

OTÁVIO PIERI ROSSATO

**REPRESENTAÇÃO E RECUPERAÇÃO DE MÚSICA EM
SISTEMAS DIGITAIS: ESTUDO EXPLORATÓRIO**

São Paulo
2011

OTÁVIO PIERI ROSSATO

**REPRESENTAÇÃO E RECUPERAÇÃO DE MÚSICA EM
SISTEMAS DIGITAIS: ESTUDO EXPLORATÓRIO**

Trabalho de conclusão de curso apresentado ao Departamento de Biblioteconomia e Documentação da Escola de Comunicação e Artes da Universidade de São Paulo como requisito parcial para a obtenção de título de Bacharel em Biblioteconomia

Orientadora: Prof^a Dr^a Nair Yumiko Kobashi

São Paulo
2011

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

ROSSATO, Otávio Pieri

Representação e recuperação de música em sistemas digitais: estudo exploratório / Otávio Pieri Rossato: Nair Yumiko Kobashi (Orientadora). São Paulo, 2011.

Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia) – Departamento de Biblioteconomia e Documentação. Escola de Comunicação e Artes. Universidade de São Paulo.

1. Recuperação de Informação da Música 2. Indexação Automática de Música 3. Música Digital I. Autor II. Título III. Orientadora.

FOLHA DE APROVAÇÃO

Nome: Otávio Pieri Rossato

Título: Representação e recuperação de música em sistemas digitais: estudo exploratório

Trabalho de conclusão de curso apresentado ao Departamento de Biblioteconomia e Documentação da Escola de Comunicação e Artes da Universidade de São Paulo como requisito parcial para a obtenção de título de Bacharel em Biblioteconomia

Banca Examinadora

Presidente da Banca: Prof^a Dr^a Nair Yumiko Kobashi

Prof. Dr.

Instituição: Universidade de São Paulo

Prof. Dr.

Instituição: Universidade de São Paulo

Aprovada em __/__/____

AGRADECIMENTOS

Agradeço ao meu pai Osvaldo Rossato e meu irmão Adonis Rossato pelo apoio e amor incondicional e por sempre acreditarem em mim.

A minha orientadora Nair Kobashi pela sua compreensão, por acreditar no meu potencial e pela contribuição para minha formação ao longo da minha graduação.

Ao Prof. Luis Milanesi pelas sábias provocações.

A Coordenadoria de Assistência Social da Universidade de São Paulo pelo apoio no período mais difícil da minha graduação.

Pela compreensão, apoio e companheirismo dos amigos: Eduardo Kenji Misawa, Daniel “Koala” Ramos, Daniel Salles, Marcelo Botta, Abraão Antunes, Robson “Ashtoffen” Gonçalves, Luiz Menici, Paula Souza e Souza, João Pedro de Quadro Moraes, Bruno Brasil e Débora “Sininho” Carvalho.

ROSSATO, Otávio Pieri. Representação e recuperação de música em sistemas digitais: estudo exploratório. 2011. Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia) – Departamento de Biblioteconomia e Documentação, Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo.

RESUMO

Na segunda metade da década de 80 a indústria fonográfica apresentava-se extremamente concentrada. Entretanto no final da década de 90, além do custo de produção musical estar acessível a músicos e a pequenas gravadoras, o rápido desenvolvimento das tecnologias de comunicação, das redes P2P e do MP3 ofereceram um cenário de distribuição da música a um custo quase-zero. Esse cenário gerou uma rápida expansão do volume de recursos musicais disponíveis, e uma mudança na lógica de mercado da música, da venda massiva de produtos culturais, para o oferecimento de serviços que atendam a demanda de muitos nichos de gostos culturais. Essa transformação social e de mercado demanda por novas formas de organização das coleções de músicas de larga escala. Por isso foi consituída uma agenda internacional de pesquisa em Recuperação de Informação da Música (RIM), que visa oferecer soluções de recuperação e descoberta de músicas. Essa agenda envolve pesquisadores de diversas áreas, como Engenharia de Áudio, Musicologia, Ciência da Computação, Ciência da Informação (CI) e Biblioteconomia. Entretanto não há ainda no Brasil uma preocupação da Biblioteconomia e CI em desenvolver pesquisas em RIM. Por outro lado a comunicação científica da área de RIM está diluída em uma literatura multidisciplinar, o que torna difícil para a neste a formulação de problemas de pesquisa relevantes. Por isso, este trabalho tem como objetivo realizar um estudo exploratório da literatura da RIM a fim de contribuir para futuras pesquisa descritivas e experimentais.

Foram identificadas na literatura duas abordagens principais de representação e recuperação da música, uma que se baseia na inferência de descritores diretamente do sinal de áudio, denominada “baseada-em-conteúdo”, e outra que se baseia na atribuição de descritores fita por pessoas, ou partir de padrões de dados por elas gerados, denominada baseada-em-texto. Os vários tipos de técnicas são apresentados e discutidos considerando suas possíveis vantagens e desvantagens.

Concluimos que pesquisas futuras relacionadas ao desenvolvimento de métodos e procedimentos que visem diminuir as deficiências de vocabulário e de distribuição das representações nos sistemas de Social Tag, ou que se concentre no desenvolvimento eficiente de bases de dados de alta qualidade metodológica, semântica e terminológica para uso em treinamento de algoritmos de aprendizado nos sistemas baseados-em-conteúdo, podem ser de grande relevância para o avanço do conhecimento e dos desenvolvimentos na área de Recuperação de Informação da Música.

ROSSATO, Otávio Pieri. Representation and retrieval of music in digital systems: an exploratory study. 2011. Completion of course work (Bachelor of Library Science) - Department of Library and Documentation, School of Communications and Arts, University of Sao Paulo, Sao Paulo.

ABSTRACT

In the second half of the 80's the music industry was extremely concentrated. However in the late 90's, plus the cost of music production to be accessible to musicians and small labels, the rapid development of communication technologies, P2P networks and MP3 provided a backdrop to distribute music to a near-zero cost. This scenario led to a rapid expansion in the volume of music resources available, and a change in the logic of the music market, the massive sale of cultural products, to providing services that meet the demand of many niches of cultural tastes. This social and market transformation demand for new forms of organization of large-scale music collections. So it was formed an international research agenda on Music Information Retrieval (MIR), which aims to offer solutions for retrieval and discovery of music. This agenda involves researchers from various fields such as Audio Engineering, Musicology, Computer Science, Information Science (IS) and Librarianship. However there is still a poor concern in Brazil's Librarianship and IS in developing research in RIM. On the other hand the communication of scientific area of MIR is diluted in a multidisciplinary literature, which makes difficult for the formulation of relevant research problems. Therefore, this paper aims to conduct an exploratory study of MIR's literature, in order to contribute to future descriptive and experimental research.

Were identified in the literature two main approaches to representation and retrieval of music, one that is based on the inference of descriptors directly from the audio signal, called "content-based", and another based on the assign of descriptors by persons or through patterns from data generated by them, called text-based. The various types of techniques are presented and discussed regarding their possible advantages and disadvantages.

We conclude that future research related to the development of methods and procedures to reduce deficiencies in vocabulary and distribution of descriptors in social tag systems, or to focus on the efficient development of databases of high methodological, semantics and terminology quality, for use on the training of content-based-system's learning-algorithms, can be of great relevance to the advancement of knowledge and developments in the area of Music Information Retrieval.

Índice de Figuras

Figura 1: Modelo de Classificação Automática Baseada-em-Conteúdo.....	29
Figura 2: Framework de Áudio Fingerprint	37
Figura 4: Consulta-por-Sintagma-Musical do Musipedia.org.....	40
Figura 5: Consulta-por-Similaridade do Last.fm	42
Figura 6: Classificação de Gêneros e Humor do AllMusic.com.....	43
Figura 7: Frequencia de Tags para o 1000 primeiros artistas do Last.fm	49

LISTA DE ABREVIATURAS E SIGLAS

AF - Audio Fingerprinting

BPM - Beats por Minuto

DCD - Descoberta de Conhecimento em Dados

GMM - Gaussian Mixture Model

HLF - High-Level Features

KNM - k-nearest neighbor

LLF - Low-Level Features

MD - Mineração de Dados

MFCC - Mel-Frequency Cespstral Coefficient

MGP - Music Genome Project

MIDI - Music Instrument Digital Interface

MLF - Mid-Level features

P2P – Peer-to-Peer

QBH - Query-by-Humming

RcMBC - Recuperação de Música Baseada em Conteúdo

RIM – Recuperação de Informação da Música

RpMBC -Representação de Música Baseada em Conteúdo

RpMBT - Representação de Música Baseada em Texto

ST – Social Tag

SVM - Support Vector Machine

SUMÁRIO

INTRODUÇÃO	10
MÉTODO.....	21
Critérios de Exclusão	23
Procedimentos de análise.....	23
APRESENTAÇÃO DOS RESULTADOS	24
Representação Baseada Em Conteúdo	25
Representação Baseada Em Texto	30
Social Tag	31
Mineração De Dados.....	33
Anotação Musical Por Especialistas	35
Recuperação De Música Em Sistemas Digitais.....	35
Consulta-Por-Fingerprint.....	36
Consulta-Por-Execução	38
Query-By-Humming	38
Consulta-Por-Sintagma-Musical	40
Consulta-Por-Texto	41
Consulta-Por-Similaridade	41
Consulta-Por-Classificação	42
DISCUSSÃO	44
CONCLUSÃO.....	50
BIBLIOGRAFIA	53

INTRODUÇÃO

Na segunda metade da década de 80 a indústria fonográfica apresentava-se extremamente concentrada, com apenas seis empresas detendo 90% do marketshare¹. O negócio da música estava estruturado em basicamente três tipos de propriedade: composição, gravação e mídia (suporte e distribuição); e naquele momento EMI, CBS, BMG, PolyGram, WEA e MCA, conhecidas como Big Six, haviam desenvolvidos modelos de negócio predatórios que visavam controlar e concentrar toda a cadeia de produção da música. De forma vertical, essas empresas assumiram o controle financeiro e administrativo dos processos de composição, produção, gravação, manufatura das mídias e distribuição. De um ponto de vista horizontal, a concentração foi obtida pela compra gravadoras e distribuidoras concorrentes e parceiras, e eventualmente "quebrando" pequenos e médios estúdios através de estratégias de mercado. Também adotaram estratégias trans-midiáticas, investindo na produção de cinema, revistas, videoclipes, roupas, acessório, souvenirs e diversos produtos, transformando cada artista em uma franquia, potencializando as vendas através da geração de demanda trans-mídia (BURNETT, 1996).

Tal concentração reproduziu-se igualmente na mediação do consumo e circulação da música, principalmente através dos gêneros musicais midiáticos, classificações gerais que procuram situar mercadologicamente os produtos e endereçá-los a um público determinado (FIRTH apud JANOTTI JUNIOR, 2003). Esses gêneros não são definidos predominantemente em termos da estrutura da forma musical, como em grande parte da música anterior ao século 20, mas pelo momento de mercado, público potencial e pelas estratégias de criação de sentido vinculados aos produtos. Para FIRTH (apud JANOTTI JUNIOR, 2003), os gêneros musicais podem ser definidos por quatro facetas: convenções musicais, convenções de performance ou imagem, convenções de mercado, e sociais, ligadas a valores e ideologias atribuídos ao produto musical, carregando assim um sentido potencial. Desse ponto de vista, "parte da comunicação dos sentidos e valores expressos pela música

¹ Quota percentual relativa ao total de receitas de um determinado mercado

popular massiva estão inscritas na codificação de gênero, ou seja, os gêneros musicais determinam, em parte, diferentes tipos de julgamentos estéticos, competências diferenciadas para que se construam determinados quadros de valor em relação a certas expressões musicais" (JANOTTI JUNIOR, 2003, p.37).

Esse grupo de empresas, como já mencionado antes, detinha a produção verticalmente, e, portanto, podia exercer certo grau de controle sobre a criação e produção musical a fim de coordená-las a estratégias mercadológicas. Por outro lado investiam na presença ubíqua de seus artistas nas revistas populares, especializadas, TV, cinema e imprensa em geral, além de tornar cada fã em um outdoor ambulante de seus artistas com as vendas de camisetas e outros produtos da sua franquia. Isso para criar um ideário comportamental e ideológico, desejável e atraente, responsável por delinear os atributos de performance e sociais de um determinado gênero. Na ponta dessa cadeia as lojas para o consumidor final reproduziam um modelo de organização a partir desses mesmos gêneros, atendendo assim às convenções tanto do fornecedor quanto dos clientes.

De uma perspectiva sistêmica, essa ação coordenada da cadeia da música, em codificar a matriz de sentidos no interior de gêneros, visa aumentar as chances de sucesso e aceitação do produto musical pelo público alvo.

Para além da estratégia de negócio, esse cenário figura como um ecossistema informacional cuja mediação de conteúdos baseia-se fundamentalmente na ideia de gêneros musicais, e esses por sua vez não necessariamente visam representar a peça musical, mas determinar o tipo de audiência, como conclui SANTINI (2011):

(...) the music industry governs some artistic classification systems with the intention of classifying primarily the audiences, not the music works, thus attempting to reduce the risks of its commercial undertakings. If the boundaries between music genres and public tastes are socially constructed, then this construction tends to be influenced in some degree by the cultural industries, which strive to organize the supply and demand according to market convenience. (p. 240)

Segundo LESSIG (2008) o século XX foi o século de uma cultura R/O ou read-only, termo proveniente da computação, para designar memórias físicas passíveis apenas

de leitura de dados, em contraposição às memórias R/W ou read-write, passíveis de leitura e inscrição. Lessig se utiliza dessa analogia às memórias R/O para caracterizar os comportamentos de produção e consumo da cultura, numa época de reprodutibilidade técnica e massiva. No caso específico da música, ele aponta para a mudança dessas relações com a introdução dos registros sonoros, no início do século XX. Isto porque até então para se ouvir música era necessário estar no lugar onde estava sendo executada por músicos, ou pela própria pessoa. Em outras palavras, era necessário ir até onde estavam os concertos, as bandas, os corais, as fanfarras, as modas, as rodas, ou participar delas. Esse contexto de consumo e produção musical proporcionava um ambiente de aproximação entre o fazer-musical e o consumo, ou seja, grande parte das pessoas participava da produção de alguma forma. Esse cenário foi muito frutífero para o desenvolvimento da música americana, por exemplo, já que uma música folclórica, era interpretada pelos mais distintos sujeitos, e ao fazê-lo, lia e escrevia sua própria cultura (R/W²), diversificando os gêneros musicais.

