

Biblioteca Digital Brasileira em Música (BDB-MuS): perspectivas para um dimensionamento multidisciplinar em ciência e tecnologia da informação para a pesquisa musicológica no Brasil

Beatriz Magalhães Castro
Universidade de Brasília
e-mail: beatriz@unb.br

Fernando Cruz
Universidade de Brasília
e-mail: fwcruz@ucb.br

Sumário:

Fundamentado nos conceitos e propostas discutidos no *World Summit on the Information Society* (WSIS), e aqueles emanantes da Cultura da Virtualidade Real sob o paradigma informacional, foi proposto em 2005 ao Instituto Brasileiro da Informação em Ciência e Tecnologia (IBICT) um projeto transdisciplinar para o desenvolvimento da Biblioteca Digital Brasileira em Música (BDB-Mus). Associado ao projeto Biblioteca Digital Brasileira daquele instituto, almeja integrar num único portal repositórios de informação digital que permita consulta simultânea e unificada aos conteúdos destes acervos, com a agregação e desenvolvimento de ferramentas especificamente modeladas para a recuperação de informação de objetos musicais.

Palavras-chave: Biblioteca Digital Brasileira de Música; IBICT; Open Archives Initiative; MIR/MDL.

Conceitos introdutórios sobre bibliotecas digitais

Neste trabalho abordamos primeiramente uma conceituação sobre as características e os recursos tecnológicos disponíveis para a construção de bibliotecas digitais em música, com o objetivo de subsidiar a construção da Biblioteca Digital em Música, e de outras iniciativas que venham a surgir no cenário brasileiro. Este trabalho busca primordialmente diferenciar o conceito de Biblioteca Digital daquele de uma digitalização de documentos (seja uma partitura musical ou outro documento) colocados em disponibilidade na *web*, mostrando que os conceitos, recursos, e dificuldades dele inerentes abrangem uma iniciativa muito mais ampla, incluindo o acesso por conteúdo interno e não somente por dados e descrições de campos bibliográficos.

No panorama da Ciência da Informação hoje estão em discussão dois modelos de arquiteturas informacionais: os ditos sistemas fechados (OPACS), e os sistemas abertos (*Open Archives*). Este último foi objeto da Declaração de Berlin (22.10.2003), que apóia o Movimento do Acesso Livre à Informação Científica, permitindo a leitura, cópia, impressão, referimento e até distribuição de textos de documentos eletrônicos. Este foi assinado pelo IBICT em 17 de março último, passando o Brasil a ser o único país representante da América Latina a compor a lista de signatários do documento europeu. O modelo permite ainda um incremento do impacto do conhecimento gerado (Steve Lawrence, 2001), por sua vez motivado pelas mudanças na tipologia do acesso à informação¹.

¹ Segundo Lawrence (2001), através da análise de 119.924 artigos de conferência em informática e áreas relacionadas. O número médio de citações para artigos *off-line* é de 2.74, enquanto a média de citações de artigos *on-line* é de 7.03, um aumento de 157%.

A experiência em OAI (*Open Archives Initiative*) data da década de 1990 e surge a partir do Laboratório de Los Alamos. No ano de 2000, em Santa Fé, é idealizada e produzida esta iniciativa que engendra o surgimento de várias ferramentas de software para construção de repositórios digitais (OJS, E-prints, OCS, Dspace, Fedora, CDSware, entre outros).

Neste cenário as bibliotecas digitais ganham destaque, uma vez que são capazes de agregar valor aos serviços providos pelas bibliotecas tradicionais (não virtuais), e serem capazes de absorver protocolos abertos como o OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*), o Z39.50 (*International Standard, ISO 23950*) e os padrões mais recentes para descrição de metadados, como o *Dublin Core*. Por este motivo tornaram-se um suporte essencial para armazenamento, indexação, recuperação e distribuição de documentos digitais pela Internet.

