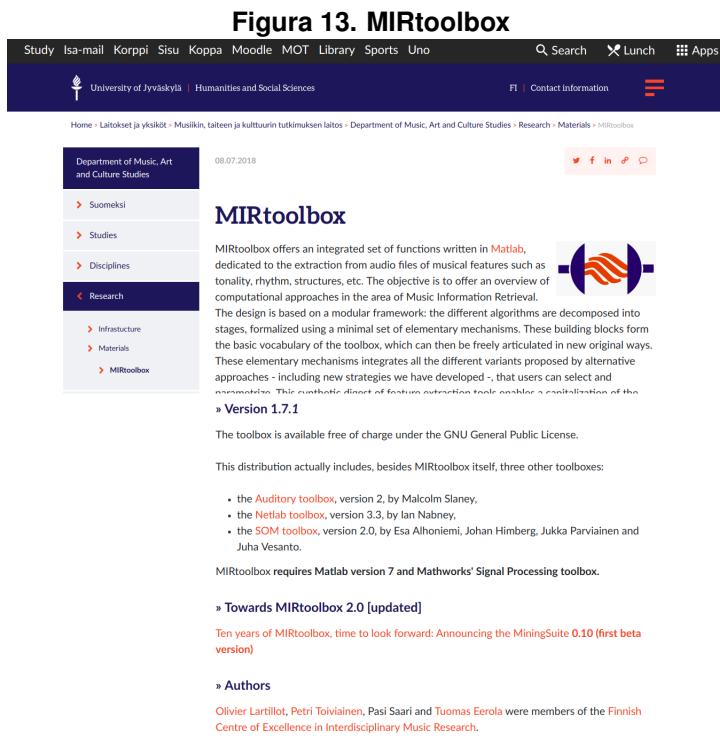
- Quadro modular: É baseado em um conjunto de blocos de construção que podem ser parametrizados, reutilizados, reordenados, etc.;
- Sintaxe simples e adaptativa: Os usuários podem se concentrar no design geral e o MIRtoolbox que cuida das tarefas laboriosas subjacentes;
- Software livre e código-fonte aberto: A ideia é propor a capitalização da expertise da comunidade de pesquisa e oferecê-la de volta à comunidade;
- Recursos: O MIRtoolbox inclui cerca de 50 extratores de recursos de áudio e música e descriptores estatísticos.

O MIRtoolbox é baseado em técnicas de mineração de dados e, por ser um pacote de ferramentas, são variadas as técnicas para recuperação da informação musical, como a classificação.

Desta forma, ele pode ser útil para a comunidade de pesquisa em Recuperação da Informação Musical (MIR), mas também para fins educacionais. Mais informações sobre o projeto podem ser consultadas em [Lartillot and Toiviainen 2007, Lartillot 2013] e [of Jyväskylä 2018]

⁹<http://jmir.sourceforge.net/publications.html>

¹⁰<https://www.jyu.fi/hytk/fi/laitokset/mutku/en/research/materials/mirtoolbox>



Fonte: [of Jyväskylä 2018], elaborado pela autora

2.2.5. MusicMiner

O *Databionic MusicMiner*, desenvolvido como parte de um projeto de pesquisa do Grupo de Pesquisa em Databionics da Universidade de Marburg, na Alemanha, é um navegador para dados musicais baseado em técnicas de mineração de dados, como clusterização e visualização com base no paradigma dos mapas geográficos ESOM (ver Figura 14).

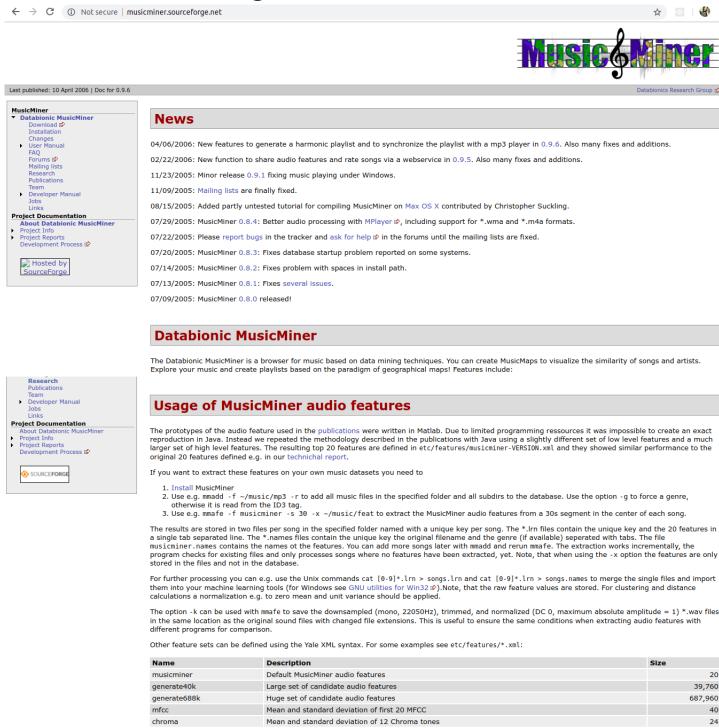
A coleção de músicas é recuperada e apresentada em forma de mapa topográfico com pequenos pontos para as músicas. O usuário pode interagir com o mapa de diferentes formas para selecionar e ouvir músicas; explorar suas músicas; e criar playlists baseadas no paradigma de mapas geográficos.

O site do projeto¹¹ apresenta as seguintes características da solução:

- Análise automática de uma árvore de pastas com arquivos de música (MP3, OGG, WMA, M4A, MP2, WAV);
- Descrição automática de arquivos de áudio digital por som;
- Criação de *MusicMaps* para navegar pelo espaço sonoro com base no paradigma dos mapas geográficos ESOM;
- Criação visual de *playlists*;
- Pesquisa por similaridade na coleção de músicas com base no som;
- Navegação hierárquica personalizável da base de dados, como por exemplo, por gênero/artista/álbum ou ano/artista;
- Base de dados flexível, incluindo o armazenamento separado de vários artistas por música, álbuns e listas de reprodução como parte de uma lista de reprodução;

¹¹<http://musicminer.sourceforge.net/>

Figura 14. MusicMiner



Fonte: [of Marburg 2005], elaborado pela autora

- Importação e exportação de meta informações baseadas em XML.

