

## Bases de dados qualificadas para a descoberta de conhecimentos úteis à gestão de acervos culturais

*Qualified databases for discovering knowledge to support the management of cultural collections*

**Daniela Lucas da Silva Lemos<sup>1</sup>**



<http://lattes.cnpq.br/9280443047358807>



<https://orcid.org/0000-0003-1565-7366>

**Henrique Monteiro Cristovão<sup>2</sup>**



<http://lattes.cnpq.br/5035919384923489>



<https://orcid.org/0000-0003-2011-7022>

### Resumo

Padrões de documentação são usados amplamente nas ações de organização e tratamento da informação para a constituição de bases de dados com qualidade, na intenção de se obter informações estratégicas para a gestão de acervos culturais. O presente estudo relata a experiência com a construção das bases de dados do acervo do Fundo de Cultura do Espírito Santo, sob a gestão da Secretaria de Estado da Cultura, disponibilizadas na rede de provedores da plataforma MEDIATECA Capixaba. Assim, a partir da análise exploratória destas bases, a pesquisa objetivou identificar agrupamentos coesos nos autores contemplados nos editais por meio de similaridades nos eixos temáticos de seus respectivos projetos. Metodologicamente, a modelagem de metadados, incluindo o uso de vocabulários controlados e do padrão de metadados *Dublin Core*, contou com a técnica conhecida como Matriz de Restrições para a elaboração de perfis de aplicação para a constituição das bases de dados qualificadas. Por conseguinte, as redes de informações projetadas a partir da exploração de dados dessas bases foram construídas a partir do método de Análise de Redes Sociais. A análise de redes de informação revelou relações entre autores, editais e temas, evidenciando padrões de participação e distribuição de recursos. Além disso, a análise exploratória de dados permitiu identificar tendências para subsidiar políticas públicas culturais. Assim, a pesquisa reforça a importância dos padrões de documentação para qualificar os dados e contribuir para a gestão da informação cultural em ambientes digitais, indicando caminhos para aprimoramentos na organização de acervos e na democratização do acesso à cultura.

**Palavras-chave:** padrões de documentação; acervos culturais; qualidade de dados; análise de redes de informação; descoberta de conhecimento.

### Abstract

Documentation standards are widely used in information organization and processing actions to create quality databases, with the intention of obtaining strategic information for the management of cultural collections. This study reports the experience with the construction of databases of the Espírito Santo Culture Fund collection, under the management of the State Secretariat of Culture, made available on the

<sup>1</sup> Doutora em Ciência da Informação, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil. Professora e Pesquisadora, Universidade Federal do Espírito Santo (UFES), Vitória, ES, Brasil. [daniela.l.silva@ufes.br](mailto:daniela.l.silva@ufes.br).

<sup>2</sup> Doutor em Ciência da Informação, Universidade de Brasília (UnB), Brasília, DF, Brasil. Professor, Universidade Federal do Espírito Santo (UFES), Vitória, ES, Brasil. [henrique.cristovao@ufes.br](mailto:henrique.cristovao@ufes.br).

network of providers of the MEDIATECA Capixaba platform. Thus, based on the exploratory analysis of these databases, the research aimed to identify cohesive groupings in the authors contemplated in the calls for proposals through similarities in the thematic axes of their respective projects. Methodologically, the metadata modeling, including the use of controlled vocabularies and the Dublin Core metadata standard, relied on the technique known as the Restriction Matrix to develop application profiles for the creation of qualified databases. Consequently, the information networks designed from the exploration of data from these databases were built using the Social Network Analysis method. The analysis of information networks revealed relationships between authors, notices and themes, highlighting patterns of participation and distribution of resources. In addition, the exploratory data analysis allowed the identification of trends to support cultural public policies. Thus, the research reinforces the importance of documentation standards to qualify data and contributes to the management of cultural information in digital environments, indicating paths for improvements in the organization of collections and the democratization of access to culture.

**Keywords:** documentation standards; cultural collections; data quality; information network analysis; knowledge discovery.

