

Digitalizando a Discoteca L. C. Vinholes: Processos no tratamento documental da coleção de discos de 78 RPM

FELIPE FRANÇA DE ANDRADE JUNQUEIRA¹
EDUARDO VETROMILLA FUENTES²
PROF. DR. RAFAEL HENRIQUE SOARES VELLOSO³

¹UFPEL – felipe.junqueira@ufpel.edu.br

²UFPEL – vetromilla@gmail.com

³UFPEL – rafavelloso@gmail.com

1. INTRODUÇÃO

De acordo com SAYÃO (2016), instituições de patrimônio cultural por todo o mundo têm reconhecido a urgência de digitalizar suas coleções. Essa ação envolve não somente os tradicionais documentos em papel ou fotografias, mas também objetos tridimensionais, além de outros recursos visuais e sonoros, tais como instrumentos musicais históricos, figurinos folclóricos e discos de vinil, por exemplo. A digitalização de objetos dessa natureza destaca-se pela contribuição à preservação de heranças culturais intangíveis, tal como a memória oral e tradições locais. A crescente oferta e uso de uma larga escala de bancos de dados pela *internet* têm amplificado a expectativa em torno de serviços digitais oferecidos por instituições que lidam com patrimônio. Assim, bibliotecas, museus e arquivos têm produzido representações digitais dos seus acervos físicos, consequentemente expandidos com conteúdos “nato-digitais”, de modo a dispor de infraestrutura para gestão de coleções e serviços *online*. No entanto, Sayão chama a atenção para além da possibilidade de liberdade destes acervos quanto a mera provisão de serviços virtuais para usuários externos à instituição, pois os seus ativos informacionais podem ainda se constituir como ferramentas auxiliaadoras para a gestão dos acervos originais com os processos de “documentação, conservação, preservação, segurança, marketing e editoração”. O autor complementa este raciocínio com a seguinte reflexão:

Num patamar mais inovador, o acervo digital, que está paralelo ao acervo físico original, pode ir além de uma representação funcional deste, ampliando o seu potencial informacional, comunicacional e de reinterpretação e apresentação. Os processos de digitalização permitem que os objetos culturais digitais possam ser agregados com outros objetos formando novos constructos, reinterpretados em outros contextos para outros propósitos, compartilhados, recriados, enriquecidos, anotados com informações que podem ser compartilhadas, incorporados em outras coleções e em outras memórias, e analisados sob outros olhares, fomentando a pesquisa interdisciplinar. (SAYÃO, 2016, p. 272).

Com a finalidade de possibilitar o acesso eletrônico a pesquisadores e entusiastas, especialmente sob essa ótica dos meios digitais como um meio de expansão, reuso e ressignificação de acervos, é que se inseriram as atividades realizadas pelos autores no tratamento documental da coleção de discos de 78 RPM da Discoteca L. C. Vinholes, enquanto bolsistas de iniciação científica da FAPERGS, no projeto “Registros Fonográficos da Discoteca L. C. Vinholes do Centro de Artes da UFPEL: Catalogação, organização, usos e nexos socioculturais”. De acordo com COELHO; VELLOSO (2019), pelo consumo de massa e o alcance que tiveram, especialmente através de meios de difusão como

o rádio, os fonogramas representam a formação de uma cultura musical potencialmente reveladora das identidades sociais às quais estão relacionados e, em última instância, possuem elementos que ao mesmo tempo compõem e constroem a sociedade. Desse modo, o acervo da Discoteca conserva identidades sociais, registra a memória cultural e musical e ainda materializa questões de gosto, com implicações estéticas e ideológicas na cidade de Pelotas e região. Nisto reside a importância da guarda, gestão, conservação, documentação, e também a digitalização, divulgação e promoção do acervo da Discoteca L. C. Vinholes. No entanto, a diversidade das formas de representação das fontes em fonogramas dificultam o estabelecimento de padrões tanto para documentação das obras quanto para intercâmbio e integração das informações na forma de metadados (ALBUQUERQUE, 2009 apud FUENTES; GOLDBERG, 2019), e, por isso, exigem procedimentos especiais no tratamento documental.

2. METODOLOGIA

A partir da natureza colaborativa prevista pelo projeto, adotou-se como parâmetro que os processos envolvidos na digitalização deste gênero documental¹ fossem aplicados inicialmente a coleção de 78 rotações por minuto. Esse critério foi adotado devido à representatividade desta coleção em relação aos gêneros e estilos da música popular e à superabundância destes materiais no acervo. Em busca de formas mais eficientes de organização deste tipo de material, a metodologia de trabalho vem seguindo a prática de colecionadores e acervos especializados (COELHO; VELLOSO, 2019).