As unidades de informação das bibliotecas digitais são os objetos digitais, que na prática são os documentos não digitalizados que antes eram tratados pelas bibliotecas tradicionais. Objetos digitais, por sua natureza, criam inovações em relação aos modelos anteriores, sendo possível citar algumas:

- Os documentos digitais passaram a ser pesquisáveis por conteúdo interno. Em outras palavras, o conceito de “*pesquisa/busca*”, antes usado apenas para localizar um documento não digital, estende-se agora para “*recuperação da informação*”, que permite não apenas localizar o documento (objeto) digital, mas também consultá-lo por atributos (metadados) internos (que podem estar embutidos ou não no documento).
- A catalogação de objetos digitais passou a ser mais cooperativa entre as bibliotecas, e está havendo uma ampla discussão sobre quais protocolos (Dublin Core, MARC e variações, Z39.50 e OAI-PMH, dentre outros), são mais adequados para garantir essa melhor integração. Essa visão vai em contraposição aos modelos mais “solitários”, que marcaram o início da história da catalogação usada nas bibliotecas tradicionais.
- Houve um foco maior na definição de linguagens e padrões a serem aplicados a objetos digitais. Surgem então padrões como o XML para descrição de objetos (além de formatos associados para representação visual dos objetos), o RDF (*Resource Description Framework*) para descrição de ontologias (semânticas) para diferentes áreas do conhecimento, além de padrões como o PURL (*Persistent URL*) para ajudar na identificação única desses objetos digitais na Internet, dentre outros.

No Brasil, seguindo uma tendência internacional, os modelos de integração de bibliotecas digitais têm seguido a iniciativa dos *Open Archives*, já citada. Segundo Kuramoto (2005), coordenador de projetos especiais do IBICT:

Trata-se de uma iniciativa para desenvolver e promover padrões de interoperabilidade para facilitar a eficiente disseminação de conteúdos. O termo “*archive*” no nome Open Archives Initiative reflete a origem da OAI, na comunidade *e-prints* onde esse termo é geralmente aceito como um sinônimo para repositórios de papers científicos. A OAI usa o termo *archive* no seu sentido mais amplo: como um repositório para armazenar informação. Protocolo para coleta de metadados (OAI-PMH) – Trata-se de um protocolo de comunicação para permitir a coleta de metadados entre dois serviços, portanto o compartilhamento de metadados entre dois serviços de informação. Protocolo é um conjunto de regras de comunicação entre dois sistemas. Exemplos: FTP, HTTP, Z39.50 etc.

Do ponto de vista funcional, um modelo de *Open Archives* é composto de dois agentes: os provedores de dados e os provedores de serviços, conforme pode ser visto na Figura 1, a seguir.

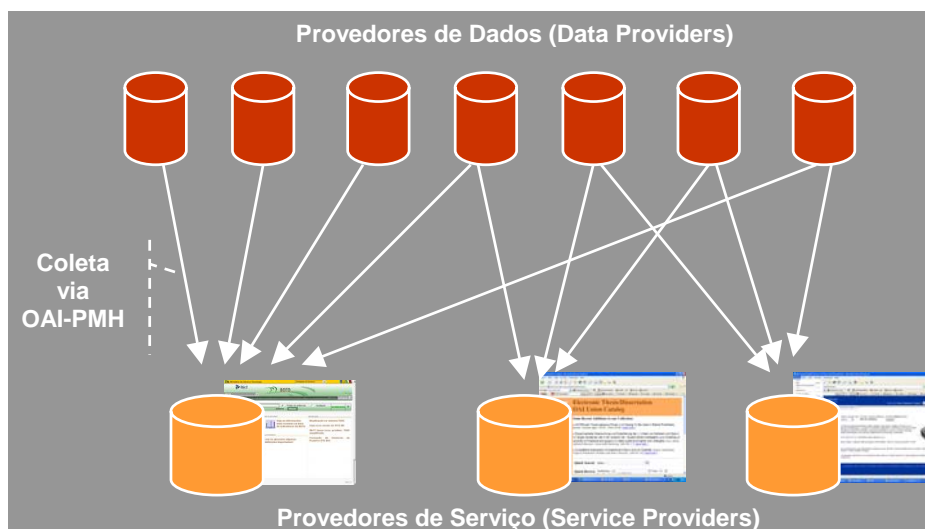


Fig. 1 – Provedores de dados e serviços. In: Kuramoto, 2005.