## 1 INTRODUÇÃO

A padronização é um elemento crucial, discutido amplamente na Ciência da Informação, especialmente na área de Organização e Representação da Informação e do Conhecimento. Padrões de documentação são usados nessa perspectiva para apoiar práticas de catalogação e são essenciais na crescente digitalização de coleções de objetos culturais (Gilliland, 2016). Na plataforma digital de agregação de dados culturais, denominada MEDIATECA Capixaba (Lemos; Martins; Macedo, 2024), os padrões de documentação são fundamentados em teorias de catalogação descritiva e de assunto, focando na criação e no uso de metadados, vocabulários controlados e regras de catalogação pertinentes a cada perfil de aplicação.

O presente estudo relata a experiência com as bases de dados do acervo do Fundo de Cultura do Espírito Santo (Funcultura), sob a gestão da Secretaria de Estado da Cultura (SECULT/ES), implementadas em um dos repositórios institucionais da rede de provedores da MEDIATECA Capixaba, para a qual se buscou qualificar seus dados a partir de padrões de documentação na intenção de se obter informações e conhecimentos úteis à gestão de seu acervo cultural.

O Funcultura foi criado pela Lei Complementar Estadual nº458 de 21/10/2008 e, desde então, é o principal mecanismo do Governo do Estado para o fomento e incentivo a programas, projetos e ações culturais (Espírito Santo, 2023). Os recursos do Fundo são aplicados em projetos, ações e programas baseados em três dimensões culturais: cultura como expressão simbólica de um povo, cultura como direito e cidadania, e cultura como

economia que gera renda e trabalho. Nesse ensejo, há uma necessidade real de otimização do acesso à cultura frente ao mundo contemporâneo, com sua dinâmica digital que rege o cotidiano da sociedade e de grande parte dos cidadãos. Assim, a entrada deste acervo no projeto Midiateca Capixaba é justamente para oportunizar estudo, análise, pesquisa e organização do conjunto de dados acerca dos editais, projetos e produtos culturais em bases de dados qualificadas, possibilitando a sua exploração para fins de difundir informações úteis à sociedade bem como conhecimentos que permitam aos gestores tomarem decisões pertinentes para a elaboração de políticas públicas que sejam de interesse ao cidadão capixaba, em especial.

A análise de redes de informação tem sido usada para a projeção visual de descoberta de conhecimento em bases de dados e, principalmente, por ser uma ferramenta metodológica que viabiliza enxergar o que outras abordagens não permitem (Wasserman; Faust, 1994; Higgins; Ribeiro; 2018), sendo também uma boa estratégia para investigar conexões entre elementos sociais e culturais. Contudo, ela precisa estar suportada por redes com dados bem estruturados, com conteúdos bem formatados e com integridade, além de contar com o apoio de informações estatísticas para auxiliar na interpretação, o que depende de dados bem descritos.

Assim, o problema de pesquisa encontra-se em saber de que maneira o uso de padrões de documentação consegue favorecer o tratamento de dados endereçados à análise exploratória de dados, de modo a entender suas características, identificar padrões, detectar *outliers* e formular hipóteses preliminares. Por conseguinte, o objetivo geral da pesquisa é identificar agrupamentos coesos nos autores contemplados nos editais por meio de similaridades nos eixos temáticos de seus respectivos projetos, a partir da análise exploratória das bases de dados do Funcultura.

## 2 PROCEDIMENTOS METODOLÓGICOS

A pesquisa é de natureza aplicada, de abordagem quali-quantitativa, e de caráter exploratório e descritivo, envolvendo bases de dados do acervo cultural do Funcultura, disponibilizado em repositório digital institucional. Essas bases são exploradas e seus dados tratados para a geração de gráficos estatísticos e redes de informação tripartida,

bipartida e monopartida. Os procedimentos metodológicos foram conduzidos em duas etapas, a saber: (i) modelagem e criação das bases de dados; e (ii) análise e exploração dos dados.

A primeira etapa metodológica pode ser considerada a espinha dorsal para a obtenção de dados curados para transformação em fontes de informações qualificadas e destinadas à plataforma de agregação MEDIATECA Capixaba, bem como matéria-prima para gestores da área da cultura em produzir informações estratégicas à tomada de decisões ao Fundo.