Com a introdução do registro sonoro e outras formas de reprodução acessíveis, a música passou a ser mediada, por um suporte físico, um documento. Para além da reprodutibilidade, tornava-se possível também delimitar fronteiras de propriedade à algo intangível como a música. Sendo possível de se reclamar propriedade por um produto de cultura antes intangível, também passou a ser possível regular a sua veiculação e reprodução, e dar origem a um novo mercado, ou seja, a indústria fonográfica. Desta forma o pianista do bar foi substituído por uma pianola³, o concerto erudito por um gramofone na sala de jantar, como também a banda de baile. Apesar de soar como uma visão romantizada do processo de industrialização da música, de fato esse cenário afetou de forma importante as relações de produção e consumo da música, que com a evolução da industrialização, tornou a produção cada vez mais profissionalizada, dentro de uma cadeia de produção, e cada vez mais vinculada à um suporte e menos a um fazer.

² Read and Write. Referência ao tipo de memória, na área de computação, que possibilitam que os dados sejam lidos e escritos.

³ Piano que executa músicas automaticamente através de uma programação inscrita em papel perfurado.

Essa visão de Lessig, de certa forma, apoia-se nas concepções de Walter Benjamin (1994) sobre a perda de Autenticidade e Unicidade da obra de arte. Embora BENJAMIM não se refira diretamente à música em seu trabalho *A obra de arte na era de sua reprodutibilidade técnica* os conceitos, os conceitos de Autenticidade e Unicidade, como constituintes da aura da obra de arte, são importantes para o entendimento das transformações nas condições de produção e consumo. Autenticidade para BENJAMIM (1994) refere-se ao o “aqui e agora da obra de arte, sua existência única, no lugar em que ela se encontra”. Antes do registro sonoro, a música só existia no espaço e no tempo de sua execução, no momento preciso em que era tocada, no “aqui e agora”, diferentemente das obras de arte às quais Benjamin se referia, que não se corporificavam materialmente. Já a Unicidade da obra refere-se ao caráter único e tradicional da obra, fundamentada no valor de culto, de sacralização da obra. Esse fundamento teológico confere a obra de arte um caráter transcendental, fora da realidade material e histórica. Essa noção idealista

restringe a atividade de produção da cultura a um número limitado de sujeitos, na medida em que descreve a cultura como uma esfera superior, atingível apenas por aqueles dotados do ‘dom’, a habilidade especial de produzir cultura (ARAUJO, 2010).

Entretanto Benjamin preconizava que a introdução das tecnologias de reprodutibilidade e produção cultural, e a conseqüente destruição da aura, aumentaria o caráter coletivo da criação, subestimando as forças econômicas adjacentes a avanços de tecnologia (ARAUJO, 2010). Não houve, por conta da introdução dessas tecnologias, a superação dessa aura, mas sim a adaptação de seus elementos centrais, que passaram a constituir elementos essenciais ao modo de produção industrial de cultura, como preconizado por Theodor Adorno e Max Horkheimer (1986). O caráter autêntico e único da obra foi mantido, mas na forma da: rigidez da tutela jurídica estabelecida pelo direito autoral sobre as obras artísticas; promoção da ideia de criação individual subjetiva, do “gênio criativo”; promoção da escassez e da originalidade das obras pela indústria; rigidez dos substratos nos quais se fixaram as obras reprodutíveis no decorrer do século XX, até a popularização das tecnologias digitais. (ARAUJO, 1985).

Tal qual uma tecnologia de registro sonoro possibilitou uma profunda mudança nos padrões de produção, consumo e distribuição de música, outra tecnologia traria profundas mudanças novamente no final da década de 90. À parte das tecnologias digitais de compressão de áudio da indústria fonográfica, voltadas ao armazenamento em CDs⁴, a Motion Pictures Expert Group lançava, na primeira metade da década, o formato de compressão de áudio MPEG 1, Audio Layer 3, ou MP3. Esse formato além de possibilitar a codificação de músicas numa qualidade próxima aos dos CDs, mas em arquivos muito menores⁵, é um formato aberto, ou seja, qualquer um pode criar softwares de produção ou reprodução de MP3 (HACKER, 2000).

Para a indústria fonográfica, cópias não autorizadas de seus produtos limitavam-se a cópias privadas, não sendo consideradas um problema importante. Mesmo nos países com um mercado musical ainda não desenvolvido, onde o volume de cópias era importante, na verdade, essa prática não impedia o forte crescimento do mercado; ao contrário, acostumava a audiência à produção musical daquelas indústrias, tornando esses mercados mais rentáveis quando desenvolvidos. (BOURREAU, 2008). É possível dizer que por isso mesmo, a indústria não se preocupou de forma concreta com o desenvolvimento desse formato.

Mas no final da década a Nullsoft desenvolveria o Winamp, um software gratuito de playback de áudio no formato MP3, que juntamente com os modems à velocidades de 56.6Kbps⁶, abriam as possibilidades de redes de distribuição online. Mas o custo de tráfego online para grandes quantidades de dados era muito caro, o que de fato inviabilizava a possibilidade de grandes coleções de MP3, até que em 1999 surgiram as redes peer-to-peer (P2P). Nessas redes não é necessário armazenar os dados em um servidor central, bastando dispor de uma espécie de servidor-mestre que endereça os downloads diretamente para os computadores de quem possui o

⁴ Compact Disc,

⁵ Formatos não comprimidos como os usados em CDs em média codificam 1 minuto de música em 10 Megabytes de dados, enquanto o MP3 codifica 1 minuto em 1 Megabyte (HACKER, 2000, p.9)

⁶ Kilobits por segundo

arquivo. Em outras palavras, transformava o computador de cada usuário em um servidor, e distribuía o tráfego de dados. Em pouco tempo redes como o Napster tinham dezenas de milhões de músicas e de usuários cadastrados. (COLLARD, 2006).

De outro lado, o desenvolvimento da indústria de computadores e softwares diminuiu significativamente o custo das tecnologias de gravação e produção musical, ao proporcionar a possibilidade de se substituir equipamentos especializados e sofisticados, necessários para a produção de matrizes em fitas magnéticas, por computadores pessoais, como aponta NAKANO (2010):

O uso dos computadores pessoais determinou expressiva queda nos investimentos, que passaram a ser acessíveis não só a empresas, como também para indivíduos, permitindo, por exemplo, que os próprios artistas e produtores montassem pequenos estúdios de gravação, os chamados bedroom studios [...], e passassem a gravar com qualidade aceitável para o mercado (p.631)

Ao longo da história da indústria fonográfica nota-se uma relação entre o custo da gravação e reprodução em massa e a diversidade e volume da produção de produtos musicais. No final do século 19, a gravação era realizada em discos de cera, por meio de transcrição direta do sinal, ou seja, para se gravar 10 discos, era necessário executar dez vezes a mesma performance, ou uma vez em 10 gravadores simultâneos, e, nesse último caso, havendo um erro de performance, seria necessário substituir 10 cilindros de cera. Portanto o custo de produção em massa era enorme, e poucas empresas tinham o capital necessário para investir. Esse custo se refletia no volume da produção, que em 1892, era de aproximadamente 320 minutos de música original (ALEXANDER, 1994). Já no começo do século 20, o desenvolvimento de tecnologias que permitiam fazer cópias a partir de uma matriz de gravação (master) diminuiu significativamente o custo de gravação, incluindo gastos com horas de estúdio e de músicos, devido a redução da necessidade da repetição de performances. Desta forma entre 1914 e 1919, o número de empresas de gravação aumentou a um taxa de 40% ao ano, além de terem surgido um grande número de novas gravadoras e selos durante esse período. Da mesma forma, a gravação em fitas magnéticas, que permitia a regravação no mesmo suporte, edição e montagem, causaram a diminuição dos

custos e conseqüentemente o aumento da produção e a diversificação do mercado como aponta GELATT citado por ALEXANDER (1994):

[The] economic attribute of tape recording transcended all others in its effect on phonographic history. Compared to the old method, tape was enticingly cheap. For an investment of a few thousand dollars one could buy a first-class tape recorder. [As a result], between 1949 and 1954 the number of companies in America publishing LP recordings increased from eleven to almost two hundred. (p.119)

Ainda que o custo de produção tenha diminuído com os avanços das tecnologias de gravação, esse cenário não evoluiu para a desconcentração do mercado. Ao contrário, as companhias mais capitalizadas começaram a comprar os distribuidores mais importantes colocando as distribuidoras independentes sob enorme pressão financeiras, levando-as a falência. Dessa forma chegou-se à uma distribuição extremamente concentrada na década de 80, como dito anteriormente. (BURNETT, 1996)

Entretanto no final da década de 90, além dos custos de produção estarem acessíveis à artistas e a pequenas gravadoras, o rápido desenvolvimento das tecnologias de comunicação, da redes P2P e do MP3 ofereceram um cenário de distribuição a um custo quase-zero, apresentando-se como alternativa ao modelo concentrado de distribuição e controle da música. Então, cópia e distribuição não autorizadas se tornaram um ponto importante de preocupação para a indústria, pelas grandes perdas financeiras que provocava. Estima-se que entre 1999 e 2002 a venda anual de música tenha diminuído em 2 bilhões de dólares (STRAUSS apud VACCARO, 2004). Também possibilitou que o número de artistas, selos, produtores e gravadoras independentes, de pequeno e médio porte, ressurgissem de tal forma que já em 2005, os selos independentes somavam mais de 28% do marketshare do mercado mundial (IFPI, 2005). Ou seja, houve uma transferência de poder das grandes gravadoras para os artistas e pequenos produtores (HUGHES, 2003). De 2000 a 2009 mais que dobrou o número de lançamentos anuais de discos em todo mundo, demonstrando que mais do que as vendas, aumentou também a diversidade de produção de músicas.

De acordo com FREEDMAN (2003), além do imperativo tecnológico, a insatisfação dos consumidores com o modelo restritivo da indústria tradicional também

impulsionou essa virada nos padrões de consumo, transformando a troca de arquivos de música umas das atividades mais populares da internet e parte de uma cultura global (VACARO, 2004).

Dessa conjuntura originou-se além de uma explosão da produção, digitalização e distribuição de músicas, também a necessidade da estruturação de novos modelos de negócio da indústria da música (VACARO, 2004). Enquanto a resposta inicial dos grandes conglomerados era constituída basicamente de processos judiciais contra redes P2P e consumidores, e o desenvolvimento de tecnologias digitais de proteção de copyright (FREEDMAN, 2003) durante a primeira metade da década de 2000, algumas organizações começaram a aproveitar as oportunidades que essa mudança do mercado proporcionava. Foi o caso da Apple que em Abril de 2003 lançou o iTunes ⁷ e em Julho de 2004 já havia vendido 100 milhões de músicas digitais (VACARO apud ITUNES TOPS, 2004). Esse modelo de distribuição legal de música digital se desenvolveu e em 2004 o mercado europeu, por exemplo, já contava com mais de 40 serviços de distribuição de música digital online, incluindo o Napster⁸ (VACARO, 2004). Embora o volume de downloads desses serviços possa parecer pouco se comparado ao volume de download ilegal, os modelos de distribuição digital se revelaram como uma alternativa importante e responderam em 2009 por mais de 25% da receita das gravações musicais em todo o mundo (IFPI, 2010).

Outro modelos emergentes são as plataformas de música online que desenvolvem redes sociais baseadas no compartilhamento de interesses musicais, como é o caso do MySpace⁹, Last.fm¹⁰, Blip.fm¹¹, Pandora¹², Spotify¹³, entre inúmeros outros.

⁷ Aplicativo de computador que organiza e gerencia arquivos de multimídia e permite a compra de músicas em formato digital.

⁸ <http://www.napster.com>. Um dos primeiros e mais populares software de P2P.

⁹ <http://www.myspace.com/>

¹⁰ <http://www.lastfm.com.br>

¹¹ <http://blip.fm>

¹² <http://www.pandora.com> (não disponível no Brasil)

¹³ <http://www.spotify.com>

Tanto essas plataformas online, quanto os modelos de distribuição digital se desenvolveram sob um modelo de negócio baseado na noção de música como serviço, diferentemente do modelo tradicional centrado na comercialização de bens físicos, como recomenda VACARO (2004):

Use technology to offer customizability of music compilations selected by consumers in stores and online. All services should enhance ease of use and consumer satisfaction with the process of search, acquisition, and consumption of music delivered via technology and people (e.g., employees, musicians, customer service, etc.). Do research on needs of various consumer segments in regards to process and satisfaction.(p.56)

Em muitas dessas plataformas online, quanto em serviços digitais de distribuição, ou mesmo de venda de itens físicos, como é o caso da Amazon¹⁴, um dos serviços mais importantes é o de recomendação de músicas. Os sistemas de recomendação (SR) são, se acordo com SANTINI (2010), “sistemas informáticos de classificação, organização e recomendação de informação sobre bens culturais baseados nas práticas e gostos dos usuários” que visam principalmente a descoberta de novas músicas pelos usuários. Esses sistemas demonstraram capacidade de promover a substituição do consumo de produtos culturais massivos, ou blockbusters, por outros de nicho (HINZ, 2010), como de reduzir os custos de busca, o que figura como uma importante vantagem competitiva (GHOSE apud HINZ, 2010). Além disso, esses sistemas passaram a desempenhar importante papel na mediação social da música na internet (SANTINI, 2010).

Assim, ao considerarmos que em 2009 mais de um quarto da receita de gravações musicais em todo mundo teve origem em serviços digitais (IFPI, 2010), como os sistemas de recomendação e as distribuidoras digitais, infere-se também um deslocamento não só de marketshare, mas também de mediação social da música, diminuindo assim o poder da indústria tradicional em mediar a organização do consumo através dos gêneros midiáticos. Inferimos como indicativo dessa diluição das convenções em relação aos gêneros musicais que, por exemplo, no Last.fm, que se utiliza da filtragem colaborativa para classificar seus conteúdos, embora os

¹⁴ <http://www.amazon.com>

usuários atribuem tags de gêneros musicais às músicas como forma predominante de classificação, essa tendência não se reproduz no momento da busca (BOOSTELLS et al. apud SANTINI, 2010). Enquanto 68% dos usuários utilizam tags de gênero no momento da classificação, apenas 51% a utilizam no momento da busca, sendo que a maior parte desses 68% que mudaram de estratégia no momento da busca se utilizou de subgêneros e estilos. (BOOSTELLS et al apud SANTINI, 2010). Acreditamos que com a diversificação da produção musical e a diminuição da influência da indústria em estabelecer as convenções sociais dos gêneros, mudou também a percepção de categorização dos consumidores, tornando-se necessário desenvolver novas formas de organização.