O *MusicMiner* é escrito em Java para máxima portabilidade e publicado sob os termos da GPL (General Public License). Seu foco principal é a pesquisa e o ensino. Informações mais detalhadas sobre o projeto podem ser encontradas em [Morchen et al. 2005] e [of Marburg 2005].

2.2.6. Tunebot

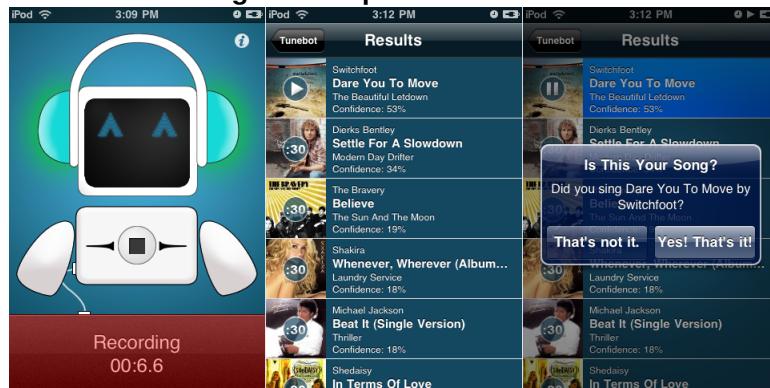
*Tunebot*¹² é um projeto criado em 2015, desenvolvido e mantido por *Interactive Audio Lab*¹³ na Universidade de Northwestern, nos Estados Unidos. Segundo [Huq et al. 2010b], o Tunebot está disponível como um serviço web (ver Figura 16) e está atualmente em teste beta como um aplicativo do iPhone (ver Figura 15). A interação do usuário nas versões da Web e do iPhone é idêntica: (i) cante, e (ii) escolha. O usuário simplesmente canta uma parte da música desejada para o Tunebot e o sistema retorna uma lista ordenada de músicas. Cada música pode ser reproduzida por um simples clique. Enquanto a música está tocando, o sistema apresenta uma caixa de diálogo perguntando se esta é a música correta. Se o usuário clicar em "sim", a consulta será armazenada no banco de dados como um exemplo pesquisável para essa música. O usuário é conectado à Amazon.com ou ao iTunes, onde a música pode ser comprada.

O sistema não exige chaves de pesquisa codificadas manualmente, pois atualiza

¹²<http://music.cs.northwestern.edu/data/tunebot/>

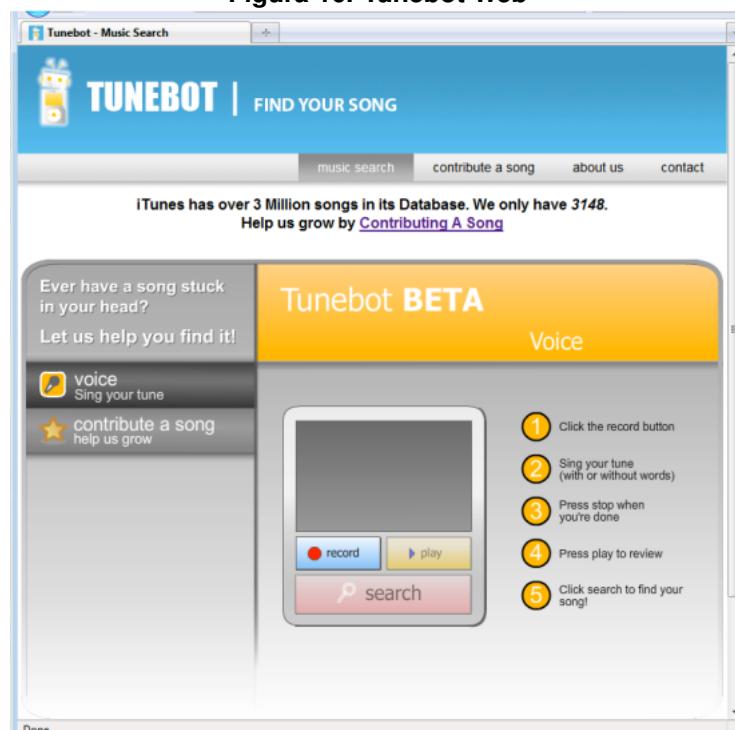
¹³<http://music.eecs.northwestern.edu/>

Figura 15. Aplicativo Tunebot



Fonte: [Lab 2008a], elaborado pela autora

Figura 16. Tunebot Web



Fonte: [Huq et al. 2010b]

automaticamente o banco de dados com novas chaves de pesquisa derivadas de consultas e contribuições do usuário [Huq et al. 2010b]. O banco de dados do *Tunebot* compara as músicas com as músicas cantadas pelos usuários utilizando o algoritmo *query by humming*.

Outro objetivo do projeto é ajudar pesquisadores na área de reconhecimento de músicas que utilizam o algoritmo *query by humming*, facilitando uma pesquisa mais precisa do desempenho do mundo real do que seria possível com conjuntos de dados existentes.

Informações mais detalhadas sobre o projeto podem ser consultadas em

[Huq et al. 2010b, Cartwright and Pardo 2012] e [Huq et al. 2010a].

3. Análise Comparativa

Esta seção apresenta os critérios definidos para a análise das soluções existentes, a análise comparativa propriamente dita e os resultados obtidos com a análise. A metodologia de comparação entre as soluções leva em conta documentações públicas disponíveis que possuam informação pertinente.