Assim sendo, a etapa de Modelagem do Esquema de Metadados foi assistida por meio da técnica de criação de Matriz de Restrições (Baker; Nilsson; Johnston, 2008), em que se integra e prevê um conjunto de registros de metadados que são instâncias válidas para um dado perfil de aplicação, neste caso, o acervo cultural do Funcultura.

A partir do ambiente de desenvolvimento da tecnologia Tainacan<sup>3</sup>, pôde-se criar o esquema de dados à luz da Matriz de Restrições. Nessa perspectiva, a produção das bases de dados do Funcultura contou com o uso de padrões de documentação (Lemos; Martins; Macedo, 2024) na organização e representação dos documentos do Fundo, destacando o uso de vocabulários controlados (taxonomias e arquivos de autoridade) e do padrão de metadados *Dublin Core* como um modelo de referência, visando obter informações normalizadas e consistentes para a extração de *insights* valiosos à instituição.

Na segunda etapa metodológica, voltada à análise exploratória, os dados oriundos das coleções do acervo do Funcultura, usados na pesquisa, foram organizados em três tabelas relacionadas. A primeira tabela representa as informações dos 876 autores proponentes dos projetos que foram contemplados. A segunda tabela contém as informações dos 1370 projetos inscritos que foram contemplados bem como o edital ao qual participaram, sendo conectada à primeira tabela por meio de um identificador único do autor proponente. A terceira tabela possui informações sobre 262 produtos vinculados aos projetos contemplados, sendo conectada à segunda tabela por meio de um

---

<sup>3</sup> Tainacan é o *software* usado na implementação das bases de dados na rede de provedores da MEDIATECA Capixaba. Disponível em: <https://tainacan.org/>.

identificador único do projeto inscrito. Todos os dados da pesquisa, inclusive os arquivos obtidos nos processamentos dos resultados estão disponíveis em repositório para acesso público<sup>4</sup>.

Com apoio do *software OpenRefine*<sup>5</sup>, os dados das três tabelas passaram por um processo de preparação e de transformação com foco na qualidade e padronização. Essa etapa envolveu ações, tais como a limpeza de dados com símbolos irrelevantes, eliminação de registros em branco, padronização de nomes de variáveis, e a criação de variáveis a partir da extração de dados existentes para representar, com menor granularidade, informações sobre o ano do edital, a categoria cultural e o eixo temático ao qual os projetos se enquadram. Ao todo foram organizados, por meio de um vocabulário controlado do tipo taxonomia, oito categorias culturais e 49 eixos temáticos dentre os 1370 projetos contemplados. A distribuição da quantidade de projetos nessa taxonomia pode ser vista no Gráfico 1, na seção de resultados.

Com foco na identificação de agrupamentos coesos entre os autores contemplados, por meio de similaridades nas categorias culturais e eixos temáticos de seus respectivos projetos, criou-se um mapeamento dos dados das tabelas para o formato de uma rede de informação tripartida contendo os seguintes nós: autores contemplados, editais de inscrição, e eixos temáticos. O referido mapeamento, Quadro 1, foi processado no *software OpenRefine* e gerou um arquivo no formato GML<sup>6</sup> para ser interpretado pelo *software Gephi*<sup>7</sup>.

No próprio *software Gephi*, a partir da rede tripartida gerada pelo mapeamento do Quadro 1, foram aplicadas projeções bipartidas para a obtenção de outras redes simplificadas, especialmente uma rede monopartida de autores contemplados.

---

<sup>4</sup> Repositório de dados da pesquisa. Disponível em:

<https://drive.google.com/drive/folders/1AQUF007OGKcxI1097FmYHGJy0v3Lyeku?usp=sharing>.

<sup>5</sup> OpenRefine é considerado um *Interactive Data Transformation*, sendo uma ferramenta de código aberto utilizada para a limpeza e transformação de volumes massivos de dados estruturados. Disponível em:

<https://openrefine.org/>.