Nos novos modelos de negócio, baseados em serviços de oferta de música, a busca e a recomendação de conteúdos desempenham papel fundamental, e ambos demandam formas de organização e classificação eficientes que permitam o acesso a uma quantidade de recursos musicais em escala global como nunca houve (NANOPOULOS, 2009).

Para acompanhar essa demanda por novas formas de recuperação da música, ao longo desse período de transformação também se constituiu uma agenda de pesquisa internacional em Music Information Retrieval ou Recuperação de Informação da Música (RIM), principalmente através da ISMIR¹⁵. Uma vez que a classificação da música pode se dar a partir de facetas acústicas, psicoacústicas, sociais, etc., essa comunidade de pesquisa se estabeleceu sob uma abordagem multidisciplinar envolvendo principalmente os campos de Ciência da Informação, Biblioteconomia, Musicologia, Teoria Musical, Engenharia de Áudio, Ciência da Computação, Direito e Administração (DOWNIE, 2003).

Entretanto a literatura em RIM não apresenta uma origem multidisciplinar, prejudicando assim a comunicação científica do campo, como aponta SANTINI (2007):

(...) não há uma sociedade (inter)disciplinar de RIM; um periódico ou livro-texto fundador onde pessoas interessadas podem adquirir as bases teóricas e

¹⁵ The International Society for Music Information Retrieval

práticas de RIM. Com exceção de alguns pequenos encontros interdisciplinares, muitos pesquisadores estão apresentando seus resultados para membros das suas próprias disciplinas. A literatura de RIM é difícil de ser localizada, lida e estudada, o que dificulta construir e sustentar uma área de pesquisa respeitável, próspera. (p.11)

SANTINI (2007) refere-se ainda a falta de esforços de aproximação dos saberes e da linguagem técnica das diferentes áreas envolvidas, prejudicando a comunicação entre não-pares, além tá presença limitada da área da CI:

“Nota-se que, para uma área de pesquisa que contém “recuperação da informação” em seu nome, o número de trabalhos publicados relacionados com algumas das técnicas da Ciência da Informação é relativamente baixo” (p.10).

O isolamento das áreas também se reproduz na literatura em Ciência da Informação. A quase totalidade dos estudos limita-se à representação bibliográfica de registros sonoros ou partituras, conservação e gestão de acervos de registros sonoros, classificação por folksonomia e estudos de usuário.

Portanto torna-se inviável observar problemas e formular hipóteses consistentes, em uma abordagem multidisciplinar, nesse campo sem antes entender a evolução das técnicas de classificação e organização da música, propostas pelas diversas áreas envolvidas nas pesquisas em RIM, nesses ambientes digitais ao longo desse período de transformação do mercado, consumo e mediação da música.

Por isso, o objetivo deste trabalho é realizar uma revisão de caráter exploratório da literatura sobre técnicas de classificação e recuperação de músicas em sistemas digitais, em um recorte multidisciplinar, ou seja, circunscrevendo as diversas áreas envolvidas na pesquisa em Recuperação de Informação da Música, afim de contribuir para futuras pesquisas descritivas e experimentais. Portanto não temos a pretensão de discutir especificamente as complexidades técnicas envolvidas, mas obter um panorama da RIM, a fim de possibilitar a formulação de hipóteses de pesquisas futuras.

MÉTODO

A fim de colaborar com o avanço do campo da RIM no Brasil optou-se por realizar um estudo exploratório, uma vez que este "tem por finalidade aumentar a familiaridade do pesquisador com o fenômeno ou com o ambiente que pretende investigar, servindo de base para uma pesquisa futura mais precisa" (DENCKER & VIA, 2001, p.59), sendo portanto adequado ao objetivo deste trabalho.

O método de coleta de dados foi escolhido a pesquisa bibliográfica e documental nas seguintes fontes: Library na Information Science Abstracts (LISA)¹⁶; Library, Information Science & Technology Abstract (LISTA)¹⁷; Information Science & Technology Abstracts (ISTA)¹⁸; Web of Science¹⁹; Scielo²⁰; Biblioteca Digital de Teses e Dissertações da USP (BDTDUSP)²¹; DEDALUS (USP)²²; Anais de International Society for Music Information Retrieval Conferece ²³; Anais do Encontro Nacional de Pesquisa em Ciência da Informação e Biblioteconomia (ENANCIB)²⁴ ; e Base de dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI)²⁵.

Foram adotadas diferentes estratégias de busca de acordo com o tipo de base. Nas bases, LISA, LISTA, ISTA, ENANCIB, DEDALUS, Scielo, BDTDUSP e BRAPCI, as

¹⁶ <http://www.csa.com/factsheets/lisa-set-c.php>

¹⁷ Consultada através da plataforma EBSCOHost
(<http://www.ebscohost.com/academic/ebscohost/>)

¹⁸ Consultada através da plataforma EBSCOHost
(<http://www.ebscohost.com/academic/ebscohost/>)

¹⁹ <http://apps.webofknowledge.com>

²⁰ <http://www.scielo.org>

²¹ <http://www.teses.usp.br/>

²² <http://200.144.190.234/F>

²³ <http://ismir2011.ismir.net>

²⁴ <http://www.ancib.org.br/pages/anais-do-enancib.php>

²⁵ <http://www.brapci.ufpr.br/>

estratégias de busca utilizadas foram: *Recuperação AND Informação AND Música; Recuperação AND Música; Classificação AND Música.*²⁶

Para a base Web of Science além das estratégias adotadas para as bases de Biblioteconomia e Ciência da Informação, foi também utilizada a expressão *Query*²⁷ *AND Music*, porque, em pesquisa preliminar, constatou-se que no léxico de áreas como engenharia de áudio e ciência da computação o termo *Query*, que em computação refere-se a uma solicitação precisa para recuperação de informação, apresenta-se como termo equivalente à *Information Retrieval*²⁸ por uma relação de causalidade entre os termos. Por tratar-se de uma base multidisciplinar a pesquisa restringiu-se aos periódicos das áreas de Ciência da Computação, Acústica, Música, Ciência da Informação e Biblioteconomia, Engenharia (Multidisciplinar) e Ciências Sociais (Multidisciplinar)²⁹

Por envolver várias áreas de conhecimento, o volume de publicações relacionadas a RIM é extenso e, como este trabalho não tem por objetivo realizar uma revisão exhaustiva do assunto, mas sim exploratória, foi estabelecido um recorte de período de publicação, entre 2006 e setembro de 2011, com exceção dos Anais do congresso ISMIR.

Os Anais da International Society for Music Information Retrieval Conference não estão indexados em bases de dados e embora fosse possível aplicar as mesmas estratégias adotadas para as outras bases nos sites onde estão publicados (ISMIR.net), através da Busca Avançada do Google³⁰, trata-se de um evento especializado que aborda exclusivamente recuperação de informação da música. Assim, seriam necessárias estratégias mais específicas de busca, incoerentes com o objetivo exploratório deste trabalho. Portanto optou-se por utilizar integralmente os Anais da conferência de 2011.

²⁶ Foram adotadas equivalentes linguísticas de acordo com o idioma da base de dados.

²⁷ Consulta, em língua portuguesa.

²⁸ Recuperação de Informação

²⁹ Classificação de periódicos do Web Of Science

³⁰ <http://www.google.com.br/>

Cr terios de Exclus o

O escopo do trabalho   obter um panorama significativo das t cnicas de classifica o e recupera o de informa o de m sica em ambientes digitais, e portanto n o haveria contribui o ao objetivo do trabalho incluir no corpus pesquisado pesquisas de especificidade limitada somente ao desenvolvimento de um determinado dom nio, ou de assunto adjacentes a representa o e recupera o da informa o de m sica. Por isso, os itens encontrados que atenderem a algum dos cr terios abaixo ser o exclu dos da an lise:

- Abordagem exclusiva de arquiteturas de softwares e elabora o de algoritmos;
- Abordagem exclusiva de fen menos econ micos ligados   RIM;
- Abordagem exclusiva dos processos sociais de valida o simb lica e media o de produtos culturais;
- Abordagem exclusiva de descri o de fen menos ac sticos;
- Representa o e Recupera o de material multim dia, sem foco na m sica;
- Trabalhos de recupera o de partituras e transcri o autom tica;

Procedimentos de an lise

O corpus encontrado ser  analisado atrav s da leitura do resumo dos artigos, a fim de identificar as t cnicas de representa o e recupera o de m sicas.

Em seguida a fim de descrever as t cnicas encontradas e criar uma hip tese de categoriza o de acordo com o atributo ou faceta da m sica que   alvo da indexa o/classifica o e de acordo com o tipo de consulta (query) usado na recupera o da m sica, seguir-se-  uma busca pelos textos completos dos artigos

mais relevantes encontrados, e uma segunda pesquisa bibliográfica, quando necessário para descrever algum dos assuntos encontrados.

APRESENTAÇÃO DOS RESULTADOS

A busca nas bases selecionadas retornou 593 registros bibliográficos, sendo que, aplicados os critérios de exclusão e eliminadas as duplicidades, foram selecionados 161 trabalhos, entre artigos e anais de congresso.

Na primeira fase da análise dessa literatura foram identificadas duas abordagens principais sobre a representação e recuperação de músicas, uma denominada “Content-Based” ou Baseada-em-Conteúdo e outra “Text-Based” ou Baseada-em-Texto. A abordagem baseada em conteúdo desenvolve-se a partir da ideia de que a recuperação de música pode se dar através da descrição de alguns atributos extraídos diretamente do sinal de áudio (ORIO, 2006), ou seja, baseia-se na determinação de atributos acústicos e de características básicas musicais, como melodia e ritmo, para proporcionar a classificação automática de grandes coleções de músicas e sua recuperação.

A abordagem baseada em texto, ou metadados textuais, desenvolve-se na atribuição de metadados de informações editoriais e bibliográficas, como também, e principalmente, representações culturais e subjetivas da música como humor, emoção, gênero e estilo (CASEY, 2008).

Ambas as correntes procuram estabelecer relações de similaridade entre músicas e entre artistas, mas enquanto a baseada-em-conteúdo privilegia as representações acústicas e musicais, a baseada-em-texto foca na dimensão socioemotiva da música, sendo, portanto, complementares.

Representação Baseada Em Conteúdo

No campo da Análise Documentária pode-se definir Indexação como um processo de representação documentária com a finalidade de recuperar informação (CINTRA et al., 2002), ou seja, criar enunciados linguísticos significativos do conteúdos explícitos e implícitos de documentos, que possibilitem estabelecer relações de similaridade entre eles para fins de recuperação. O objetivo da Representação de Música Baseada em Conteúdo (RpMBC) é equivalente ao conceito descrito de Indexação, entretanto difere fundamentalmente quanto a linguagem em que opera. Enquanto a Análise Documentária apoia-se nas teorias da Linguística (CINTRA et al., 2002) para operar com a linguagem verbal, seja natural ou documentária, a RpMBC trabalha sobre a linguagem musical e, portanto, apoia-se na Teoria Musical e na Musicologia para produzir suas representações documentárias. Assim, para melhor compreender a RpMBC é preciso antes deixar explícitos alguns conceitos fundamentais da linguagem musical e de sua terminologia, definidos a seguir de acordo com MED (1996) e ORIO (2006).

Altura - é a percepção da frequência fundamental de uma nota musical, grosso modo, se ela é mais grave, ou mais aguda. A sucessão de diferentes alturas forma a melodia;

Intensidade - é a amplitude, e portanto a energia, das vibrações de um determinado som. Relacionado à força ou ao volume do agente que a produz;

Timbre - é a percepção da qualidade do som, determinada pela espécie do agente que produz o som, ou seja, qual instrumento ou voz que produz o som. É o que nos faz distinguir entre um saxofone e um piano tocando uma mesma nota musical em um mesmo volume;

Ritmo - é a ordem e proporção em que estão dispostos os sons em função do tempo, ou seja, os padrões temporais da presença de sons e silêncio.

Melodia - conjunto de sons musicais dispostos em ordem sucessiva, ou ainda notas musicais dispostas em uma orientação horizontal;

Harmonia - conjunto de sons dispostos simultaneamente, ou ainda notas musicais que ocorrem simultaneamente, ou seja, numa concepção vertical;

Tempo - é a velocidade que uma peça musical é tocada, geralmente expressa em beats³¹ por minuto (BPM) ou, principalmente em partitura por convenções como Largo, Adagio, Allegretto, que não expressam um BPM exato, mas uma faixa de BPMs;

Estrutura - é a dimensão horizontal da música mas em um nível macro, ou seja, os padrões de repetição, disposição de motivos e frases musicais, refrões, pausas, etc. ao longo do tempo;

Acústica - é a contribuição da acústica de uma sala, pós produção de áudio, filtros e equalização para a composição do timbre de um instrumentos;

Orquestração - é a escolha dos instrumentos que serão empregados na execução de cada linha melódica e/ou harmônica de uma peça musical;

A princípio, a combinação de qualquer dessas facetas, enquanto elementos de uma linguagem, seria relevante para descrever uma música. Entretanto, grande parte desses atributos e seus correspondentes acústicos, exceto para especialistas, não são diretamente ligados às propriedades intrínsecas da música percebidas por humanos. Em geral, a forma como as pessoas entendem e interpretam música estão mais ligados a rótulos de gênero e estilo musical, humor e emoção relacionada a uma determinada música. Assim grande parte do propósito da RpMBC é inferir relações entre esse atributos acústico-musicais e rótulos semânticos significativos para as pessoas. (FU, 2011)

³¹ Beat, ou pulso, é a unidade básica de tempo na música

Por outro lado, do ponto de vista da engenharia de áudio, existem atributos mais simples de serem extraídos e analisados e outros mais complexos e dependentes de alta capacidade de processamento de hardware. Considerando ainda que assim como no campo da Análise Documentária, a escolha dos enunciados que representam um documento e sua classificação dependem das necessidades informacionais presentes em um determinado contexto de uso do documento, este ou aquele atributo pode ser mais adequado ou menos a um determinado contexto informacional. Embora não haja na literatura coletada discussões propriamente sobre contextos informacionais e contextos de uso, os diversos trabalhos abordam a eficiência e eficácia de determinado atributo na recuperação e descoberta de músicas.