Deste modo, a primeira etapa da adequação do formato físico ao virtual reside no planejamento da documentação, ou seja, o registro formal dos objetos e informações do acervo. A estruturação adotada para a documentação é essencial para a gestão, controle, recuperação e adaptação de seus dados, e deve ser pensada a partir da natureza e necessidades do acervo, além de discutida e definida entre os membros do grupo LABET, grupo de pesquisa responsável pelo projeto. Quanto à distribuição dos processos nessa etapa, aderiu-se, em primeira instância, pela fotodocumentação dos rótulos dos discos em ambas faces, acompanhada pela identificação e acondicionamento nos armários e, em segunda, pela catalogação em planilhas eletrônicas separadas por gravadoras.

A etapa seguinte se concentrou no tratamento documental, procedimentos aplicados pela “Ciência de Dados” que são elementares para o estudo e análise das informações obtidas durante a documentação (PADILHA, 2014), pois possibilitam uma reflexão crítica acerca dos processos de tratamento através da detecção de padrões e obtenção de *insights* para tomadas de decisão. Assim é realizada a revisão, organização e otimização da planilha através da disputa de dados², responsável por transformar uma base com informações “caóticas” em uma base apta para publicação. A metodologia deste trabalho se configura em uma série de processos que vão desde o mapeamento, à permuta de dados brutos por dados refinados³, à padronização das informações através de técnicas

¹ Em síntese, a metodologia adotada para a digitalização da coleção de discos de 78 RPM pode ser esquematizada da seguinte maneira: fotografia dos discos; ordenação e acondicionamento nos armários; catalogação dos campos de informação deliberados em planilhas; tratamento documental (disputa de dados/*data wrangling*); criação e gestão de conteúdos das páginas eletrônicas (programação e desenvolvimento de designs).

² Termo traduzido para o português, originado na língua inglesa: *data wrangling*.

³ Os dados brutos são submetidos a um processamento de limpeza em que são removidas discrepâncias, erros de leitura dos *softwares* ou de entrada de dados, dentre outros.

como o *clustering*⁴, e ao enriquecimento que complementa os campos de informação do acervo, utilizando para isso bancos globais de dados eletrônicos em projetos como o *Wikidata*. Neste processo, em acordo com as normativas museológicas, adotou-se que os objetos fossem analisados e categorizados em suas múltiplas possibilidades de organização, de modo a não comprometer as informações dos objetos originais, mas adaptá-los às funções de busca no espaço virtual. Esse conjunto de procedimentos conduzem ao que se denomina de “validação”: a confirmação da precisão dos campos e das informações neles contidas por meio da verificação cruzada dos dados. O *software* manipulado para este fim, o tratamento através da disputa de dados, foi o OpenRefine⁵.

3. RESULTADOS E DISCUSSÃO

Com quinze colunas, doze representando os campos de informação selecionados do disco (rótulo/gravadora, número de catálogo, face, título, gênero, compositor, intérprete, acompanhamento, número de matriz, data de gravação, data de lançamento e título da outra face) e três representando outras informações (*status* de publicação, armário e coluna de acervo físico), aproximadamente 36 mil células foram tratadas através desta metodologia. Assim, diversas transformações foram aplicadas, desde a padronização no uso de letras maiúsculas e minúsculas, a correção de grafias no nome dos artistas, ao ajuste na ordenação de termos para os campos de acompanhamento, a sistematização das datas no formato numérico, dentre outras. Além disso, uma técnica de classificação hierárquica para arquivos digitais conhecida como taxonomia foi utilizada nos campos de gênero, compositor e intérpretes, resultando em novas planilhas com as informações relacionadas para a ampliação das possibilidades de busca por filtros, a partir da interpretação musicológica dos metadados. A criação destas “taxonomias” teve como finalidade a modelagem e adaptação da base para a importação através do *plugin* Tainacan⁶, para o WordPress⁷, contribuindo para uma navegação mais fluida no repositório digital da Discoteca⁸, que tem como resultado prático melhorar o acesso aos dados do acervo e auxiliar na consulta do acervo físico.