<sup>6</sup> *Graph Modelling Language* (GML) é um formato para representação de grafos de fácil leitura por humanos e com uma capacidade semântica razoável para configurar as características da rede, dos nós e das arestas. Disponível em: [https://en.wikipedia.org/wiki/Graph\\_Modelling\\_Language/](https://en.wikipedia.org/wiki/Graph_Modelling_Language/).

<sup>7</sup> Gephi é um *software* de código aberto utilizado para visualização, análise e manipulação de redes e grafos. Disponível em: <https://gephi.org/>.

Quadro 1 - Mapeamento de dados para o formato de rede executado no *software* OpenRefine

```
graph [
comment "Rede de contemplados por eixo e chamada do edital"
directed 0

node [ id {{jsonize(cells.idContemplado.value)}}
categoria "idContemplado"
grupo "contemplado"
eixo {{jsonize(cells.eixo.value)}}
chamada {{jsonize(cells.chamadaEdital.value)}} ]

node [ id {{jsonize(cells.eixo.value)}}
categoria "eixo"
grupo "eixo-chamada" ]

node [ id {{jsonize(cells.chamadaEdital.value)}}
categoria "chamada edital"
grupo "eixo-chamada" ]

edge [ source {{jsonize(cells.idContemplado.value)}} target
{{jsonize(cells.eixo.value)}} ]

edge [ source {{jsonize(cells.idContemplado.value)}} target
{{jsonize(cells.chamadaEdital.value)}} ]
]
```

Fonte: Dados da pesquisa (2024).

O trabalho recorreu à técnica de análise de redes sociais (ARS) para a descoberta de conhecimento e, principalmente, por ser uma ferramenta metodológica que viabiliza enxergar o que outras abordagens não permitem (Wasserman; Faust, 1994; Higgins; Ribeiro, 2018), sendo também uma boa estratégia para investigar conexões entre elementos sociais e culturais. Aliado a ARS, a inspeção visual foi amplamente usada por ser uma ação que direciona a investigação em novas explorações. Contudo, ela precisa estar suportada por redes com uma boa formatação dos dados, além de contar com o apoio de gráficos estatísticos para auxiliar na interpretação. Dessa forma, foram usados os *softwares* Gephi, Looker Studio<sup>8</sup> e uma planilha eletrônica nessa etapa de análise e interpretação, cujos artefatos visuais e resultados são apresentados na próxima seção.

### 3 RESULTADOS E DISCUSSÃO

O levantamento de requisitos (Quadro 2) foi fundamental para a organização dos dados do acervo em três coleções (bases de dados): projeto, contemplado (autor) e

<sup>8</sup> Looker Studio é um *software* gratuito para *data visualization* e criação de relatórios e painéis interativos a partir de dados. Disponível em: <https://lookerstudio.google.com/>.

produto. Cada coleção foi organizada com as suas principais entidades que demandaram metadados para as suas descrições e representatividades nas bases de dados.

**Quadro 2 - Especificação de requisitos das principais entidades das bases de dados do Fundo**

| Entidade   | Descrição   | Coleção Tainacan         |
|--|---|--------------------------|
| Dados gerais   | Fornecer informações administrativas sobre os lançamentos anuais dos editais, suas respectivas categorias e eixos.  | Funcultura - Projeto     |
| Projeto contemplado  | Fornecer informações sobre o projeto contemplado, incluindo título, proponente contemplado, o valor recebido, contatos do agente cultural e seu relatório final de prestação de contas.   |                          |
| Produto [outra coleção a ser criada no Tainacan, e representada via tipo de metadado relacionamento] | Fornecer informações sobre as diferentes ações entregues pelo proponente em forma de resultados do projeto contemplado. Logo, um edital pode se desdobrar em vários produtos. O produto, por sua vez, é o objetivo de um edital, ou seja, o resultado esperado pelo proponente ao submeter uma proposta de projeto. |                          |
| Contemplado  | Pessoas envolvidas com o desenvolvimento do produto participante de um edital e projeto.  | Funcultura – Contemplado |
| Categoria de Metadados   | Com título, descrição e tipo de produto (lista de seleção).   | Funcultura – Produto     |
| Acervo   | Fornecer informações sobre um conjunto de objetos, históricos ou não, que necessitam estar reunidos para ocorrer sua compreensão.   |                          |
| Audiovisual  | Fornecer informações sobre um produto que necessariamente apresente imagem e som de forma engajada.   |                          |
| Documento técnico  | Fornecer informações registradas em papel ou meio eletrônico que possa ser usado para consulta, estudo ou validação técnica.  |                          |
| Evento   | Fornecer informações sobre um acontecimento que reúna pessoas, presencialmente ou virtualmente, para determinado fim.   |                          |
| Livro  | Fornecer informações sobre uma publicação física ou digital, composto por páginas, que necessariamente deve ter ISBN ou ISSN.   |                          |
| Mestres da Cultura Popular   | Fornecer informações biográficas de personalidades da cultura capixaba.   |                          |
| Música   | Fornecer informações de uma prática cultural que combine sons e ritmos.   |                          |
| Site   | Fornecer informações sobre um sítio eletrônico criado como ou para o produto, acessado através de endereço HTTP ou HTTPS, disponível ou não, na web.  |                          |