Os atributos mais simples de se extrair são classificados como *Low-Level Features* (LLF)³². Embora esses atributos sejam amplamente utilizados na classificação de músicas (FU, 2011), eles são descrições em nível de sinal de áudio, portanto não necessariamente determinam o conceito musical abstrato codificado no sinal digital. Ou seja, são medidas baseadas na análise do espectro de áudio, em geral com amostragens de áudio de 10 a 100ms de duração (FU, 2001), por isso denominados também de “Short-Term Features”³³. A medida mais utilizada para a definição de LLFs é o “*Mel-Frequency Cepstral Coefficient*” (MFCC) (KNEE, 2007). A definição e uso desse coeficiente é altamente técnica e especializada, mas pode ser genericamente definida como um coeficiente relacionado à distribuição de amplitudes e frequências de um sinal de áudio ao longo do tempo, com base na percepção humana dos sons (LOGAN, 2000).

As classes mais importantes de LLF são relacionadas ao timbre da música, e tem se mostrado apropriadas em identificar músicas, similaridade acústicas e de orquestração entre músicas. Por isso têm sido utilizadas, entre outros, na classificação automática de gêneros (FU, 2011, HOLPZAPFEL, 2008), principalmente quando associadas a atributos de nível superiores (BOGDANOV, 2011).

³² Atributo de baixo-nível, em língua portuguesa

³³ Atributos de curta-duração, em língua portuguesa

Outras classe de atributos são os Mid-Level features (MLF)³⁴, que se aproximam mais das qualidades musicais percebidas e apreciadas pelas pessoas, como Ritmo, Altura, Melodia e Harmonia (FU, 2011; ORIO, 2006). Extraí-los e processá-los envolve tarefas mais complexas porque além demandarem amostragens maiores de áudio, sua determinação envolve maior consideração à aspectos psicoacústicos complexos. A altura de uma nota está mais ligada ao julgamento psicoacústico, ou seja, do ouvido e do cérebro do indivíduo, do que propriamente a medidas objetivas da nota fundamental de um sinal. (MACY apud FU, 2011). Determinar aspectos da harmonia, por sua vez, que é feita basicamente através da identificação de acordes, que depende da identificação das alturas das notas. Por isso mesmo o ritmo é o MLF mais usado para classificações baseadas em áudio, principalmente na inferência do humor da música (FU, 2011; LU, 2006). Por outro lado os atributos de Altura e Harmonia são mais importantes na determinação de similaridade e a identificação de músicas no nível melódico (FU, 2011). A partir da combinação desses atributos, extraídos de vários frames ao longo do áudio, combinado a modelos estatísticos de reconhecimento de padrão, é possível definir similaridades entre as Estruturas das músicas.

Os atributos descritos até agora não figuram por si só como formas de classificação e indexação. É necessário ter classificadores e indexadores, no caso da RpMBC, automáticos. Por isso esses atributos são combinados e aplicados em um conjunto de dados de treinamento a fim de gerar a “aprendizagem” dos classificadores automáticos. No caso da classificação, por exemplo, a partir de um conjunto de músicas já classificadas, desenvolvido especificamente para o treinamento dos algoritmos, é possível desenvolver regras de classificação com base nos atributos escolhidos, a fim de prever, com o máximo de eficácia, a classificação de músicas fora da base de treinamento, ou seja, sem rótulos de classificação. (FU, 2011). Portanto, podemos dizer que, de forma genérica, a extração de atributos do áudio, combinado às base de treinamento, e algoritmos de aprendizado, resultam no

³⁴ Atributos de nível-médio, em língua portuguesa

desenvolvimento de classificadores automáticos de músicas, conforme *Figura 1*. Esse processo também é conhecido como “Content-Based AutoTag”³⁵.

Os principais algoritmos de aprendizado encontrados foram são a Support Vector Machine (SVM), Gaussian Mixture Model (GMM) e k-nearest neighbor (KNM), que podem ser definidos como modelos estatísticos de aprendizado (MANDEL, 2006).

Há dois tipos de classificadores, os single-label, que atribuem um único rótulo para cada música e portanto desenvolvidos para tarefa de classificar itens, e os multi-label, que atribuem um conjunto de rótulos, ou descritores para cada item. Na área de Recuperação de Informação da Música, esse processo de atribuição de um conjunto de descritores é denominado Music Annotation, ou anotação de música, que é equivalente ao processo de Indexação no campo da Análise Documentária.

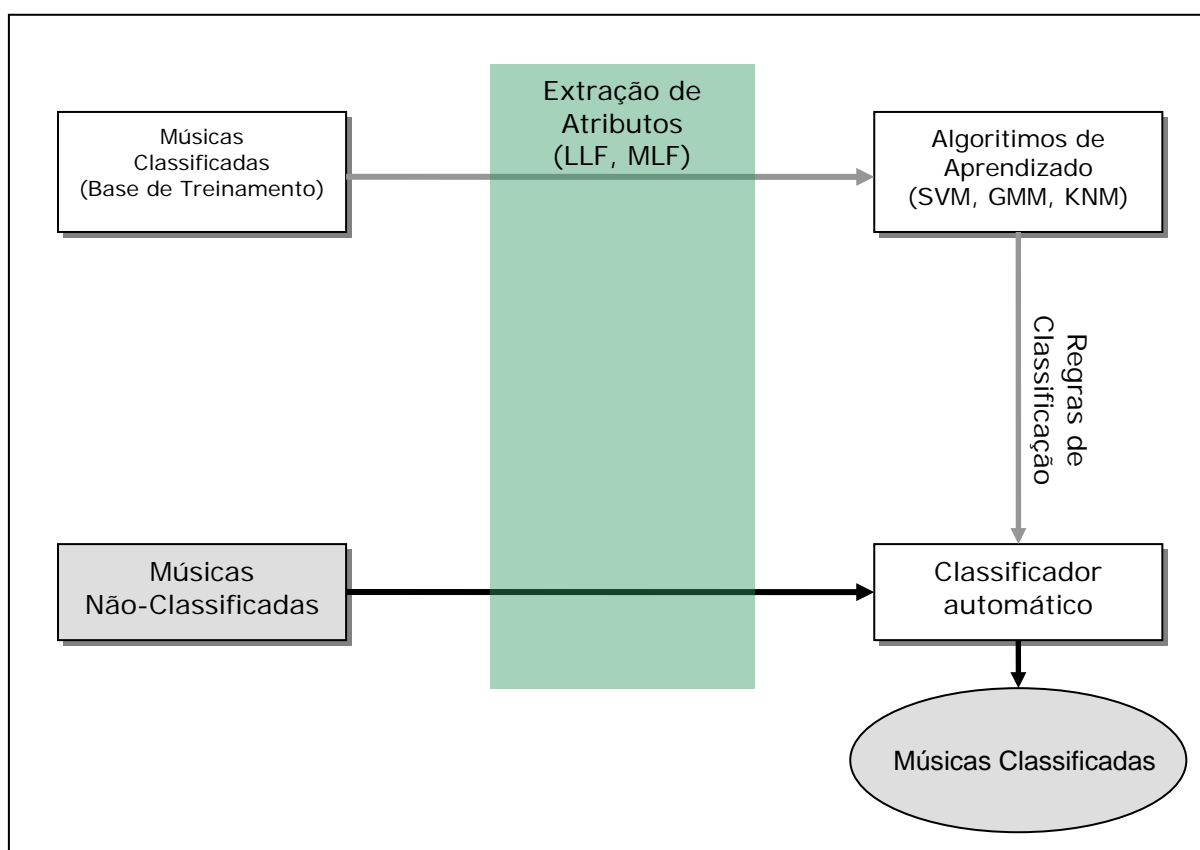


Figura 1: Modelo de Classificação Automática Baseada-em-Conteúdo

³⁵

Taggeamento automático baseado no conteúdo do sinal de áudio

Representação Baseada Em Texto

A Representação de Música Baseada em Texto (RpMBT) procura criar representações documentárias de músicas com rótulos e descritores atribuídos por pessoas ou por autômatos, a partir de documentos em linguagem verbal. Em contraste com a RpMBC, não extrai diretamente do sinal de áudio as inferências de significado, e sim atribui significados às músicas que podem estar relacionados ao seu conteúdo musical, bem como contexto social, emotivo, de uso, de preferências e culturais em geral.

Dado o volume das coleções de músicas na internet a RpMBT procura desenvolver formas automáticas de elaboração dessas representações a fim de proporcionar a organização e recuperação de música em escala global.

Uma vez que não opera diretamente com a linguagem musical, e sim com a verbal, a RpMBT tem uma aproximação maior, visto sua finalidade, com a Análise Documentária. Entretanto, os trabalhos encontrados diferem dos procedimentos metodológicos da Análise Documentária, pois em geral não se baseiam em Linguagens Documentárias, ou seja, em linguagem construídas especialmente para organizar e facilitar o processo de comunicação documentária (LARA, 2004), mas sim, operam na esfera da linguagem natural. Em geral os rótulos e descritores são atribuídos diretamente por usuários, procedimento que no campo da Biblioteconomia “tendem a ser uma instância complementar de indexação dos sistemas de informação”. (KOBASHI, 2007, sem página)

Nas abordagens encontradas na literatura, para além da representação descritiva das músicas a partir de metadados como autor, nome da música, compositor, etc., predomina a preocupação com a descrição temática das música, principalmente relacionada a três facetas: emoção ou humor da música, que se refere à dimensão do sentimento evocado pela audição de uma determinada música, como alegria, tristeza, descontração, etc; contexto funcional de consumo da música, como

músicas para estudar, para festas, meditação, etc.; gênero, estilo ou subgêneros musicais.

A seguir são apresentadas técnicas de classificação e indexação de músicas baseada em texto.

Social Tag

Também descrito como “folksonomia” e “filtragem colaborativa”, Social Tag (ST) é uma prática e uma tecnologia empregada em sistemas cujas unidades denominativas são atribuídas pelos próprios usuários, ou seja, é a própria rede de usuários que indexa seus conteúdos, através da atribuição de tags à músicas, artistas, álbuns e playlists. Tags são *“palavras-chave, rótulos ou breve descrições criadas pelos usuários para classificar conteúdos na internet a partir de vocabulário próprio”* (SANTINI, 2010, sem página) e embora os usuários, obviamente, atribuam tags de maneira predominantemente subjetivas, quando as tags de milhares de usuários são combinadas é possível obter uma visão rica, complexa e multifacetadas dos documentos indexados (LAMERE, 2008).

As representações empregadas não figuram somente como atributos ligados ao documento, mas também expressam opiniões e experiências dos usuários, como forma de explicitar a relevância dos conteúdos (SANTINI, 2010). Além disso os usuários de sistemas ST não só participam de uma rede de filtragem colaborativa de informação, como, de fato, “taggeam” os conteúdos de acordo com suas preferências pessoais, figurando assim como uma forma de recuperação personalizada de informação. Essa personalização dos comportamentos de busca também torna possível estabelecer perfis de busca e gosto dos usuários, bem como tendências, tornando possível não somente propiciar recuperação, mas uma forma eficiente de descoberta e recomendação de músicas (LEE, 2010). Por isso, a filtragem colaborativa costuma estar associada a Sistemas de Recomendação de Música.

Outra característica dessas redes de usuários é a potência na atividade de classificação. Em geral são compostas por milhares de usuários, adeptos do “taggeamento” como uma prática social (LAMERE, 2008). Um sistema como o do Last.Fm, por exemplo, recebe 2 milhões de tags por mês (PAMPALK apud LAMERE, 2008). Esse volume de usuários atribuindo tags também faz com que a representação nesses sistemas sejam bastante fluida e dinâmicas, em contraposição ao perfil estáticos de sistemas com linguagens estruturadas.

Entretanto apesar desse modelo “crowdsourced”³⁶ apresentar certa potência na tarefa de taggeamento, em geral apresenta a dificuldade em consolidar os enunciados nominativos, apresentando grande diluição de tags, e portanto diminuindo significativamente a precisão do sistema. Até setembro de 2008, o Last.fm havia coletado 25 milhões de tags para músicas e 20 milhões de para artista, utilizando-se 1.2 milhões de tags únicas, entretanto, apenas cerca de 11% dessas tags foram usadas mais de 10 vezes (KIM, 2009)

Os sistemas ST possuem arquiteturas distintas que se adequam melhor a um ou outro objetivo de recuperação e recomendação. Assim algumas características principais devem ser observadas na elaboração desses sistemas, que segundo LAMERE (2008) são:

Direito de Taggeamento - alguns sistemas permitem ao usuário atribuir tags apenas ao conteúdo que ele mesmo contribuiu, ou seja, que fez o upload ³⁷. Esse tipo de política é mais frequente em sites com grande quantidade de conteúdo gerados por usuários, como é o caso do Youtube³⁸ e do Flickr³⁹. Outros permitem que qualquer usuário de sua rede atribua tags a qualquer conteúdo. Sites que tem como foco a descoberta de novas músicas, em geral adotam esse tipo de procedimento como é o caso do Last.Fm.

³⁶ Modelo de produção que utiliza a inteligência coletiva de pessoas espalhadas pela Internet

³⁷ Transferência de um arquivo de um disco de memória local para servidores da World Wide Web

³⁸ <http://www.youtube.com>

³⁹ <http://www.flickr.com>

Suporte para Taggeamento - o sistema pode sugerir as tags que já foram utilizadas por um determinado usuário para indexar um determinado artista, pode sugerir as tags mais popularmente usadas para um determinado artista, ou não ter nenhuma espécie de apoio. Sistemas que oferecem sugestões de tag tendem a consolidar melhor um grupo de tags para um determinado conteúdo do que os que não oferecem suporte, que tendem a ter indexações mais difusas;

Agregação - Alguns sistemas permitem que uma mesma tag seja adicionada várias vezes a um único item, outros limitam de forma que uma tag só pode ser adicionada uma vez a um item;

Tipo de objeto - os objetos passíveis de taggeamento como o artista, o álbum, a música, uma playlist⁴⁰;

Observa-se nessas características, principalmente o Direito e o Suporte de taggeamento, uma preocupação em tornar menos difuso os enunciados dos itens e diminuir a ambiguidade e erros ortográficos nas tags a fim de minimizar a principal falha desse tipo de sistema que é a baixa precisão na recuperação de informação (MCGREGOR, 2006).

Mineração De Dados

Mineração de Dados (MD), ou Data Mining é uma disciplina estabelecida na área da Ciência da Computação como um subprocesso, de um processo maior chamado de Knowledge Discovery in Data, ou Descoberta de Conhecimento em Dados (DCD). A DCD é entendida como o *“processo de identificar padrões em dados, válidos, novos, potencialmente úteis e em última instância, inteligíveis”* (FAYYAD apud COENEN, 2011, p.25; tradução nossa), composta por três processos: a preparação de dados, a descoberta de informação desconhecida, ou a MD, e a análise e visualização dos dados. Portanto embora tecnicamente a MD seja um

⁴⁰ Lista de execução de músicas

subprocesso da DCD, para propósitos práticos, elas podem ser entendidas, como sinônimos (COENEN, 2011).