Nesse caso, os primeiros têm a tarefa de alimentar e construir as bases, enquanto os últimos providenciam a interface para disponibilização e recuperação dos dados de várias bibliotecas. Essa é uma forma interessante e prática de integrar bases de dados nacionais.

Em relação ao interesse da comunidade científica pelas bibliotecas digitais, atualmente já é possível perceber uma significativa melhoria nos processos de captação, armazenamento, indexação e recuperação de informações, com um foco grande em informações textuais. Vários formatos têm sido investigados com o objetivo de aumentar a qualidade das pesquisas e as formas de recuperação de textos, seja por atributos de metadados descritivos, seja pelo uso de vocabulários controlados. Com um enfoque crescente, diversos outros formatos de arquivo tais como imagens, filmes e músicas têm sido beneficiados pelas pesquisas e, no estágio atual, esses temas têm sido bastante abordados em congressos de bibliotecas digitais específicos para cada caso. Destaca-se o 3º. SIBD - Simpósio Internacional de Bibliotecas Digitais ocorrido em São Paulo em 2005².

Bibliotecas digitais de música e sistemas de recuperação musical

Como uma das formas de expressão mais populares, a música é uma importante parte da cultura, presente em praticamente todos os espaços sociais, sejam estes públicos ou privados. No campo tecnológico, é possível perceber que a reprodução e disseminação de músicas ganharam em escala e qualidade, passando das mídias analógicas (discos de vinil e fitas cassete) para mídias digitais como os Discos Compactos (ou CDs). Além disso, as *TIC - Tecnologias de Comunicação e Informação* – promoveram avanços na tecnologia subjacente e na infra-estrutura de rede, possibilitando um maior interesse do público (incluindo-se aqui a comunidade científica) pela recuperação musical, face aos recursos de softwares musicais e de bases de conteúdos musicais digitais disponibilizados pela Internet.

Dentro da Ciência da Informação, surgem então as bibliotecas digitais de música que buscam tratar a música como informação (e não apenas como expressão de arte) passível de ser preservada, armazenada, representada, catalogada e disponibilizada de maneira similar ao que já ocorre com as informações textuais.

Nessa investida, o grande problema é identificar quais os atributos importantes de uma música e como tratá-la como informação. Afinal, a música possui um conjunto de atributos (melodia, ritmo, harmonia, dentre outros) que a tornam uma informação mais complexa do que um texto, que geralmente possui apenas uma dimensão. Um outro aspecto que dificulta é que as

² Veja maiores detalhes em www.sibi.usp.br/sibd/

consultas musicais são complexas e dependem muito da musicalidade e dos desejos do usuário. Sistemas que atendam diferentes requisitos de diferentes tipos de usuários não são triviais e há um conjunto substancial de investigação ainda a ser feita nesse campo.

Um outro desafio está na forma de representação musical e os metadados utilizados. Antigamente uma música se restringia a fontes impressas ou manuscritas (partituras), e as bases de dados eram acessíveis apenas por metadados bibliográficos tais como título, autor, e outros atributos. Com a evolução tecnológica, a música foi se desvinculando do suporte informacional para assumir diversas representações aceitáveis, como um arquivo de áudio ou mesmo uma partitura eletrônica, podendo inclusive ser recuperada a partir de atributos internos. Dessa forma, a música não é mais um bloco indivisível e com um formato único com metadados associados, mas sim algo que possui diferentes representações incluindo estratégias para indexação e recuperação pelo próprio conteúdo interno da música.

Dentre as inovações possíveis com essa nova abordagem é possível citar algumas: (i) a recuperação de músicas tendo como parâmetro de entrada uma melodia cantada pelo próprio usuário numa interface de áudio, (ii) a identificação do título de uma música a partir das frases musicais presentes na obra musical, ou a (iii) a identificação de padrões musicológicos entre músicas de diferentes épocas, por exemplo.