3.1. Critérios de Análise

Com base nos conceitos apresentados por [Wazlawick 2012], que define os atributos de qualidade internos, externos e de uso de produtos de software; e pela autora [rer. nat. Christiane Gresse von Wangenheim 2017], que apresenta heurísticas definidas por [Nielsen 1995] para assegurar que os produtos são usáveis, esta seção tem por objetivo apresentar os critérios que devem ser considerados na avaliação das soluções para busca de dados musicais, de modo a compará-las e facilitar a escolha pela mais adequada para uma determinada situação.

Neste trabalho foi realizada uma adaptação de heurísticas para permitir uma boa análise da eficiência e adequação funcional das soluções apresentadas, sendo utilizado informações disponibilizadas nas documentações próprias de cada solução comercial e acadêmica. Já para a análise da usabilidade foi realizada uma adaptação do *MATcH Checklist* disponibilizado pelo Grupo de Qualidade de Software (GQS) da UFSC, onde o conjunto de perguntas possui uma escala de resposta com 2 opções: Sim (a solução atende o objetivo) e Não (a solução não atende o objetivo).

Assim sendo, os seguintes critérios e subcritérios são considerados:

- Eficiência: Trata da otimização do uso de recursos de tempo e espaço. Espera-se que o sistema seja o mais eficiente possível de acordo com o tipo de problema que ele soluciona:
 - *Comportamento em relação ao tempo*: Mede o tempo que o sistema leva para processar suas funções, ou seja, o tempo de reconhecimento/busca de uma música;
 - *Utilização de recursos*: Avalia a complexidade das estratégias e algoritmos utilizados na recuperação de informação musical;
 - *Bitrate*: Mede a qualidade do áudio. Essa qualidade consiste no número médio de bits que será comprimido em um segundo de dados. A unidade utilizada é o KBPS ou 1000 BITS por segundo;
- Adequação Funcional: Mede o grau no qual o produto oferece funções que satisfazem necessidades estabelecidas e implicadas quando o produto é usado sob condições especificadas:
 - *Disponibilidade*: Avalia a disponibilidade da aplicação em diferentes plataformas;
 - *Modelo de desenvolvimento*: Avalia se a solução é de código aberto, dando a possibilidade para que qualquer um consulte, examine ou modifique o produto;
 - *Interações*: Avalia se a solução permite extensões e/ou integrações com outras aplicações;

- *Acessibilidade*: Avalia se a solução possui acesso ao acervo de músicas on-line e/ou off-line;
- *Busca de dados*: Avalia se a solução foi projetada para *matching* exato ou por similaridade (aproximado);
- *Inclusão da dados*: Avalia se a solução permite o envio de músicas feito pelo usuário;
- *Modelo de pagamento*: Avalia o custo da solução, como por exemplo: Gratuito, Pago ou Freemium.
- Usabilidade: Avalia o grau no qual o produto tem atributos que permitem que seja entendido e que seja atraente ao usuário, quando usado sob condições específicas:
 - *Visibilidade do status do sistema*: O sistema deve sempre manter o usuário informado sobre o que está acontecendo. Por exemplo, os componentes interativos selecionados são claramente distintos dos demais?
 - *Prevenção de erros*: Mensagens de erros devem ser claras e objetivas, devem indicar o problema com precisão e sugerir uma solução;
 - *Flexibilidade e eficiência de uso*: A solução deve permitir configuração de ações frequentes. Por exemplo, as funções mais utilizadas são facilmente acessadas?
 - *Estética e Design minimalista*: Mensagens de diálogos não devem conter informações irrelevantes. Por exemplo, o menu é esteticamente simples e claro, com opções fáceis de encontrar, dispostas em uma ordem lógica e com títulos curtos?
 - *Pouca interação homem/dispositivo*: Qualquer informação deve ser fácil de pesquisar e deve ser focada na tarefa do usuário. Por exemplo, a navegação da solução é intuitiva, é fácil chegar à tela desejada?

3.2. Análise

Uma comparação é primeiramente realizada entre as soluções comerciais, seguida das soluções acadêmicas.

3.2.1. Soluções Comerciais

Inicialmente, para a análise das soluções comerciais, foram criadas 3 tabelas, uma para cada grande critério analisado para melhor visualização.

Em relação ao critério de Eficiência mostrado na Tabela 1, o comportamento das soluções em relação ao tempo foram constatados as diferenças verificando o tipo de método utilizado para o reconhecimento/busca de músicas. Os testes foram realizados usando conexão de Internet Wi-Fi e Dados Móveis (4G). Com um smartphone e um cronômetro, foram feitas 5 execuções em cada aplicação com músicas diferentes. E então, calculado a média de todos os tempos para cada aplicação.

Para as soluções utilizando RPC, foi testada a busca de músicas pela correspondência de texto através de metadados (ou seja, através da informação do título, álbum ou gênero da música). Neste caso, a busca demorou até 30s para retornar uma amostra de resultados. Para as soluções que utilizam *Fingerprint* e fazem reconhecimento de música

de forma exata (E), foi testada a busca através da informação de uma parte de áudio "original". Já para as soluções de reconhecimento de músicas de forma aproximada (A), foi testada a busca através da informação de uma parte de áudio "original", do canto ou do cantarolar. Ambas as soluções podem variar de 5s até 13s para retornar uma amostra de resultados.

Cabe observar aqui que uma análise feita apenas sobre o comportamento geral das soluções em relação ao tempo não seria relevante, já que tempos entre 5s e 30s são relativos à situação em que o usuário está no momento. Por exemplo, o usuário pode estar em lugares onde a conexão de Internet não é boa ou quase nula, ou pode estar em lugares onde a conexão de Internet é ótima, já que o reconhecimento de uma música depende de uma conexão de Internet. Quanto melhor a conexão, mais rápido será o reconhecimento da música.