No caso específico da RIM, a MD tem como propósito extrair de conjuntos de documentos, padrões de dados e informações relevantes para a indexação de documentos musicais. Na literatura encontrada, as técnicas de mineração ocorrem predominantemente em documentos disponíveis na Web, e por isso são genericamente denominadas como “Web-based Mining Techniques”, ou Técnicas de Mineração Baseadas na Web, que tem como principal objetivo estimar relações de similaridade entre artista, associar termos descritivos, membros de uma banda, orquestração, capas de álbum, etc. (SCHEDL, 2011)

Os trabalhos encontrados usam diferentes abordagens quanto a técnica de coleta e o conjunto de dados que mineram. Mas de modo geral, baseiam-se em correlações textuais e de co-ocorrências de músicas e artistas presentes em sites de revisões musicais (WHITMAN apud TURNBULL, 2008a), sistemas de social tag, perfil de preferências de usuários, playlists de rádios (PACHET (2001) apud SCHEDL, 2011), histórico de busca em buscadores Web (KNEES, 2008) e páginas da Web em geral relacionadas à música (SCHEDL, 2005).

Entretanto embora a MD se mostre relevante em proporcionar representações do contexto social dos documentos musicais, ela se mantém restrita ao nível de dados de artista, uma vez que há usualmente poucos dados disponíveis no nível das músicas, individualmente. Por isso a MD têm se mostrado mais promissora quando associada a informações baseadas em áudio, como os LLF. (SCHEDL, 2011). Outro revés dessa abordagem é a quantidade de informação irrelevante e de ruído no processo de mineração de textos relacionados (TURNBULL, 2008b; SCHEDL, 2011).

Anotação Musical Por Especialistas

Dos métodos de RpMBT encontrados esse é o mais próximo dos procedimentos metodológicos de representação documentária para fins de recuperação estabelecidos na área de Biblioteconomia e Ciência da Informação. Isto porque se utiliza de vocabulários controlados para realizar a anotação⁴¹, das músicas e artistas. O caso mais notório desse tipo de abordagem é o “Music Genome Project” (MGP) do Pandora*. O MGP é composto de aproximadamente 50 especialistas em Música, treinados para anotar as músicas a partir de um vocabulário controlado de 150 a 500 tags, dependendo do gênero musical, com inter-revisões entre os indexadores (TURNBULL, 2008b). Embora esse processo produza uma indexação de alta qualidade, é uma abordagem de alto custo e que requer grande quantidade de recursos humanos, visto que cada música leva de 20 a 30 minutos para ser indexada. Em 2007 havia 600,000 músicas anotadas, e em 2008 a média era de 15,000 músicas por mês, entretanto em uma base que já possuía milhões de músicas e por isso o MGP também precisa ser bastante seletivo em relação as músicas que são anotadas. (TURNBULL, 2008b)

Recuperação De Música Em Sistemas Digitais

A Recuperação de Música Baseada em Conteúdo (RcMBC), procura desenvolver formas de recuperação alternativas à tradicional, baseada em texto. Na recuperação baseada em texto os documentos são indexados com palavras-chave. Entretanto o usuário pode não saber nenhuma palavra-chave associada à música que procura, como autor, ou nome da música (JINHEE, 2008). Ao considerarmos a abundância de recursos musicais disponíveis digitalmente, é natural que muitas vezes,

⁴¹ Anotação Musical consiste em atribuir alguns termos ou pequenos enunciados que descrevam a música, para fins de recuperação, muito similar portanto ao proceso ao entendimento de Indexção, na área de Análise Documentária

principalmente o usuário não especializado, não disponha de informações verbais suficientes para formular uma expressão de busca que o permita recuperar determinada música ou artista desejado, mas se lembre, por exemplo, da melodia da música. Nesse caso abordagens baseadas em conteúdo são importantes para a recuperação como aponta Sinanan (2010):

“Content-based information retrieval uses a piece of the sought after item as a query instead of using metadata such as name or composer. This is useful when information such as creator, title or publisher is not known.” (sem página)

Por outro lado a RcMBT oferece representações semanticamente mais significativas, por estarem mais próximas da percepção e dos processos de significação das pessoas em relação as músicas. As pessoas tendem a entender, classificar e procurar por músicas através de conceitos altamente abstratos, denominados High-Level Features (HLF) ⁴², como o gênero e emoção de uma música. Embora haja esforços na RpMBT para desenvolver classificações e descritores semanticamente relevantes para as pessoas, ainda há um grande hiato entre a computação de atributos de áudio, e o significado deles para as pessoas (FU, 2011). Entre a potência e a objetividade da representação e recuperação dos métodos propostos pela abordagem Baseada-em-Conteúdo, e o alto grau de relevância semântica produzido pelas abordagens Baseada-em-Texto, há o objetivo comum em desenvolver tecnologias e sistemas que ajudem às pessoas a encontrar e descobrir músicas em um ambiente de abundância de recursos musicais. Da soma desse esforços surgiram formas diversas de recuperação de músicas, apresentadas a seguir, organizadas a partir do ponto de vista da modalidade de consulta, ou seja, de como a consulta é executada no sistema .

Consulta-Por-Fingerprint

Na abordagem da Consulta-por-fingerprint, um trecho do áudio de uma música é utilizado como expressão de busca, para recuperar a música exata correspondente

⁴² Atributos de alto-nível

ao trecho. A técnica de Audio Fingerprinting (AF), ou “impressão digital” do áudio, “consiste em derivar de um fragmento de áudio uma representação compacta da música que pode ser eficientemente comparada à outros fragmentos” (CHANDRASEKHA, p. 801, 2011). Essa representação em geral é feita na forma de um espectrograma de áudio, que é uma representação das frequências musicais em função do tempo. Dito de outra forma é uma representação das características acústicas, utilizada para encontrar músicas com características idênticas (CHANDRASEKHA, 2011). A “fingerprint” da música se baseia na extração de LLFs, em geral MFCC, aplicados tanto para base de músicas quanto na expressão de busca (Figura 2).

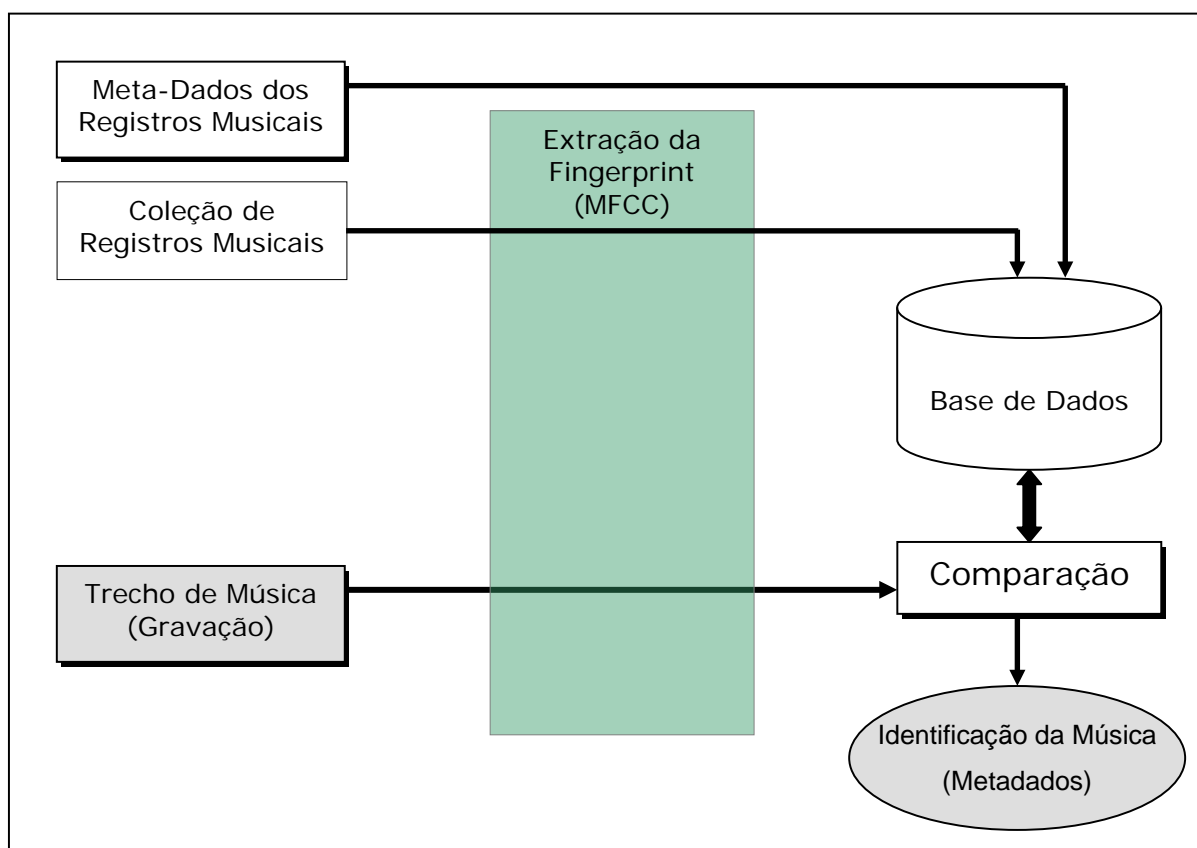


Figura 2: Framework de Áudio Fingerprint (adaptado de CANO, 2005, p.274)

As aplicações mais comuns da AF têm sido em: sistemas de identificação de músicas para celulares e dispositivos móveis de comunicação, como o caso do

Shazam⁴³ e do Midomi⁴⁴, aplicativos que identificam músicas através de pequenos trechos captados pelo microfone do celular (COLVIN, 2009); detecção de material protegido por copyright (GOMES, 2003; BATTLE 2004)

Consulta-Por-Execução

Essa abordagem baseia-se em uma expressão de busca ligada ao conteúdo musical executada por humanos. Por exemplo, uma pessoa cantando, ou assoviando uma música, ou executando uma sequência de notas em teclado virtual. A seguir são descritas as principais formas encontradas desse tipo de consulta.

Query-By-Humming

A abordagem do Query-by-Humming (QBH)⁴⁵ consiste em recuperar músicas que contenham melodias similares, ou idêntica, à uma melodia cantada ou assobiada em um microfone pelo usuário (JINHEE, 2008).

Esses sistemas podem ser considerados como a integração de três estágios: a construção da base de dados de música; a transcrição do áudio cantado do usuário em uma informação melódica (expressão de busca); a comparação, ou matching, da expressão de busca com a base de dados, para recuperação. (BANDERA, 2011).

Os sistemas de QBH utilizam representações simbólicas das músicas, em geral de formato MIDI⁴⁶, para codificar sua base de dados, bem como o áudio cantado pelo usuário (BIRMINGHAM, 2006; SINANAN, 2010). Assim, a comparação entre a expressão e a bases de dados não se dá em nível de sinal de áudio, mas sim entre representações codificadas em MIDI. Ainda sobre o estágio de comparação, há um consenso na literatura de que ter um ou mais índices que podem ser pesquisados

⁴³ <http://www.shazam.com>

⁴⁴ <http://www.midomi.com>

⁴⁵ Aproximadamente, Consulta-por-sussurro, em língua portuguesa.

⁴⁶ O formato MIDI (Music Instrument Digital Interface) contém informações detalhadas quanto à notas musicais, sua ordem de execução, bem como a duração e o volume de cada uma delas, codificadas em uma notação numérica simples.

antes de entrar no banco de dados completo por trás do sistema é útil para tornar o processamento das pesquisas mais rápido (SINANAN, 2010)

As proposições de QBH presentes na literatura variam bastante em relação ao desenvolvimento dos índices, técnicas de transcrição da expressão de busca, e de comparação, porém, pode-se definir genericamente a arquitetura desse sistema, de acordo com a *Figura 3*.

Alguns trabalhos, entretanto, propõem sistemas de QBH a partir da extração e comparação de MLFs (GUO, 2008; SONG, 2002), ou de áudio fingerprints, tanto do áudio como da expressão de busca, independentemente assim de representação simbólicas, como o MIDI (BANDERA, 2011).

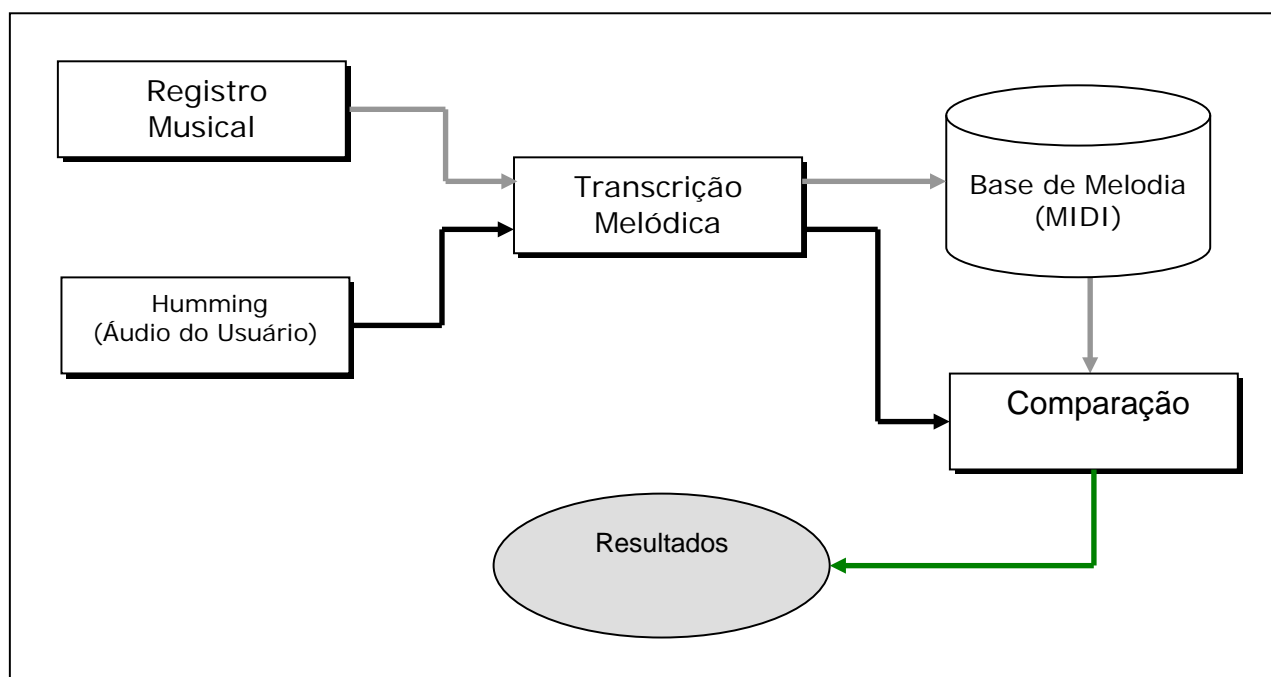


Figura 3: Framework de Query-by-Humming (adaptado de Rayynane, 2008, p.2249)

Consulta-Par-Sintagma-Musical

Esse tipo de recuperação utiliza como expressão de busca a execução ou escrita musical de uma frase musical realizada pelo usuário. A frase é executada em teclados virtuais simulacros de pianos, teclados musicais conectados ao computador ou mesmo no próprio teclado do computador (IRWIN, 2008). Em geral baseiam-se na comparação melódica entre a expressão dada e um banco de dados (TYPKE, 2007). Podem ainda basear-se na similaridade rítmica, técnica conhecida como “query-by-tapping”, proposta por Eisenberg (2004), que consiste em executar um padrão rítmico no teclado de um computador, ou ainda na similaridade rítmica-melódica (KEN, 2008; MUSIPEDIA, 2011).