O estado da arte das bibliotecas digitais de música procura trabalhar bases compostas por um ou mais dos seguintes tipos de informações musicais: (i) as informações desestruturadas, como as descritas em formato MP3 e WAV, (ii) as informações semi-estruturadas como as descritas em formato MIDI, e (iii) as informações musicais estruturadas como as descritas em partituras no formato CMN (*Common Music Notation*) e as músicas descritas em formato texto e cifradas. Alguns exemplos podem ser encontrados em www.musedata.org, www.themefinder.org, e o www.dlib.indiana.edu/variations/.

Apesar das vantagens dessa nova abordagem, há muito a ser avançado em pesquisas sobre informação musical e o interesse no tratamento organizado da música vem crescendo rapidamente. Dentre as iniciativas internacionais está uma ação multidisciplinar de pesquisadores provenientes das áreas de Ciência da Informação, da Ciência da Computação, da Musicologia e da Engenharia que fizeram emergir uma disciplina específica sobre o assunto, denominada MIR/MDL - *Music Information Retrieval/Music Digital Libraries*. Atualmente este grupo realiza encontros anuais³ para discutir a música e sua complexidade sob diversos aspectos incluindo técnicas de processamento de sinais e organização de bases de dados em bibliotecas digitais para atender a demanda do público interessado em informação musical, produzindo resultados significativos para o mercado.

No Brasil, até ao momento existem poucos relatos sobre estudos de MIR/MDL, apesar do Brasil ser uma fonte inesgotável de material musical. Embora prolifere uma quantidade imódica de sites e acervos musicais espalhados por diversas bibliotecas, nenhum estudo sistemático sobre necessidades de informação musical foi relatado nas referências pesquisadas, em especial nas bases de dados providas pelo IBICT (www.ibict.br). Algumas poucas iniciativas tem sido encabeçadas pela Biblioteca Nacional do Rio de Janeiro, e pelo CDMC – Unicamp, com acervos específicos, mas a área carece ainda de muito esforço para uma estruturação efetiva.

³ O site oficial do grupo está disponível em www.ismir.net

Biblioteca Digital de Brasileira em Música (BDB-Mus): propostas e metodologias

O projeto BDB-Mus surgiu a partir de uma iniciativa interdisciplinar entre os departamentos de Música e Ciência da Informação da Universidade de Brasília, e da Ciência da Computação da Universidade Católica de Brasília, no intuito de desenvolver um estudo sistemático para compartilhar conhecimentos e investigar a aplicação de tecnologias às informações musicais. Apesar de ser um campo de pesquisa vasto, algumas possibilidades de pesquisa foram assumidas:

a) Investigar formas de representação musical, ou seja, de objetos musicais e seus atributos, a fim ajudar a propor um padrão aceito nacionalmente. A premissa inicial é que tais objetos musicais contemplem as normativas RISM atualmente disponíveis, visando compatibilidade, mas, ao mesmo tempo, contemplando testes com formatos abertos propostos pela comunidade MIR/MDL, como os padrões MEI e MusicXML.

b) Investigar alternativas de arquiteturas de bibliotecas digitais de música que sirvam aos interesses dos departamentos de música e demais órgãos interessados a fim de propor diretrizes nacionais. Objetiva-se compor uma proposta de base de dados que possa servir como embrião acessível a todos os programas de ensino e pesquisa em música.

c) Investigar metadados bibliográficos que atendam as necessidades de pesquisa e de recuperação de informações musicais que sirvam aos propósitos nacionais.

d) Investigar algoritmos voltados para o tratamento do conteúdo musical a fim de identificar o que poderia servir às bibliotecas brasileiras.

e) Participar mais ativamente do grupo MIR/MDL e de outros órgãos que trabalhem com padronizações musicais em nível internacional.