É possível verificar que tanto soluções para o reconhecimento de músicas de forma aproximada, como o MusicID e o ACRCLOUD, quanto soluções para o reconhecimento de forma exata, como o Shazam, Deezer e Musixmatch, podem utilizar o mesmo método de *Fingerprint*. Apesar de utilizarem o mesmo método, a forma como pode ter sido desenvolvido é que gerou a vantagem competitiva dentre os concorrentes do mesmo ramo.

O SoundHound utiliza a tecnologia de Inteligência Artificial (IA) para o reconhecimento de músicas. O desenvolvimento de uma solução baseada em IA se dá através de métodos e algoritmos de aprendizagem de máquina, como Classificação e Clustering, porém não foi encontrada documentação que especificasse o método ou algoritimo utilizado pelo SoundHound para o reconhecimento das músicas. Com essa solução, é possível recuperar músicas através da voz, cantarolando ou até mesmo informando uma parte do áudio "original". Conforme os usuários cantam e conforme a escolha da música do resultado amostrado, a IA associa a cantoria àquela música. Então, quanto mais cantar e buscar as músicas "certas" para a cantoria, a taxa de acerto aumenta, formando uma rede totalmente interligada para o reconhecimento correto de músicas.

Há também as soluções que utilizam o método de Recuperação por Conteúdo (RPC), que seriam as cadeias de caracteres, em buscas, por exemplo, por título, álbum ou gênero da música - o que chamamos de metadados -, conforme já explícitado em seções anteriores deste trabalho. Destas soluções, pode-se citar o Spotify e o SoundCloud. A busca das músicas no Spotify, por exemplo, é feita exclusivamente através de texto, não sendo possível adicionar músicas criadas pelo usuário. Entretanto, é possível criar playlists com as suas músicas preferidas.

Para o Musipedia, não foi encontrada documentação que especificasse o método ou algoritimo utilizado para o reconhecimento das músicas.

Quanto ao critério de Bitrate, das 9 soluções analisadas, foi encontrado essa informação em apenas 4. Diferentes bitrates podem produzir qualidade de som variável. Assim, quanto maior o bitrate, mais vezes por segundo o som original é amostrado, produzindo assim uma reprodução mais autêntica e melhor do som [SanDisk 2018].

O Deezer e Spotify apresentam bitrate de até 320kbps, possuindo 3MB por minuto de áudio e uma qualidade aproximada à de um CD. Já o MusicID e o SoundCloud apresentam bitrate de até 128kbps, possuindo 1MB por minuto de áudio e uma qualidade mínima aceitável em termos de alta fidelidade do áudio [F 2012].

Tabela 1. Análise das soluções comerciais: Critério de Eficiência

CRITÉRIOS	Eficiência		
	Comportamento em relação ao tempo	Utilização de recursos	Bitrate
C MusicID	Até 8s	FP	Até 128kbps
O Shazam	Até 8s	FP	-
M SoundHound	Até 5s	IA	-
E Deezer	Até 10s	RPC e FP	Até 320kbps
R Spotify	Até 30s	RPC	Até 320kbps
C SoundCloud	Até 30s	RPC	Até 128kbps
I Musixmatch	Até 13s	RPC e FP	-
A ACRCloud	Até 5s	FP	-
L Musipedia	-	-	-

Tabela 2. *

Legenda: FP - Fingerprint; IA - Inteligência Artificial; RPC - Recuperação por Conteúdo;

Em relação ao critério de Adequação Funcional mostrado na Tabela 3, das 9 soluções analisadas, 7 são multiplataformas, sendo em sua maioria voltadas para uso em smartphones e 2 de uso exclusivo via Web.

Pode-se verificar que as soluções comerciais, em sua grande maioria, possuem um modelo de desenvolvimento fechado (F), ou seja, seu código não pode ser alterado, mas disponibilizam APIs Web para a comunidade de desenvolvedores, para que a solução possa ser incorporada a seus próprios sites e aplicações.

Quanto à acessibilidade, todas precisam de conexão com a Internet para o uso de suas funcionalidades e apenas 2 das 9 soluções analisadas possuem acesso off-line, que se dá através do download de músicas para a memória interna do dispositivo. As 2 soluções são Spotify e Deezer.

Das 9 soluções comerciais analisadas, 4 possuem reconhecimento de músicas de forma aproximada (A), ou seja, é possível realizar buscas de músicas por similaridade através da voz, sendo que 2 delas permitem cantarolar. Já as 5 demais soluções realizam o reconhecimento de músicas de forma exata (E), sendo necessária uma parte de áudio "original" e/ou uma busca por conteúdo através de metadados.

Apenas 3 soluções permitem a inclusão de músicas criadas por usuários, enquanto as outras 6 permitem a inclusão de novas músicas somente através de contatos com gravadoras e/ou artistas. Destas 3 primeiras, duas reconhecem músicas de forma aproximada: ACRCloud e Musipedia. A primeira é um serviço na nuvem, sendo, até o momento, o maior banco de dados de músicas e um serviço utilizado pela maioria das outras soluções aqui analisadas, como Musixmatch, que permite o reconhecimento de músicas por similaridade, e o Musipedia é uma wikipedia de músicas que aceita contribuições musicais de diversas formas: através da voz, partes de músicas, ou até em formato MIDI (quando o som é criado digitalmente).

Com exceção do ACRCLOUD, todas as demais soluções possuem versões gratuitas para uso, com a possibilidade de pagar uma mensalidade e não ter interrupções e propagandas entre as músicas. Ainda, poucas soluções possuem integrações com outros serviços e/ou aplicações. Com o Shazam, por exemplo, é possível integrar-se ao Spotify e, então, ao encontrar uma música, poder ouvi-la por completo. Da mesma forma para o SoundHound, além de ser possível o compartilhamento da sua pesquisa com o Twitter. Com o Musixmatch integrado ao Spotify, por exemplo, ao encontrar uma música, você acompanha a música com a letra em tempo real. O ACRCLOUD e o Musixmatch são as soluções que possuem o maior número de integrações.