Outra modalidade utiliza a escrita musical, ou seja, uma partitura produzida pelo usuário no sistema (*Figura 4*), como expressão de busca (SUYOTO, 2008; MUSIPEDIA, 2011).

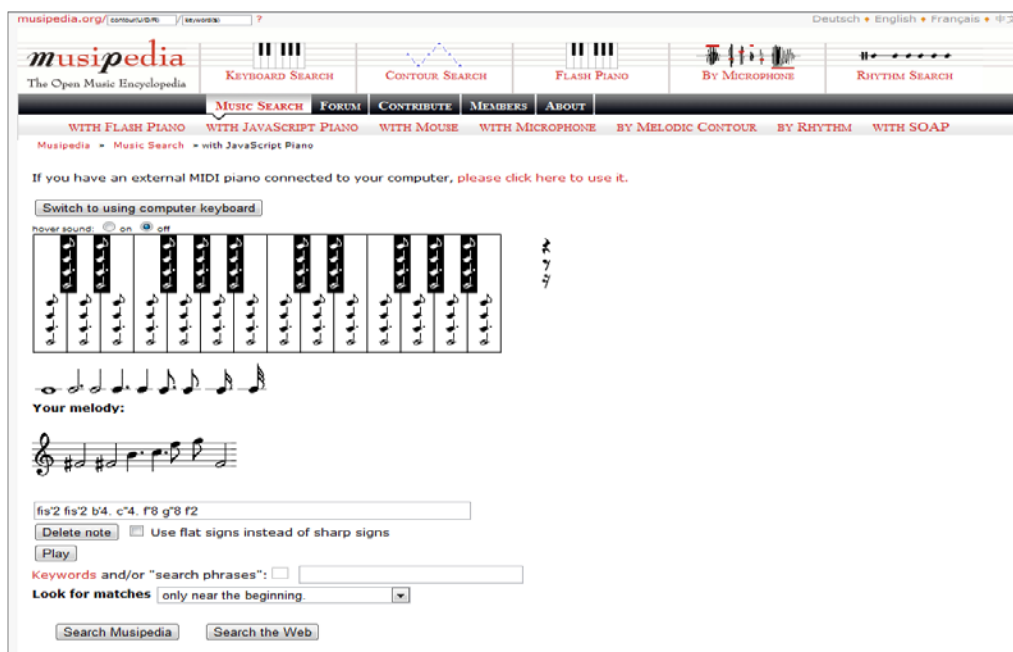


Figura 4: Consulta-par-Sintagma-Musical do Musipedia.org

Esse tipo de consulta, entretanto, requer conhecimentos musicais prévios, e portanto tendem a ser mais adequados em ambientes acadêmicos ou em populações com treinamento musical.

Consulta-**Por-Texto**

Essa categoria circunscreve os sistemas em que a expressão de busca é expressa em linguagem verbal, seja natural ou estruturada. Portanto inclui a busca tradicional por metadados como título de músicas, álbuns, artistas, informações geográficas, cronológicas, etc. Inclui a busca por as tags, sejam elas produtos de técnicas de filtragem colaborativa, anotação por especialistas, autotag ou de processos de mineração de dados. É forma mais usual e mais recorrente de busca em sistemas digitais e em geral não prescinde de conhecimentos musicais, nem da disponibilidade de fragmentos de áudio ou dispositivos de captação de áudio.

Consulta-**Por-Similaridade**

A consulta por similaridade encontra músicas que potencialmente sejam relevantes para uma pessoa a partir de um documento musical ou artista presente na base de dados (*Figura 5*). Dito de outra forma, recupera músicas e artistas que sejam similares ao dado como referência pelo usuário. Portanto mais do que um processo de recuperação propriamente, é um processo de recomendação e descoberta de músicas. As medidas de similaridades podem ser de três tipos: *similaridade acústica*, através da análise e comparação de atributos do sinal de áudio; *similaridade semântico-social*, através da análise e comparação de tags, dados colaborativos, opiniões, perfis de preferência e padrões de consumo; *similaridade semântico-socio-acústica*, que combinam medidas de similaridade semânticas, sociais e de áudio, portanto formas híbrida das duas anteriores (NANOPOULOS, 2009).

Embora as medidas possam ser divididas nesses três tipos, devido ao caráter multifacetado da música, os estudos recentes optam por desenvolver sistemas híbridos, que possam recomendar músicas com base em similaridades, tanto sociais, quanto semânticas e acústicas (BOGDANOV, 2011; KNEES, 2007; PAMPALK, 2006)

The image shows a screenshot of the Last.fm website interface. It displays four artist profiles in a vertical list. Each profile consists of a square cover image on the left, followed by the artist's name, a play button icon, the number of executions (e.g., 318.488 for Sabotage), and a button labeled 'Adicionar à sua biblioteca'. Below this is a list of genre tags (e.g., 'rap, rap nacional, hip hop, brazilian, heavy metal' for Sabotage) and a short biographical paragraph. To the right of each profile is a box titled 'Artistas parecidos em sua biblioteca' (Artists similar to those in your library), which contains two smaller artist thumbnails and their names (e.g., 'Racionais MC's' and 'Criolo' for Sabotage). A large teal double-headed arrow is positioned between the Sabotage and EMICIDA profiles, indicating a similarity or relationship between them.

Figura 5: Consulta-por-Similaridade do Last.fm

Consulta-Por-Classificação

Assim como a Consulta-por-Similaridade apresenta-se como uma forma de descoberta de músicas, uma vez que não recupera itens a partir de uma expressão específica, mas organiza uma coleção de músicas ou artistas que compartilhem similaridades conceituais estabelecidas através de ontologias. São as ontologias de gênero, por exemplo, muito populares ainda, mas que atualmente também utilizam outras facetas de classificação, como humor, ou emoção, como é o exemplo do AllMusic.com (Figura 6). Essa ontologias são criadas por especialistas, que no caso do AllMusic, também produzem revisões sobre artista e gêneros. Assim apresenta-

se como uma abordagem tradicional de organização, classificando os conteúdos através de vocabulários controlados, e os disponibilizando em forma de catálogos virtuais.

The screenshot shows the AllMusic.com website interface. At the top, there is a search bar with a 'GO' button and a dropdown menu for 'Artist/Group'. The site is powered by Rovi and AllMusic. A navigation bar lists various music genres: POP/ROCK, JAZZ, R&B, RAP, COUNTRY, BLUES, ELECTRONIC, LATIN, REGGAE, INT'L, and CLASSICAL. Below this, a 'Moods' section is highlighted, with a search bar and the text 'explore music...'. The main content area displays a grid of mood categories, each with a small icon and a link. The categories are arranged in three columns:

- [Acerbic](#)
- [Amiable/Good-Natured](#)
- [Atmospheric](#)
- [Bitter](#)
- [Boisterous](#)
- [Brittle](#)
- [Campy](#)
- [Cerebral](#)
- [Complex](#)
- [Crunchy](#)
- [Detached](#)
- [Dreamy](#)
- [Earthy](#)
- [Elegant](#)
- [Epic](#)
- [Fierce](#)
- [Freewheeling](#)
- [Gleeful](#)
- [Gritty](#)
- [Hedonistic](#)
- [Hypnotic](#)
- [Intense](#)
- [Irreverent](#)
- [Aggressive](#)
- [Angrv](#)
- [Austere](#)
- [Bittersweet](#)
- [Brash](#)
- [Brooding](#)
- [Carefree](#)
- [Cheerful](#)
- [Confident](#)
- [Cynical/Sarcastic](#)
- [Distraught](#)
- [Druggy](#)
- [Eccentric](#)
- [Energetic](#)
- [Ethereal](#)
- [Fierv](#)
- [Fun](#)
- [Gloomy](#)
- [Happy](#)
- [Hostile](#)
- [Indigent](#)
- [Intimate](#)
- [Jovous](#)
- [Ambitious](#)
- [Angrt-Ridden](#)
- [Autumnal](#)
- [Bleak](#)
- [Bravado](#)
- [Calm/Peaceful](#)
- [Cathartic](#)
- [Cold](#)
- [Confrontational](#)
- [Delicate](#)
- [Dramatic](#)
- [Earnest](#)
- [Eerie](#)
- [Enigmatic](#)
- [Exuberant](#)
- [Fractured](#)
- [Gentle](#)
- [Greasy](#)
- [Harsh](#)
- [Humorous](#)
- [Innocent](#)
- [Ironic](#)
- [Laid-Back/Mellow](#)

On the right side of the page, there is a promotional graphic for a 'SEND A CALLOUT CARD' campaign. The graphic features a speech bubble with the text: 'WHEN YOU PRESSURE ME FOR NUDE PICS, I THROW UP IN MY MOUTH A LITTLE'. Below this, it asks 'Feeling pressured or disrespected?' and encourages users to 'SEND A CALLOUT CARD' and 'And draw your digital line.'

Figura 6: Classificação de Gêneros e Humor do AllMusic.com

Entretanto a classificação em grandes classes de gêneros, embora necessárias, apresentam problemas intrínsecos de ambiguidade e falta de consenso (AUCONTURIER, 2003). PACHET e CASALY (2000) ao compararem 3 taxonomias de gênero da Internet, - Allmusic, Amazon e MP3.com, com 531, 719 e 430 gêneros, respectivamente – encontraram 70 palavras em comum entre as três, além de não apresentarem similaridades estruturais.

O rigidez do controle de vocabulário também pode ser um revés, ao não acompanhar o caráter dinâmico da produção musical, que pode significar um distanciamento da percepção dos usuários em relação às categoria de gênero e estilos, como aponta SANTINI (2011):

“(...) when comparing the similarity criteria used by the industry to group artists and organize the music market against user classifications and uses (...) the compatibility is minimal”. (p.241)

DISCUSSÃO

A música sofreu uma revolução nesta primeira década do século XXI, em todas as suas instâncias. Não só por conta do rápido desenvolvimento das tecnologias de comunicação e informática, que trouxeram novas formas de produção, de comunicação e organização da música, mas também por alterações nas práticas sociais, de consumo, nos processos de significação, mediação, validação cultural associadas à música (SANTINI, 2011).

Até o final da década de 1990 o consumo e as práticas sociais relacionadas à música eram conduzidos pelas estratégias de mercado de uma indústria extremamente concentrada. Essa indústria em seu modelo tradicional, bem como toda a indústria cultural e de entretenimento, orienta-se por uma lógica de consumo e propagação massiva, o que equivale a dizer que o objetivo primário é maximizar o tanto quanto possível a penetração e difusão global de um determinado produto cultural. Dessas duas matrizes de pensamento, concentração econômica e consumo cultural massivo, surge o modelo de consumo da música vigente até então, baseado no desenvolvimento de um portfólio reduzido de artistas responsáveis pela maior parcela possível do marketshare. Justifica-se assim o investimento massivo em propaganda e mídia relacionadas a esse portfólio de artistas, garantindo a penetração e conseqüente aumento das vendas.

Esse cenário incorreu em um processo de retroalimentação em que quanto maior o sucesso das vendas de um produto cultural, maiores eram os investimentos em publicidade, e quanto maior os investimentos em publicidade, maior o sucesso de vendas. Esse chamado “efeito super-star”, aliado a alta concentração econômica e administrativa da distribuição da música impôs limites bastante rígidos à diversidade

de produtos musicais, trazendo certa homogeneidade e empobrecimento das práticas de consumo musical, levando a uma cultura de “hits”:

Poucas músicas que se destacam nas paradas são presenças constantes nas emissoras, enquanto os executivos da indústria de entretenimento suam a camisa em busca da próxima grande mina. (...). A mídia de massa e a indústria do entretenimento cresceram nos últimos cinquenta anos nas costas dos campeões de bilheteria, dos discos de ouro e dos níveis de audiência de dois dígitos (ANDERSON, 2006, p.1)

Entre o fim da década de 90 e começo da de 2000 surgiram quatro fatores determinantes para a mudança desse cenário: barateamento da produção musical, Internet, MP3 e redes p2p. Isso porque a combinação desses fatores fragilizava pontos importantes da estratégia de concentração do modelo anterior já que: a) muitos passaram a gravar músicas com qualidade técnica no mínimo aceitável para o mercado, com recursos financeiros muito menores; b) a duplicação do registro sonoro caiu para quase-zero com o advento do MP3; c) a distribuição, que antes demandava grandes estruturas logísticas, reduziram-se a um computador com conexão à internet e um software de rede p2p.