Sendo assim, o projeto BDB-Mus atuaria em três áreas de aplicabilidade usando o aporte das tecnologias de informação e comunicação (TIC), no domínio do patrimônio material e imaterial, como: (i) na descoberta do que ainda é desconhecido (como é o caso de várias expressões musico-culturais brasileiras, como os arquivos de música indígena); (ii) no conteúdo de arquivos conhecidos, mas ainda não tratados (como é o caso do compositor Claudio Santoro); (iii) no conteúdo de arquivos já sistematizados, mas ainda sem padrão de aquisição de metadados (como os acervos presentes em diversas arquivos e bibliotecas no país).

Do ponto de vista da tecnologia da informação, estão previstas diversas tarefas para possibilitar a aquisição da informação (*IN phase*) e aquela da recuperação da informação pelo usuário (*OUT phase*), que incluem: definição dos metadados (de conteúdo e bibliográficos) relevantes; alternativas de captação de conteúdo; estratégias de preservação; formas de interação com o usuário; integração com a base RISM e outras bases OPAC.

Para a difusão do projeto estão previstas a promoção e realização de debates nacionais com o objetivo de homologar os resultados da pesquisa tecnológica, e das normalizações que possam vir a ser aplicadas de forma mais ampla. Deve ser discutida ainda a conceituação de uma proposta de *framework* para o uso das normas ajustadas, e a elaboração de roteiros de acompanhamento e certificação das ações implantadas.

Por sua natureza transdisciplinar, este projeto buscará sua legitimidade junto às comunidades brasileiras nas áreas de Tecnologia da Informação e Comunicação, Gestão do Conhecimento, Ciência da Informação e Musicologia. Por este motivo, a discussão desta temática tem como objetivo principal apresentar e encetar a discussão, avaliação, modelagem e integração do projeto junto à comunidade musicológica, com o objetivo de encontrar os principais recursos a partir da experiência dos profissionais da área musical.

Espera-se ainda contribuir para uma ampliação do quadro patrimonial brasileiro que poderá engendrar novos paradigmas conceituais sobre dos processos e produtos da cultura musical no Brasil, como previsto por Cook (2005) ao afirmar que “...*working with larger data sets will open up new areas of musicology.*” Esta discussão, no entanto, e os conceitos daí inferidos poderão ser

desenvolvidos em oportunidade subsequente sobre a contribuição a aspectos epistemológicos inerentes aos processos disponíveis hoje em ciência e tecnologia para o conhecimento em música.