Tabela 3. Análise das soluções comerciais: Critério de Adequação Funcional

CRITÉRIOS	Adequação funcional							
	Disponibilidade	Modelo de desenv.	Integrações	Acessibilidade	Busca de dados	Inclusão de dados	Modelo de Pagamento	
C	MusicID	iOS, Android	F	não	On	A	não	G
O	Shazam	iOS, Android	F	Spotify, Google Music, Apple Music	On	E	não	G
M	SoundHound	iOS, Android	F	Spotify, Youtube, Twitter	On	A	não	F
E	Deezer	iOS, Android, Windows, Web	F	Google Music, Twitter, Facebook	On/Off	E	não	F
R	Spotify	iOS, Android, Windows, Web Linux, OS X	F	não	On/Off	E	não	F
C	SoundCloud	iOS, Android	F	não	On	E	sim	F
I	Musixmatch	iOS, Android	F	Spotify, Deezer, Google Music, Youtube, Twitter, Facebook, Last.fm	On	E	não	F
A	ACRCLOUD	Web	F	Spotify, Deezer, Youtube, ISRC, UPC, LyricFind, Music Story, SyncPower	On	A	sim	P
L	Musipedia	Web	A	não	On	A	sim	G

Tabela 4. *

Legenda: F - Fechado; A - Aberto; E - Busca Exata; A - Busca Aproximada; G - Gratuito; F - Freemium; P - Premium;

Com relação à usabilidade mostrado na Tabela 5, testes foram realizados considerando os critérios e subcritérios descritos na seção 3.1. Foi realizada uma adaptação do *MATCh Checklist* disponibilizado pelo Grupo de Qualidade de Software (GQS) da UFSC, onde o conjunto de perguntas possui uma escala de resposta com 2 opções: Sim

(a solução atende o objetivo) e Não (a solução não atende o objetivo). Cada solução possui particularidades que podem atrair ou afastar o usuário. Os resultados dos testes são discriminados abaixo, feitos em um smartphone, modelo Asus Zenfone 4, com sistema operacional Android.

Todas as soluções são simples, possuem navegação intuitiva e são de fácil execução. As funções mais utilizadas são facilmente acessadas. Os componentes interativos são claramente distintos uns dos outros, com ícones intuitivos. Com exceção do Musipedia, as soluções possuem uma linguagem clara, concisa e funcionam corretamente, não apresentando problemas. Com relação à Musipedia (ver Figura 9), ela é uma enciclopédia de músicas e sua navegação não é muito intuitiva, além de possuir textos longos, não apresentando uma linguagem clara e concisa. Não funciona corretamente, não apresentando mensagens de erros claras e objetivas.

No caso do SoundHound, Deezer e Spotify, o menu é esteticamente simples, mas suas versões gratuitas possuem propagandas, o que polui a tela. Ao se adquirir a versão paga, as propagandas são retiradas. Já o MusicID e o Musixmatch não possuem muitas funcionalidades e o menu é esteticamente simples, claro e sem propagandas.

Deezer, Spotify e SoundCloud executam buscas de músicas pela correspondência de texto (metadados) para identificar arquivos de música digital, ou seja, através da informação do título, álbum ou gênero da música.

O MusicID (ver Figura 1) é de uso exclusivo para o reconhecimento de músicas ouvindo você cantar, cantarolar ou que são tocadas ao seu redor. O mesmo vale para o SoundHound (ver Figura 3), com o diferencial de possuir mais funcionalidades além do reconhecimento de músicas, como a possibilidade de criar playlists e o compartilhamento com o Twitter. Além das informações básicas sobre a música, o aplicativo sugere alguns vídeos que podem ser assistidos diretamente no SoundHound.

O Shazam (ver Figura 2) é similar ao MusicID e ao SoundHound. A diferença é que faz o reconhecimento de músicas apenas tocadas ao seu redor. Ele também possui a funcionalidade de tirar fotos de QRCodes para realizar a busca da música. O Deezer (ver Figura 4) é similar ao Shazam, pois também possui a funcionalidade de reconhecimento de músicas que são tocadas ao seu redor e a possibilidade do download das músicas para uso off-line, além da criação de playlists.

O Spotify (ver Figura 5) é similar ao Deezer em termos de funcionalidades. A diferença é que o Spotify não possui reconhecimento das músicas pelo som (trecho de áudio ou voz). Para usufruir de todas as possibilidades do aplicativo, como o download de músicas para uso off-line, é necessário obter a versão premium. Por sua vez, o SoundCloud (ver Figura 6) é similar ao Spotify, com a diferença de ser o único aplicativo comercial analisado que permite a inclusão de músicas criadas pelos usuários.

O Musixmatch (ver Figura 7) é um aplicativo que sincroniza letras de músicas e também permite o reconhecimento de músicas que são tocadas ao seu redor. Ainda, o ACRCLOUD (ver Figura 8) é um serviço na nuvem, com um grande banco de dados musical. Além de oferecer o serviço de reconhecimento de músicas, ele também permite o reconhecimento de músicas pela web, como um teste do seu serviço.