Mas acima de tudo, as pessoas se apropriaram dessas tecnologias e fizeram da troca de arquivos de música umas das atividades mais populares da internet e parte de uma cultura digital global (VACARO, 2004). Nessa cultura digital há descentralização da distribuição, da produção e da digitalização das músicas, sendo que todas essas ações passam a não ser diretamente dependentes de uma estratégia de mercado. Assim emerge novamente com naturalidade a diversidade de gostos e preferências musicais das pessoas, que compartilham músicas, opiniões e informações sobre elas, como forma de expressar seus sentimentos e identidades. Ou seja, a música volta, aos poucos, a emergir como uma prática social rica e diversa. Assim inicia-se um movimento de ramificação do consumo em nichos de gostos culturais e praticas diversas, ou seja, a transformação do mercado de massa em um mercado de milhões de nichos, postulada pela teoria da Cauda-Longa de ANDERSON (2006), exemplificado, no caso específico da música:

Nela [na Cauda Longa] é possível encontrar qualquer coisa, como velhos sucessos, ainda lembrados com carinho por fãs dos bons tempos ou redescobertos por novos aficionados. Há gravações ao vivo, faixas que não fizeram tanto sucesso, remixagens e até mesmo capas. Os nichos chegam às centenas, gêneros, dentro de gêneros, dentro de gêneros (imagine toda uma grande loja de CDs inteiramente dedicada às bandas de cabeludos da década de 1980 (hair bands) ou a músicas ambientais rítmicas (ambient dub). Também há bandas estrangeiras, cujos preços as tornavam inacessíveis nas prateleiras de importados, e bandas obscuras ou até selos ainda mais desconhecidos - muitas das quais não têm força suficiente para entrar nas grandes lojas. (ANDERSON, 2006, p.21)

No centro do paradigma dessa teoria, no lugar da onipresença de um produto cultural, a descoberta e acessibilidade da maior quantidade possível de produtos, de forma à atender o maior número de nichos. (p.21)

Essa nova lógica pode ser observada em sites como do Last.fm, MusicBeta do Google⁴⁷, iCloud da Apple⁴⁸ e Pandora, que não baseiam-se propriamente na venda de músicas, ou álbuns, mas no oferecimento de um serviço de organização, acesso e descoberta de músicas. Isso fica ainda mais claro na forma de monetarização, uma taxa mensal, tal como a do provedor de Internet, ou da companhia de luz. Esses serviços, assim estruturados atendem à uma demanda contemporânea:

Os consumidores estão mergulhando de cabeça nos catálogos, para vasculhar a longa lista de títulos disponíveis, muito além do que é oferecido na Blockbuster Vídeo e na Tower Records. E quanto mais descobrem, mais gostam da novidade. A medida que se afastam dos caminhos conhecidos, concluem aos poucos que suas preferências não são tão convencionais quanto supunham (ou foram induzidos a acreditar pelo marketing, pela cultura de hits ou simplesmente pela falta de alternativas. (ANDERSON, 2006, p.15)

Nesse contexto a RIM, ao se propor a, e efetivamente, desenvolver soluções de recuperação e descoberta de músicas, apresenta-se como área fundamental para os modelos de negócio de música baseado em serviços. Mas o seu grau de importância é equivalente aos desafios que enfrenta no cumprimento de sua proposição.

⁴⁷ <http://music.google.com>

⁴⁸ <http://www.apple.com/icloud>

Os estudos de representação e recuperação baseada-em-conteúdo tentam, a partir do sinal de áudio inferir automaticamente descritores significativos para humanos, capazes de possibilitar recuperação. Apesar da evolução que apresenta desde a constituição dos diálogos multidisciplinares em RIM, ainda está em fase embrionária, se considerado o seu propósito. Isso porque os desenvolvimentos estão por enquanto em nível de sinal, sua análise, computação etc., ou seja, debruçam-se sobre a faceta física da informação. A informação, enquanto processo, demanda de um sistema interpretante, sociocognitivo, para criar significados. Esse papel essencialmente humano, nos processos automatizados é desempenhado, de certo modo, por algoritmos de aprendizagem e estes, por sua vez, são autômatos, que aplicam uma lógica probabilística em um conjunto de exemplos de associação entre sinais de áudio e palavras, ou seja, a base de treinamento. Nessa base, está cristalizada a memória de associações música-palavras de um grupo de pessoas, e é a partir dela que os algoritmos reconhecem padrões para criar regras de associação que norteiam as decisões do classificador automático. Portanto, para inferir significados potenciais a partir de atributos acústicos, tão importante quanto a compreensão, análise e síntese do sinal de áudio, é o desenvolvimento das bases de treinamento. Apesar disso, é recorrente na literatura o problema em compor bases de larga escala, dado as restrições de licenças e direitos autorais (SANTINI, 2007; BERTIN-MAHIEUX, 2011; TINGLE, 2010). Também a maioria dos datasets disponíveis para treinamento e teste de algoritmos são resultado de filtros colaborativos ou de mineração de dados, e portanto, de baixo grau de precisão e unicidade de termos, como resume TINGLE (2010):

One problem facing the Music Information Retrieval (Music-IR) research community is the lack of a large, cleanly-labeled data set. This problem persists partly because of the inability to freely distribute a large corpus of high-quality music without violating copyright law. Another obstacle is the development of a standard vocabulary of music tags. However, the (...) cost of training experts and collecting annotations is perhaps the biggest problem for the research community (p. 55)

Por outro lado temos a prática de Social Tagging como a forma mais popular de representação de músicas nos sistemas da Web. Embora essa técnica seja difundida em serviços de música Web, como no caso do Last.fm, ela apresenta o problema intrínseco da baixa precisão e unicidade das tags utilizadas

(MACGREGOR, 2006; LAMERE, 2008). Para além desse revés há o “efeito de popularidade” (CELMA, 2008), que faz com que, embora o montante de tags nesses sistemas seja enorme, haja uma alta concentração delas na descrição de um número reduzido de artistas (*Figura 7*). Em resumo, poucos artistas recebem a maioria das tags, a maioria dos artistas recebe poucas, ou nenhuma tag. Esse é um revés importante ao considerarmos que os sistemas de recomendação operam em um modelo baseado na teoria da Cauda-Longa.

Apesar desse efeito de distribuição ser esperado como um componente socialmente inerente (CELMA, 2008), ele é amplificado pelo fato de a prática de tagging estar em parte associada a um subgrupo social, que não necessariamente reflete o gosto do grupo como um todo, como aponta LAMERE, 2008:

A typical tagger is likely to be young, affluent, and Internetsavvy(...). The music taste of these taggers may not be representative of music tastes of the general population. This can lead to tagging bias where some types of music receive more than their fair share of tags. Artists that are favorites of taggers will receive a disproportionately large number of tags, while artists that are not listened to by taggers will be tagged infrequently. This tagger bias can be demonstrated by comparing traditional music sales to tagging behavior. (p.112)

Já a anotação por especialistas apresenta uma taggemaneto de alta qualidade (TURNBULL, 2008b) e uma frequência de distribuição de tags mais homogênea na Cauda-Longa (CELMA, 2008). Entretanto demanda de uma quantidade grande de tempo, recursos humanos e financeiros (TURNBULL, 2008b), e não tem sido entendido na literatura como uma opção viável para bases de larga escala

Nota-se portanto entre essas três abordagens uma complementariedade, coerente com o aspecto multifacetado da representação musical proposta pela MIR: a possibilidade da automação em larga escala da representação de músicas baseada em atributos do áudio; complexidade semântica e social da Mineração de Dados e das práticas de Social Tagging; precisão nominativa da anotação por especialista.

Por esse aspecto complementar que muitos dos trabalhos recentes propõem abordagens híbridas, combinando atributos de áudio, Social Tagging, e Mineração de Dados.

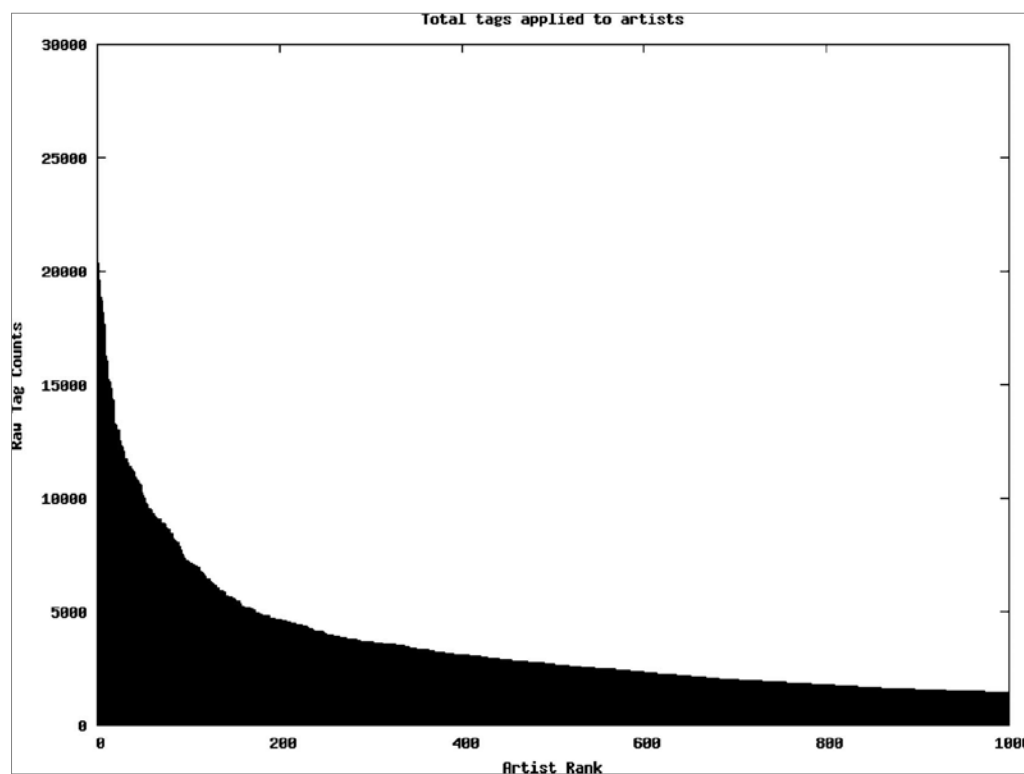


Figura 7: Frequência de Tags para o 1000 primeiros artistas do Last.fm (LAMERE, 2008)

KIM (2009) propõe uma abordagem de “propagação de tags”, que consiste em atribuir à artistas não indexados, tags de artistas que apresentem maior similaridade a ele. Entretanto essa similaridade é computada através de métricas que consideram dados de preferência de usuários coletados em sistema de filtro colaborativo, similaridade de social tags e padrões de dados encontrados através da mineração de dados em documento de texto da Web. Esse modelo apresentou melhores resultados quando comparado a modelo de propagação de dados que baseados exclusivamente em descritores acústico baseado-em-conteúdo.

WANG, J. (2011) propõe uma abordagem similar ao combinar LLFs à social tags a fim de otimizar a aprendizagem dos algoritmos classificadores e diminuir o hiato semântico entre descritores acústicos e julgamento humano de similaridade musical.

WANG, D. (2010) avalia a eficácia no uso de social tags e de atributos acústicos para o agrupamento de músicas de acordo com o estilo musical, e em seguida propôs um modelo integrando tags e atributos acústico resultando em melhores resultados do que se usados isoladamente.

Esses trabalhos corroboram a hipótese de que a combinação de atributos acústicos, sociais e semânticos pode trazer melhores resultados quanto ao estabelecimento de similaridade entre músicas, portanto, de eficácia na recomendação e recuperação de músicas.

CONCLUSÃO

As três principais hipóteses de representação e recuperação de músicas encontradas na literatura apresentam vantagens e desvantagens em seu uso na organização e coleções de músicas de larga escala.

A abordagem baseada-em-conteúdo propõe-se a classificar e indexar músicas atribuindo nominativos inferidos a partir do sinal de áudio, possibilitando assim a organização de bases de dados, independente da intervenção humana no processo de representação. Para tanto, analisa as características de sinal de áudio e as associa a descritores através de algoritmos de aprendizado. Entretanto para inferir nominativos semanticamente significativos para humanos, depende do desenvolvimento de bases de treinamento, ou seja, uma coleção de exemplos dessas associações para que os algoritmos possam estabelecer regras de classificação e indexação, úteis para o classificador automático. Assim, tão importante quanto a qualidade de extração dos atributos de áudio e dos algoritmos, é a constituição de uma base de treinamento com alta qualidade de representação e relevância semântica, o que, de acordo com a literatura, não ocorre pela dificuldade em montar bases anotadas por especialistas, e de compartilhá-las, por restrições de

direitos autorais. O principal problema enfrentado, portanto, por essa abordagem, é a atribuição de enunciados semanticamente significativos aos atributos de áudio extraídos.

Já a abordagem de Social Tag e de Mineração de Dados, baseiam-se em descrições feitas por humanos, e para humanos. Entretanto, no caso da Social Tag, depende do estabelecimento de grandes comunidades de “taggers”. Os nominativos são atribuídos livremente, sem mediação de vocabulários ou especialistas, e disto decorre que, se por um lado essas representações acompanham a dinâmica da diversidade da produção cultural e de seus significados sociais, por outro, apresenta naturalmente alto grau de ambigüidade de termos e a diluição dos enunciados, comprometendo assim a eficácia e a eficiência da recuperação. Também apresenta um viés de popularidade, que faz com que as tags sejam pouco distribuídas entre as músicas e artistas, ou seja, grande parte dos conteúdos apresenta poucos pontos de acesso.

Em contraponto às duas anteriores, a anotação por especialista resulta em enunciados de alta qualidade e de alta relevância semântica, produzindo em teoria, sistemas de maior precisão na recuperação. Entretanto, considerando bases de dados de larga escala, é um processo oneroso e lento, e, de certa forma, anacrônico. Além disso, o sistema tende a apresentar falta de mobilidade, pelo fato de o controle de vocabulário não acompanhar a evolução contínua da criação e resignificação artística e estética da música. Por isso não tem sido considerado uma solução viável para serviços de web.

Concluimos que tanto no caso das Social Tags, quanto da abordagem baseada-em-conteúdo, existe a carência de metodologias de estabelecimento de cadeias de relacionamento de unidades semânticas significativas, para fins de recuperação. No caso da Social Tag, de mecanismos que diminuam a ambigüidade das tags e aperfeiçoem a consolidação dos enunciados. No caso da RpMBC, no desenvolvimento de bases de treinamento de algoritmos com alta qualidade de indexação e classificação. Embora a Ciência da Informação e Biblioteconomia no Brasil não tenha apresentado até agora aproximação com as atividades de representação e recuperação de informação da música na Web, vimos que muitos

dos propósitos da RIM são relacionados aos fazeres tradicionais da área, principalmente no domínio da Análise Documentária.

Portanto entendemos que pesquisas futuras relacionadas ao desenvolvimento de métodos e procedimentos que visem diminuir as deficiências de vocabulário e de distribuição das representações nos sistemas de Social Tag, ou que se concentre no desenvolvimento eficiente de bases de dados de alta qualidade metodológica, semântica e terminológica para uso em treinamento de algoritmos de aprendizado nos sistemas baseados-em-conteúdo, podem ser de grande relevância para o avanço do conhecimento e dos desenvolvimentos na área de Recuperação de Informação da Música.

BIBLIOGRAFIA

ADORNO, Theodor W.; HORKHEIMER, Max. *Dialética do esclarecimento: fragmentos filosóficos*. 2. ed. Rio de Janeiro: Zahar, 1986. 254 p.

ALEXANDER, Peter. New Technology and Market Structure: evidence from the music recording industry. **Journal Of Cultural Economics**, v. 18, p.113-123, 1994.