Fonte: Dados da pesquisa (2024).

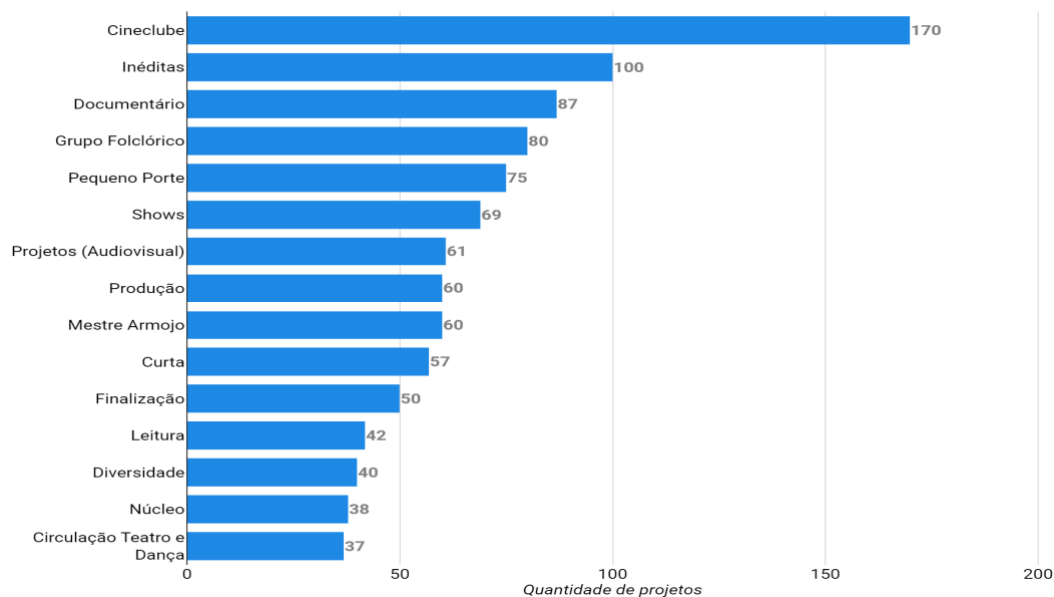
A Matriz de Restrições com a relação dos metadados e seus respectivos perfis de aplicação foi elaborada a partir desses requisitos, evidenciando as decisões de modelagem para cada elemento presente na matriz, tornando-o, assim, uma representação da informação e do conhecimento diferenciada para as bases de dados do Funcultura, em que refletiram em metadados normalizados, padronizados e qualificados à luz dos vocabulários controlados propostos na presente pesquisa.

As taxonomias criadas para o acervo digital do Funcultura são em nove: autor de livro; categoria e eixo; estilo musical; local de atuação dos mestres; local de gravação; localidade; localidade do evento; mestre da cultura popular; e tipo de evento. Para além das taxonomias, vários arquivos de autoridade foram modelados para uso nos valores de dados de alguns metadados, as cardinalidades indicadas, os principais filtros atribuídos, a obrigatoriedade ou não na entrada de dados, as orientações de uso do metadado na perspectiva do usuário, entre outras configurações, garantindo, assim, a qualidade na catalogação dos dados e, conseqüentemente, na recuperação da informação.