Tabela 5. Análise das soluções comerciais: Critério de Usabilidade

CRITÉRIOS	Usabilidade				
	Visibilidade do status do sistema	Prevenção de erros	Flexibilidade e eficiência de uso	Estética e Design minimalista	Pouca interação homem/dispositivo
C	MusicID	sim	sim	sim	sim
O	Shazam	sim	sim	sim	sim
M	SoundHound	sim	sim	sim	sim
E	Deezer	sim	sim	sim	sim
R	Spotify	sim	sim	sim	sim
C	SoundCloud	sim	sim	sim	sim
I	Musixmatch	sim	sim	sim	sim
A	ACRCloud	sim	sim	sim	sim
L	Musipedia	sim	não	sim	sim

3.2.2. Soluções Acadêmicas

Em relação ao critério de Eficiência mostrado na Tabela 6, o comportamento das soluções em relação ao tempo foi encontrada documentação apenas para o Tunebot. A solução utiliza QBH e, segundo [Lab 2008b], o comprimento mediano de uma música cantada por um usuário é de cerca de 18 segundos de áudio, levando cerca de 5s para retornar resultados a partir do momento em que a aplicação recebe a música. Para comparação, a música mais longa recebida até o momento tem cerca de 48 segundos e a aplicação demora cerca de 13s para retornar uma resposta. O tempo de retorno é uma função de vários fatores, incluindo o tamanho do banco de dados e o comprimento da consulta, tanto em termos da duração total do áudio quanto do número de notas que o usuário cantou.

Cada solução possui ferramentas variadas para uso distinto no reconhecimento de músicas, o que gera diversas formas de utilização dos recursos. Por exemplo, o MusicMiner (ver Figura 14) possui ferramentas que utilizam os métodos de Classificação e Visualização através de Mapas Musicais Topográficos.

O AMUSE (ver Figura 10) e o Java MIR (ver Figura 12) utilizam o método de Classificação. Já o Tunebot (ver Figura 16 e 15) utiliza Query by Humming (QBH). O CLAM (ver Figura 11) utiliza Spectral Modeling Synthesis (SMS), e o MIRtoolbox (ver Figura 13) é um acoplado de ferramentas baseado em técnicas de mineração de dados. Por ser um pacote de ferramentas, são variadas técnicas para recuperação da informação musical que ele oferece, como por exemplo, clusterização e classificação.

Quanto ao critério de bitrate, das 6 soluções analisadas, não foi encontrada documentação que especificasse o bitrate utilizado para o fluxo de transferência das músicas.

Em relação ao critério de Adequação Funcional mostrado na Tabela 8, as soluções acadêmicas apresentam um modelo de desenvolvimento aberto, ou seja, seu código pode ser alterado por outros usuários, contribuindo para a resolução de problemas existentes nas aplicações. Das 6 soluções analisadas, todas possuem reconhecimento de músicas de forma aproximada (A), ou seja, permitem realizar buscas de músicas através de trechos de áudio e/ou voz e/ou codificação MIDI e/ou por conteúdo (metadados).

Todas as soluções permitem a inclusão de músicas criadas por usuários. Quanto à

Tabela 6. Análise das soluções acadêmicas: Critério de Eficiência

CRITÉRIOS		Eficiência		
		Comportamento em relação ao tempo	Utilização de recursos	Bitrate
A	MusicMiner	-	C e V	-
C	CLAM	-	SMS	-
A	MIRtoolbox	-	Diversas	-
D	AMUSE	-	C	-
Ê	Java MIR	-	C	-
	Tunebot	Até 13s	QBH	-

Tabela 7. *

Legenda: C - Classificação; V - Visualização; SMS - Spectral Modeling Synthesis; QBH - Query by Humming;

disponibilidade, com exceção do Tunebot, todas as soluções estão disponíveis para Linux, Mac OSX e Windows e seu acesso é off-line, ou seja, é necessário o download e instalação do projeto no dispositivo. O download e instalação das soluções é feito de forma gratuita. O Tunebot está disponível na Web e também possui uma versão beta para smartphones iOS, tendo o seu acesso on-line (Ver Figura 15).

O AMUSE e o Tunebot são as duas soluções que possuem integração com outras aplicações. O AMUSE faz parte do rol de ferramentas do MIRtoolbox. Já o Tunebot tem integração com o *Karaoke Callout*¹⁴, que ajuda a construir a base de conhecimento de músicas do Tunebot.

Tabela 8. Análise das soluções acadêmicas: Critério de Adequação Funcional

CRITÉRIOS		Adequação funcional						
		Disponibilidade	Modelo de desenv.	Interações	Acessibilidade	Busca de dados	Inclusão de dados	Modelo de Pagamento
A	MusicMiner	Linux, Mac OSX, Windows	A	não	Off	A	sim	G
C	CLAM	Linux, Mac OSX, Windows	A	não	Off	A	sim	G
A	MIRtoolbox	Linux, Mac OSX, Windows	A	não	Off	A	sim	G
D	AMUSE	-	A	MIRtoolbox	Off	A	sim	G
Ê	Java MIR	Linux, Mac OSX, Windows	A	não	Off	A	sim	G
M	Tunebot	Web, iOS	A	Karaoke Callout	On	A	sim	G
I								
O								

Tabela 9. *

Legenda: F - Fechado; A - Aberto; E - Busca Exata; A - Busca Aproximada; G - Gratuito; F - Freemium; P - Premium;

Como explicitado anteriormente, as soluções analisadas tem seu acesso off-line.

¹⁴*Karaoke Callout* é um jogo de karaokê para a plataforma iOS da Apple que permite aos usuários “tunebot” desafiarem uns aos outros para um duelo de canto [Shamma and Pardo 2006].

Diante disso, para o funcionamento correto de cada aplicação é necessário a instalação de projetos de terceiros. Desta forma, com a impossibilidade de funcionamento das aplicações, não foi possível analisar os critérios de usabilidade das soluções acadêmicas.

3.3. Discussão dos Resultados

Conforme explicado na seção anterior, a análise de soluções foi dividida em dois grupos: soluções comerciais e soluções acadêmicas. Uma análise entre todas elas seria inviável, pois cada grupo se difere no motivo de uso e é voltado a públicos-alvo diferentes. As soluções comerciais são mais utilizadas por usuários finais. Já as soluções acadêmicas são voltadas a usuários pesquisadores da comunidade de MIR.