O trabalho de criação desta base, ao que se pode sobressair, não foi somente técnico, pois para além das ferramentas de validação dos metadados, pesquisas musicológicas foram necessárias na produção dos conteúdos nato-digitais. O caso dos gêneros musicais exemplifica bem isso, pois a complexidade deste tema musical, acrescido das diferentes possibilidades de grafias, torna difícil a tarefa de organizá-los: “Chotis, chotiça, schottish, xote e xotis são maneiras diferentes de grafar o mesmo estilo musical ou se referem a expressões artísticas diferentes?” (JUNQUEIRA, 2022, p. 4). Além disso, samba pode ser tanto samba-canção, quanto samba de partido alto, ou ainda samba-maxixe, enquanto Rocha Vianna pode ser tanto “Pixinguinha” quanto seu irmão “China”, ambos músicos que atuaram na era das gravações mecânicas. Nestes exemplos, as características musicais, no caso do samba, e as

⁴ O agrupamento automático de metadados pelo grau de semelhança através de algoritmos.

⁵ Disponível em: <https://openrefine.org/>. Acesso em: 20/08/2022.

⁶ Disponível em: <https://tainacan.org/>. Acesso em: 20/08/2022.

⁷ Disponível em: <https://br.wordpress.org/>. Acesso em: 20/08/2022.

⁸ Inaugurado em 16/05/2022, o repositório digital da Discoteca teve seu lançamento realizado na 20ª Semana Nacional de Museus. Disponível em: <https://acervosvirtuais.ufpel.edu.br/discoteca/78-rpm>. Acesso em: 20/08/2022.

informações biográficas, no caso dos músicos, podem contribuir para tomadas de decisão nos processos de tratamento.

Outro resultado derivado deste trabalho foi a criação de uma nova ordem taxonômica denominada “artistas”, resultante da junção dos campos “intérpretes” e “compositores”, enriquecidos pela ferramenta de reconciliação. Com este método, foi possível obter, de modo automatizado, informações inéditas à base. Foram identificados os locais e datas de nascimento e morte, cidadania e ocupações profissionais deste novo campo. Com auxílio desta ferramenta, este e os outros campos da base ainda poderiam ganhar novos complementos. O presente desafio reside em modelar e importar essas informações à base.

4. CONCLUSÕES

Em síntese, o tratamento documental foi indispensável, pois tornou possível o fornecimento de estatísticas com objetivo de obter dados mais precisos sobre as categorias de informações do acervo e de garantir que os metadados fossem consistentes. Se os processos de tratamento fossem negligenciados, bloqueios durante os processos de *upload* ocorreriam, tornando recorrentes as dificuldades de interpretação dos dados e sua confiabilidade, potencialmente obstruindo as ferramentas de busca e consequentemente inviabilizando o trabalho de pesquisadores no endereço digital.

O trabalho realizado foi um processo intermediário para a criação de uma base *online* com uma amostragem representativa de sua diversidade de dados. A partir deste avanço, ferramentas como o formulário do Tainacan agora podem ser utilizadas para facilitar a digitalização dos próximos itens. Com a ampliação desta amostragem, e assim um maior número de entradas de termos, mais efetivo e confiável se torna este processo, que seguirá com o objetivo de disponibilizar toda a coleção de 78 RPM, que totaliza cerca de 4.500 discos, e posteriormente as demais coleções do acervo, que chegam a aproximadamente 26 mil itens.

5. REFERÊNCIAS BIBLIOGRÁFICAS

COELHO, Luís Fernando Hering; VELLOSO, Rafael Henrique Soares. O acervo da Discoteca LC Vinholes, do Centro de Artes da Universidade Federal de Pelotas: música gravada e identidades no extremo sul do Brasil. In: **XXIX Congresso da Anppom-Pelotas/RS**. 2019.

FUENTES, Eduardo & GOLDBERG, Luiz Guilherme. **Dados em discos de goma-laca e sua aplicação em estudos musicológicos e arquivísticos**. CIC, 2019.

JUNQUEIRA, Felipe França de Andrade. **Um relato sobre a gestão da Discoteca L. C. Vinholes**: desafios da digitalização de acervos fonográficos. 14º Fórum de Museus (RS), 2022.

PADILHA, Renata Cardozo. **Documentação museológica e gestão de acervo**. Florianópolis: FCC, 2014.

SAYÃO, Luis Fernando. Digitalização de acervos culturais: reuso, curadoria e preservação. **Informação digital e suas diversas abordagens pela ótica de um cientista da informação**, p. 269, 2016.