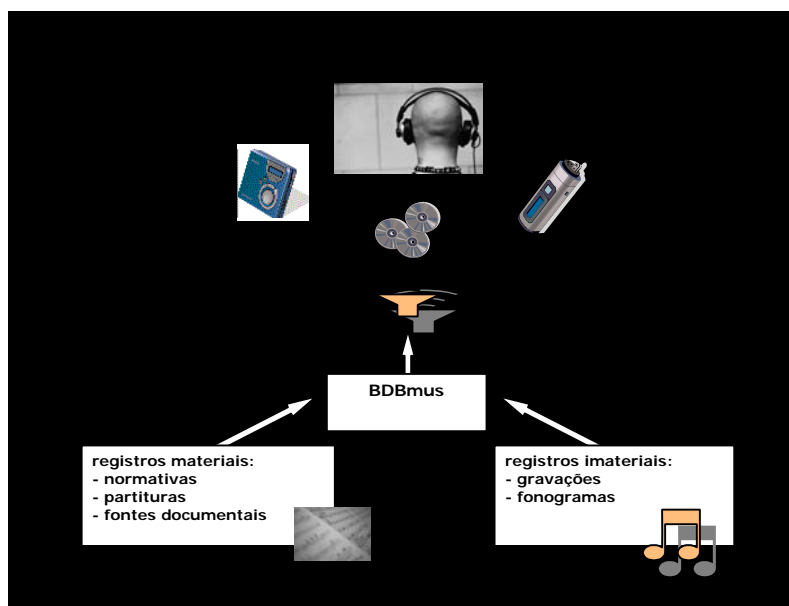


Fig. 2 – Visão do modelo BDB-Mus

Referências Bibliográficas

- Arms, W. Y. *Digital Libraries*, The MIT Press, 2000.
- Bainbridge, D. et al. - *How to Build a Digital Library*, Ed. Morgan Kaufmann, 2003.
- Bonardi, A. - IR for Contemporary Music: What the Musicologist Needs. *Proceedings of the 1st. ISMIR – International Conference on Music Information Retrieval*. Plymouth, Massachusetts - USA, 2000. Disponível em http://ismir2000.ismir.net/papers/invites/bonardi_invite.pdf Acesso em 14/12/2005.
- Byrd, D. & Crawford, T. Problems of Music Information Retrieval in the Real World (2002). In: *Information Processing and Management* 38, pp. 249-272. (disponível em <http://mypage.iu.edu/~donbyrd/Papers/RealWorldMusicIR35TR.pdf> Acesso em 03/09/2004).
- Cook, Nicholas. Towards the complete musicologist. *Proceedings of the 6th. ISMIR – International Conference on Music Information Retrieval*. London - UK, 2005. Disponível em <http://www.ismir.net/all-papers.html> Acesso em 12/04/2006.
- Downie, S. - Toward the Scientific Evaluation of Music Information Retrieval Systems. *Proceedings of the 4th. ISMIR – International Conference on Music Information Retrieval*. Baltimore Maryland - USA, 2003. Disponível em <http://ismir2003.ismir.net/papers/Downie.pdf/> Acesso em 13/06/2004.
- Downie, S., Cunningham, S. J. - Toward a Theory of Music Information Retrieval Queries: System Design Implications. *Proceedings of the 4th. ISMIR – International Conference on Music Information Retrieval*. Baltimore Maryland - USA, 2003 Disponível em <http://www.ismir.net/all-papers.html> Acesso em 08/01/2006.
- Downie et al. - Perspectives on Music Information Retrieval. *JASIST – Journal of the American Society for Information Science and Technology* - No. 12, Vol. 55, 2004.
- Grund C. M, Music Information Retrieval, Memory and Culture: Some Philosophical Remarks. *Proceedings of the 6th. ISMIR – International Conference on Music Information Retrieval*. London - UK, 2005. Disponível em <http://www.ismir.net/all-papers.html> Acesso em 12/02/2006.

- Kuramoto, Helio - Tecnologias para gerenciamento de bibliotecas, bibliotecas virtuais e digitais. *Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação*. Curitiba, 2005 (cedido pelo autor).
- Li, T. & Ogihara, M. - Detecting Emotion in Music. *Proceedings of the 4th. ISMIR – International Conference on Music Information Retrieval*. Baltimore Maryland - USA, 2003. Disponível em <http://www.ismir.net/all-papers.html> Acesso em 21/01/2006.
- Selfridge-Field, E. - What Motivates a Musical Query? *Proceedings of the 1st. ISMIR – International Conference on Music Information Retrieval*. Plymouth, Massachusetts - USA, 2000. Disponível em <http://www.ismir.net/all-papers.html> Acesso em 12/01/2006.
- Uitdenbogerd, A. - Music IR: Past, Present, and Future. *Proceedings of the 1st. ISMIR – International Conference on Music Information Retrieval*. Plymouth, Massachusetts - USA, 2000. Disponível em <http://www.ismir.net/all-papers.html> Acesso em 12/01/2006.
- Uidenbogerd, A. & Yap, Y. W. – Was Parsons Right? An Experiment in Usability of Music Representations for Melody-Based Music Retrieval. *Proceedings of the 4th. ISMIR – International Conference on Music Information Retrieval*. Baltimore Maryland - USA, 2003. Disponível em <http://www.ismir.net/all-papers.html> Acesso em 06/01/2006.
- Wiering, F, Typke, R. & Veltkamp, R. C. - A Survey of Music Information Retrieval Systems. P *Proceedings of the 6th. ISMIR – International Conference on Music Information Retrieval*. London - UK, 2005. Disponível em <http://www.ismir.net/all-papers.html> Acesso em 15/01/2006.