3.3.1. Soluções Comerciais

Com relação à análise das soluções comerciais, pode-se concluir que as soluções focam na usabilidade visando facilitar o uso para o seu público-alvo, que apenas as utilizam no seu dia-a-dia. Estes usuários não possuem muito interesse em como essas soluções funcionam, desejando apenas que a busca seja rápida e satisfaça os filtros informados.

Os testes realizados de reconhecimento e busca de músicas, por metadados ou por som, demonstraram que as aplicações atenderam bem a todas as funcionalidades oferecidas. Se o objetivo for o reconhecimento de músicas quando não se conhece nenhum valor de metadado, o Shazam e o Deezer são essenciais. Porém, se o objetivo for aquela música que não sai da cabeça e precisar cantar ou cantarolar, então o ideal é usar o SoundHound ou o MusicID. Se o conhecimento da letra da música for indispensável, é possível usar o Musixmatch.

Se o objetivo for playlists e download de músicas, para não ter a necessidade do uso de conexão com a Internet, o Spotify ou o Deezer atendem a demanda. Já o Musipedia, em particular, é uma solução voltada a usuários músicos, que compõem músicas através da codificação MIDI.

Em relação ao tempo, o ideal é a conexão de Internet Wi-Fi. Nestas condições, todas as soluções, exceto o Musipedia, fazem o reconhecimento/busca de músicas no tempo mínimo de 5s. Nesse caso, saber o objetivo do usuário com relação a desempenho é imprescindível para a escolha de qual solução usar.

Pode-se concluir que o método mais utilizado para o reconhecimento de músicas é o *Fingerprint*. Inclusive, o uso do serviço do ACRCloud tem sido incorporado em cada vez mais soluções para o reconhecimento das músicas, como é o caso do Musixmatch e do Deezer. Outro método bastante utilizado é o RPC, principalmente pelas soluções que permitem apenas buscas exatas, onde é feita a correspondência de texto para identificar arquivos de música digital, ou seja, através da informação do título, álbum ou gênero da música.

3.3.2. Soluções Acadêmicas

Com relação à análise realizada para as soluções acadêmicas, pode-se concluir que as soluções focam na otimização e desempenho no que diz respeito ao reconhecimento de

músicas. O desenvolvimento de tais aplicações é voltado para os usuários pesquisadores da comunidade de MIR, com o objetivo de trazer inovação na busca de dados musicais.

Todas as soluções incluem uma estratégia para busca por similaridade (a maioria delas utilizando técnicas de Classificação), são gratuitas e permitem a inclusão de dados musicais em seus repositórios. Ainda, a maioria são soluções off-line e estão disponíveis em diversas plataformas. Por outro lado, poucas soluções oferecem integração com outras soluções ou aplicações que tocam músicas. Inclusive apresentam um modelo de desenvolvimento aberto, ou seja, seu código pode ser alterado por outros usuários.

Houve dificuldade no uso das soluções acadêmicas e por este motivo, não houve a análise dos critérios de usabilidade, além de não ter sido possível o teste prático dos critérios de eficiência de desempenho, o que poderia ter enriquecido ainda mais a presente pesquisa. A documentação existente para cada proposta sugere que o desenvolvimento das soluções acadêmicas geralmente tem o objetivo de validar uma determinada pesquisa, sendo após oferecida gratuitamente à comunidade para futuras pesquisas e o aprimoramento da solução. Desta forma, a usabilidade nem sempre é o foco principal.

Por outro lado, certos processos já foram desenvolvidos e estão sendo utilizados como parte de uma nova solução, como é o caso do projeto AMUSE, que faz parte do rol de ferramentas do MIRtoolbox. Outro caso é o Tunebot, que possui uma versão beta desenvolvida para plataformas iOS, sendo um dos seus objetivos ajudar pesquisadores na área de reconhecimento de músicas que utilizam o algoritmo *Query by Humming*, facilitando uma pesquisa mais precisa do desempenho do mundo real do que seria possível com conjuntos de dados existentes.

Por fim, a escolha de um método para a recuperação da informação musical vai depender do objetivo da solução a ser desenvolvida. Neste sentido, quanto mais aprimorados forem os métodos e algoritmos de soluções acadêmicas, melhores soluções comerciais serão desenvolvidas para a utilização pelo usuário final.

4. Considerações Finais

Este trabalho teve como objetivo analisar soluções para a busca por similaridade de dados musicais, ou seja, soluções que não necessariamente buscam dados musicais apenas através do casamento direto de parâmetros de entrada para a busca, como título da música, palavras-chave ou um áudio como parte da música, mas também através do casamento aproximado (ou similar) destes parâmetros. Além de uma pesquisa aplicada sobre o que são dados musicais, sobre a forma como os dados são tratados e armazenados, para que possam posteriormente ser recuperados e ouvidos no dia-a-dia, este trabalho buscou reunir e comparar informações com o propósito de contribuir com futuros trabalhos que desejam desenvolver soluções para a busca por similaridade de dados musicais.

Pela observação das características analisadas nas Tabelas 1, 3 e 5, as soluções comerciais enfatizam a usabilidade, visando facilitar o uso para o público-alvo, que é o usuário final no seu dia-a-dia. O objetivo destas soluções pode ser o reconhecimento de músicas através de uma parte da música ou da voz, bem como a criação de playlists e download de músicas para uso off-line. O método mais utilizado para o reconhecimento de músicas é o Fingerprint.

As soluções acadêmicas analisadas nas Tabelas 6 e 8, por sua vez, enfatizam o

desempenho no reconhecimento de músicas. O desenvolvimento de tais soluções é destinado principalmente a usuários pesquisadores da comunidade de MIR com o objetivo de trazer inovação na maneira como são recuperados os dados musicais. Todas as propostas incluem uma estratégia para busca por similaridade (a maioria delas utilizando técnicas de Classificação), são gratuitas e permitem a inclusão de dados musicais em seus repositórios. Ainda, a maioria são soluções off-line e estão disponíveis em diversas plataformas. Por outro lado, poucas soluções oferecem integração com outras soluções ou aplicações que tocam músicas.