ANDERSON, Chris. **A cauda longa**: Do mercado de massa para o mercado de nicho. Rio de Janeiro: Campus, 2006. 256 p.

AUCOUTURIER, J; PACHET, F. Representing musical genre: a state-of-the-art. **Journal Of New Music Research**, v. 32, n. 1, p.83-93, 2003.

ARAÚJO, Bráulio Santos Rabelo de. O conceito de aura, de Walter Benjamin, e a indústria cultural. *Pós*, v. 17, n. 28, p.120-143, dez. 2010.

BANDERA, Cristina et al. Humming method for content-based music information retrieval. In: International Society For Music Information Retrieval Conference, 12., 2011, Miami. Disponível em <<http://ismir2011.ismir.net/program2011.html>>

BATLLE, Eloi. Recognition and analysis of audio for copyright protection: the RAA project. **Journal Of The American Society For Information Science And Technology**, v. 55, n. 12, p.1084-1091, 2004.

BENJAMIN, Walter. A obra de arte na era de sua reprodutibilidade técnica. In: BENJAMIN, Walter. *Obras escolhidas magia e técnica, arte e política: ensaios sobre literatura e história da cultura*. São Paulo: Brasiliense, 1985.

BERTIN-MAHIEUX, Thierry et al. The Million Song Data Set. In: International Society For Music Information Retrieval Conference, 12., 2011, Miami. Disponível em <<http://ismir2011.ismir.net/program2011.html>>

BIRMINGHAM, W; DANNENBERG, R; PARDO, B. Query by humming with the VocalSearch system. **Communications Of The ACM**, v. 49, n. 8, p.49-52, 2006.

BOGDANOV, Dimitri et al. Unifying Low-Level and High-Level Music Similarity Measures. **Ieee Transactions On Multimedia**, v. 13, n. 4, p.687-701, 2011.

BOURREAU, Marc; MOREAU, François; GENSOLLEN, Michel. The Digitization of the Recorded Music Industry: Impact on Business Models and Scenarios of Evolution. **Telecom Paristech Working Paper**, n. -08-01, 2008. Disponível em <http://economie.cnam.fr/servlet/com.univ.collaboratif.utils.LectureFichiergw?ID_FICHI ER=1295877017251>.

BURNETT, Robert. *The Global Jukebox: the international music industry*. Nova York: Routledge, 1996.

CANO, Pedro; BATLLE, Eloi; KALKER, Tom. A Review of Audio Fingerprinting. **Journal Of Vlsi Processing**, v. 41, n. 3, p.271-284, 2005.

CASEY, Michael et al. Content-Based Music Information Retrieval: Current Directions and Future Challenges. **Proceedings Of The Ieee**, v. 96, n. 4, p.668-696, 2008.

CELMA, Oscar; CANO, Pedro. From hits to niches? or how popular artists can bias music recommendation and discovery. **Proc. 2nd Workshop On Large-scale Recommender Systems And The Netflix Prize Competition**, 2008.

CHANDRASEKHA, Vijay et al. Survey and evaluation of audio fingerprinting schemes for mobile query-by-example application. In: International Society For Music Information Retrieval Conference, 12., 2011, Miami. Disponível em <<http://ismir2011.ismir.net/program2011.html>>

CINTRA, Anna Maria Marques et al. **Para entender as linguagens documentárias**. São Paulo: Polis, 2005. 92 p.

COENEN, Frans. Data mining: past, present and future. **The Knowledge Engineering Review**, v. 26, n. 1, p.25-29, 2011.

COLLARD, L. The Internet Music Revolution. **7th Annual Multimedia Systems, Electronics And Computer Science**, 2006. Disponível em <<http://mms.ecs.soton.ac.uk/2007/papers/34.pdf>>.

COLVIN, Joao. Naming That Tune: Mobile Music Information Retrieval Systems. **Music Reference Services Quarterly**, v. 12, n. 1-2, p.29-32, 2009.

DENCKER, Ada de Freitas Maneti; VIÁ, Sarah Chucid da. Pesquisa empírica em ciências humanas: com ênfase em comunicação. São Paulo: Futura, 2001.

DOWNIE, Stephen. Music Information Retrieval. **Annual Review Of Information Science And Technology**, v. 37, n. 1, p.295-340, 2003.

EISENBERG, Gunnar; BATKE, Jan-mark; SIKORA, Thomas. BeatBank: An MPEG-7 compliant Query by Tapping System. In: AUDIO ENGINEERING SOCIETY CONVENTION, 116, 2004, Berlin, Alemanha.

FREEDMAN, D. Managing pirate culture: Corporate responses o peer-to-peer networking. **International Journal On Media Management**, v. 5, n. 3, p.173-179, 2003.

GOMES, Ro De C. T. et al. Audio watermarking and fingerprinting: for which applications? **Journal Of New Music Research**, v. 32, n. 1, 2003.

GUO, Le et al. Content-based retrieval of polyphonic music objects using pitch contour. **Ieee International Conference On Acoustics, Speech And Signal Processing**, v. 12, n. 1, p.2205-2208, 2008.

HACKER, Scot. MP3: the definitive Guide. California, USA: O'Reilly & Associates, Inc, 2000.

HINZ, Oliver; ECKERT, Jochen. The Impact of Search and Recommendation Systems on Sales in Electronic Commerce. **Business & Information Systems Engineering**, v. 2, n. 2, p.67-77, 2010.

HOLZAPFEL, A; STYLIANOU, Y. Musical genre classification using nonnegative matrix factorization-based features. **Ieee Transactions On Audio Speech And Language Processing**, v. 16, n. 2, p.424-434, 2010.

HUGHES, J; LANG, K. R. If I had a song: The culture of digital community networks and its impact on the music industry. **International Journal On Media Management**, v. 5, n. 3, p.180-189, 2013.

INTERNATIONAL FEDERATION OF THE PHONOGRAPHIC INDUSTRY. **IFPI releases definitive statistics on global market for recorded music**: 12th annual Recording Industry in Numbers. Disponível em: <http://www.ifpi.org/content/section_news/20050802.html>. 2 ago. 2005. Acesso em : jul. 2011

_____. **IFPI Digital Music Report 2010**. Disponível em: <<http://www.ifpi.org/content/library/DMR2010.pdf>>. 2010. Acesso em: set. 2011

JANOTTI JUNIOR, Jeder Silveira. À procura da batida perfeita: a importância do gênero musical para a análise da música popular massiva. **Eco-pos**, v. 3, n. 2, p.31-46, 2003.

JINHEE, Y; SANGHYUN, P. An efficient frequent melody indexing method to improve the performance of query-by-humming systems. **Journal Of Information Science**, v. 34, n. 6, p.777-798, 2008.

IRWIN, Ken. Musipedia: The Open Music Encyclopedia. **Reference Reviews**, v. 22, n. 4, p.45-46, 2008.

KNEES, Peter; POHLE, Tim; WIDMER, Gehrard. A music search engine built upon audio-based and Web-based similarity measures. In: Annual International Acm Sigir Conference, 30, 2007, Amsterdam, Holanda. Disponível em: <<http://www.sigir.org/sigir2007>>

_____ et al. A document-centered approach to a natural language music search engine. **Advances In Information Retrieval**, n. 4956, p.627-631, 2008. Disponível em: <http://www.cp.jku.at/research/papers/Knees_et_al_ECIR08.pdf>.

KIM, Joon Hee; TOMASIK, Brian; TURNBULL, Douglas. Using artist similarity to propagate semantic information. In: International Society For Music Information Retrieval Conference, 10., 2009, Utrecht, Holanda. Disponível em: <<http://www.ismir.net/proceedings>>

LAMERE, P et al. Social Tagging and Music Information Retrieval. **Journal Of New Music Research**, v. 37, n. 2, p.101-114, 2008.

LEE, S K. CHO, Y H. Collaborative filtering with ordinal scale-based implicit ratings for mobile music recommendations. **Information Sciences**, v. 180, n. 11, p.2142-2155, 2010.

LESSIG, Lawrence. Remix: making art and commerce thrive in the hybrid economy. Nova York : Penguin Press, 2008.

LOGAN, B. Mel Frequency Cepstral Coefficients for Music Modeling. In: International Society For Music Information Retrieval Conference, 1, 2000, Plymouth, Estados Unidos. Disponível em <<http://www.ismir.net/proceedings>>

LIU, L; LIU, D. Automatic mood detection and tracking of music audio signals. **IEEE Transactions On Audio Speech And Language Processing**, v. 14, n. 1, p.5-18, 2006.

MACGREGOR, G; LIU, D; MCCULLOCH, E. Collaborative tagging as a knowledge organisation and resource discovery tool. **Library Review**, v. 55, n. 5, p.291-300, 2006.

MANDEL, Michael; POLINER, Graham E.; ELLI, Daniel P.W. Support vector machine active learning for music retrieval. **Multimedia Systems**, v. 12, n. 1, p.1-1, 2006. Disponível em: <<http://www.ee.columbia.edu/~dpwe/pubs/MandPE06-svm.pdf>>.

MED, Bohumil. **Teoria da Música**. 4. ed. Brasília: Musimed, 1996. 420 p.

MUSIPEDIA. **Musipedia**: the open music encyclopedia. Disponível em: <<http://www.musipedia.org>>. Acesso em: out. 2011.

NAKANO, Davi. A produção independente e a desverticalização da cadeia produtiva da música. **Gest. Prod.**, v. 17, n. 3, p.627-638, 2010.

NANOPOULOS, Alexandros et al. Music search engines: Specifications and challenges. **Information Processing And Management**, v. 45, p.392-396, 2009.

ORIO, Nicola. Music Retrieval: A tutorial and review. **Foundations And Trends In Information Retrieval**, v. 1, n. 1, p.1-90, 2006.

PACHET, François; CAZALY, Daniel. A taxonomy of Musical Genres. In: Content-Based Multimedia Information Access Conference, 6., 2000, Paris, França.

PAMPALK, Elias; GOTO, Massataka. MusicRainbow: A New User Interface to Discover Artists Using Audio-based Similarity and Web-based Labeling. In: international society for music information retrieval conference, 7, 2006, Victoria, Canada. Disponível em: <<http://www.ismir.net/proceedings>>

RYYNANEN, M. KLAPURI, A. Query by humming of MIDI and audio using locality sensitive hashing. **Ieee International Conference On Acoustics, Speech And Signal Processing**, v. 1-12, p.2249-2252, 1 2008.

SANTINI, Rose Marie; SOUZA, Rosali Fernandes. Recuperação Da Informação De Música E A Ciência Da Informação: Tendências e desafios de pesquisa. In: Encontro Nacional De Pesquisa Em Ciência Da Informação, 8, 2007, Salvador. Disponível em: <<http://www.ancib.org.br/pages/anais-do-enancib.php>>

_____; SOUZA, Rosali Fernandes. A organização do consumo de música na internet através da classificação do gosto cultural: estudo de caso do sistema de recomendação Last. fm. In: International Society For Knowledge Organization Congress - Spain, 9, 2009, Valencia, Espanha. 2009

_____; SOUZA, Rosali Fernandes. Classificação colaborativa de conteúdos não-textuais na Internet: as novas formas de mediação e organização da informação da música através da folksonomia. In: Encontro Nacional De Pesquisa Em Ciência Da Informação, 11, 2010, Rio de Janeiro. Disponível em: <<http://www.ancib.org.br/pages/anais-do-enancib.php>>

_____. Collaborative classification of popular music on the internet and its social implications. **Oclc Systems & Services: International Digital Library Perspectives**, v. 27, n. 3, p.210-247, 2011.

SCHEDL, Markus; KNEES, Peter; WIDMER, G. Discovering and visualizing prototypical artists by web-based co-occurrence analysis. **Proceedings Of The Sixth International**

Conference On Music Information Retrieval (ISMIR 05), 2005. Disponível em: <<http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.76.278>>.

_____. A music Information system automatically generated via Web content mining techniques. **Information Processing And Management**, v. 47, p.426-439, 2011.

SINANAN, Samantha. The Future is here: query by humming as an example of content-based music information retrieval. **Library Student Journal**, v. 5, 2010. Disponível em: <<http://librarystudentjournal.org/index.php/lsj/article/view/148/272>>

SONG, Jungmin; BAE, So-young; YOON, Kyoungro. Query by humming: Matching humming query to polyphonic audio. **IEEE Transactions**, v. 1, p.329-332, 2002. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1035785>.

SUYOTO, I S H; UITDENBOGERD, A L. Searching musical audio using symbolic queries. **IEEE Transactions On Audio Speech And Language Processing**, v. 16, n. 2, p.372-381, 2008.

TINGLE, Derek. Exploring automatic music annotation with "acoustically-objective". In: **Acm Sigm International Conference On Multimedia Information Retrieval**, 11., 2010, Philadelphia, Estados Unidos.

TURNBULL, Douglas; BARRINGTON, Luke. Semantic annotation and retrieval of music and sound effects. **IEEE Transactions On Audio Speech And Language Processing**, v. 16, n. 2, p.467-476, 2008a

_____; BARRINGTON, Luke. Five approaches for collecting tags for music. In: **International Society For Music Information Retrieval Conference**, 9, 2008, Philadelphia, Estados Unidos. 2008b. Disponível em: <<http://www.ismir.net/proceedings>>.

TYPKE, Rainer; WIERING, Frans; VELTKAMP, Remco C.. A survey of music information retrieval systems. In: **International Conference On Music Information Retrieval**, 6, 2005, Londres, Inglaterra.

_____. **Music Retrieval based on Melodic Similarity**. 2007. f. Tese (Doutorado) - Curso de Ciência da Computação, Universidade de Utrecht, Utrecht, Holanda, 2007. Disponível em: <<http://igitur-archive.library.uu.nl/dissertations/2007-0219-200118/index.htm>>

VACCARO, Valerie L.; COHN, Deborah Y. The Evolution of Business Models and Marketing Strategies in the Music Industry. **International Journal On Media Management**, v. 6, n. 1, p.46-58, 2004.

WANG, Dingding; LI, Tao. Are tags better than audio features?: the effect of jointuse of tags and audio content features for artistic style clustering. In: **International Society For Music Information Retrieval Conference**, 11., 2010, Utrecht, Holanda. Disponível em: <<http://www.ismir.net/proceedings>>

WANG, Ju-Chiang; LEE, Hung-Shin; WANG, Hsin-Min; et al. Learning the similarity of audio music in bag-of-frames representation from tagged music data. In: **International Society For Music Information Retrieval Conference**, 12., 2011, Miami. Disponível em: <<http://ismir2011.ismir.net/program2011.html>>