Como trabalhos futuros, sugere-se, primeiramente, a aquisição dos protótipos das soluções acadêmicas para que uma avaliação de usabilidade destas propostas possa ser realizada. Neste trabalho, devido a restrições de tempo, não foi possível ter acesso a essas soluções. Outra sugestão é um aprofundamento no estudo dos métodos e algoritmos utilizados para a recuperação da informação musical, como *FingerPrint*, Recuperação por Conteúdo e Classificação, bem como uma comparação para verificar os métodos mais eficientes. Por fim, o desenvolvimento e avaliação de soluções para recuperação de informação musical poderia contar com a participação de mais mulheres no programa de mentoria WIMIR¹⁵, incluindo soluções para busca por similaridade.

Referências

- ACRCloud (2015). Acrccloud serviços de reconhecimento automático e confiável na nuvem acrcloud.
- Amatriain, X. (2004). *An Object-Oriented Metamodel for Digital Signal Processing with a focus on Audio and Music*. Phd thesis, universitat Pompeu Fabra, Barcelona.
- Amatriain, X. (2007). Clam: A framework for audio and music application development. *IEEE Softw* 24(1), page 82–85.
- Cartwright, M. and Pardo, B. (2012). Building a music search database using human computation. *Proceedings of the 9th Sound and Music Computing Conference (SMC 2012), Copenhagen, Denmark, July 12-14, 2012*.
- Daquino, F. (2012). A evolução do armazenamento de músicas [infográfico].
- Development, C. (2004). C++ library for audio and music - clam.
- Development, M. (2018?). Musipedia the open music encyclopedia.
- Dortmund, T. (2018?). Amuse - advanced music exploration.
- Downie, J. S. (2003). Music information retrieval (chapter 7). In *Annual Review of Information Science and Technology*, number 37, pages 295 – 340. Blaise Cronin, Medford, NJ: Information Today.
- F, S. (2012). itunes's new 1080p compares well to blu-ray quality, at a fraction of file size.
- Gomes, A. C. F. (2015). Representação da informação musical: uma análise a luz dos requisitos funcionais para dados de autoridade (frad). Graduação, Universidade de Brasilia, Brasilia.

¹⁵<https://wimir.wordpress.com/mentoring-program/>

- Huq, A., Cartwright, M., and Pardo, B. (2010a). Crowdsourcing a real - world on - line query by humming system. *Proceedings of the 7th Sound and Music Computing Conference (SMC 2010), Barcelona, Spain, July 21-24, 2010*.
- Huq, A., Cartwright, M., and Pardo, B. (2010b). Crowdsourcing a real-world on-line query by humming system. In *Proceedings of the SMC 2010 - 7th Sound and Music Computing Conference*, Barcelona, Espanha.
- Lab, I. A. (2008a). Tunebot.
- Lab, I. A. (2008b). Tunebot dataset.
- Lartillot, O. (2013). Mirtoolbox user's guide 1.7.1.
- Lartillot, O. and Toiviainen, P. (2007). Mir in matlab (ii): A toolbox for musical feature extraction from audio. In *Proc. 8th International Conference on Music Information Retrieval (ISMIR 2007)*, page 127–130, Vienna.
- McKay, C. (2010). *Automatic music classification with jMIR*. Phd thesis, CIRM-MT and Marianopolis College, Québec.
- McKay, C. (2018?). jmir.
- McLane, A. (1996). Music as information. In *Annual Review of Information Science and Technology (ARIST)*, volume 31, pages 225 – 262. American Society for Information Science.
- Michels, U. (1992). *Atlas de música*. Madrid: Alianza, 2 edition.
- Morchen, F., Ultsch, A., Nocker, M., and Stamm, C. (2005). Databionic visualization of music collections according to perceptual distance. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR 2005)*, pages 396–403, London, UK.
- Muller, M. (2007). *Information Retrieval for Music and Motion*. ISBN 978-3-540-74047-6. Springer Berlin Heidelberg New York, Bonn, Germany.
- Nielsen, J. (1995). 10 usability heuristics for user interface design.
- of Jyväskylä, U. (2018?). Mirtoolbox.
- of Marburg, U. (2005). Musicminer - databionic musicminer.
- Pacheco, J. and Lopes, R. (2014). Descrição e procura de música. *Comunicação Áudio e Vídeo, Instituto Superior Técnico*.
- rer. nat. Christiane Gresse von Wangenheim, D. (2017). Critérios e recomendações ergonômicas - aula 3.
- SanDisk (2018?). Definição de bitrate.
- Santini, R. M. and de Souza, R. F. (2007). Recuperação da informação de música e a ciência da informação: Tendências e desafios de pesquisa. In *VIII ENANCIB – Encontro Nacional de Pesquisa em Ciência da Informação*, pages 1 – 14, Salvador, Bahia. Organização e Representação do Conhecimento Poster.
- Semidão, R. A. M. (2013). Dados, informação e conhecimento: elementos de análise conceitual. *DataGramZero*, 14(4).

- Shamma, D. and Pardo, B. (2006). Karaoke callout: using social and collaborative cell phone networking for new entertainment modalities and data collection. In *In Proceedings of ACM Multimedia Workshop on Audio and Music Computing for Multimedia (AMCMM 2006)*, Santa Barbara, CA, USA.
- Vatolkin, I., Theimer, W., and Botteck, M. (2010). Amuse (advanced music explorer) – a multitool framework for music data analysis. In *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR 2010)*, page 33–38, Utrecht, Netherlands.
- Wazlawick, R. S. (2012). *Engenharia de Software para Sistemas de Informação: Conceitos e práticas que fazem sentido*. UFSC, Florianópolis.
- Zuben, P. (2004). *Música e tecnologia: o som e seus novos instrumentos*, volume 1. Irmãos Vitale, second edition.