Finalmente, uma vez as bases de dados criadas com padrões de documentação fundamentados em princípios teórico-metodológicos da Ciência da Informação, a exemplo da modelagem de taxonomias apropriadas ao contexto de uso, tornou-se possível obter informações relevantes à Instituição a partir de dados qualificados, como pode-se observar nos gráficos e nas redes de informação projetados a seguir.

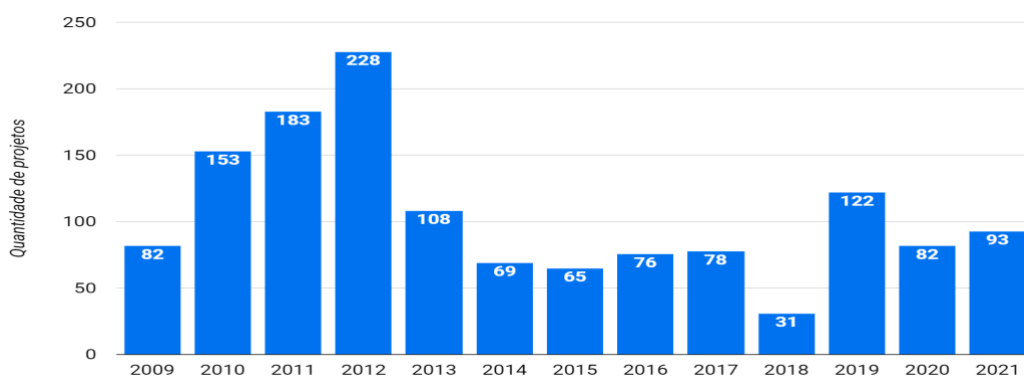
Os resultados de um breve levantamento estatístico sobre a quantidade de projetos distribuídos para os eixos temáticos do Fundo são mostrados no Gráfico 1. Destaca-se também a distribuição da quantidade de projetos ao longo dos anos de publicação dos editais, apresentada no Gráfico 2.

**Gráfico 1 - Distribuição dos projetos culturais nos 15 principais eixos temáticos**



Fonte: Dados da pesquisa (2024).

**Gráfico 2 - Distribuição da quantidade de projetos ao longo dos anos**



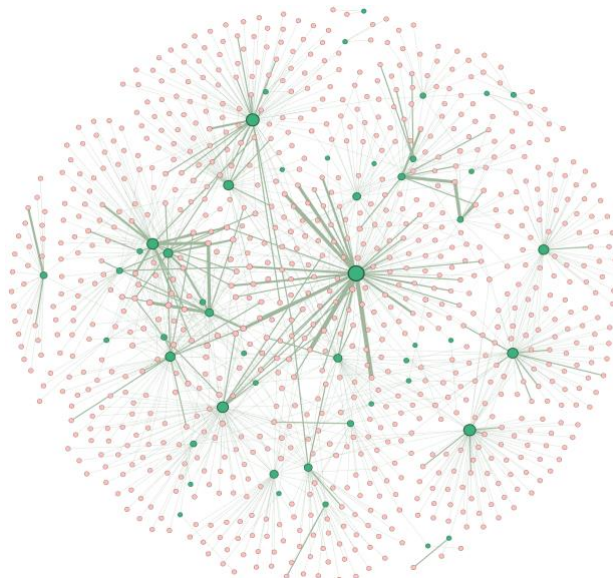
Fonte: Dados da pesquisa (2024).

No que diz respeito às redes de informação, a rede bipartida é composta por nós avermelhados, que representam os autores contemplados, e nós esverdeados, que representam os eixos temáticos, conforme representado na Figura 1.

Essa rede destaca o tamanho dos nós de forma proporcional ao maior grau ponderado. Destaca também a espessura de arestas por aquelas com maior quantidade de conexões coincidentes entre autores com projetos de mesmo eixo temático. Nesse caso, o eixo temático “Cineclube”, na parte central da rede, é destacado, por estar associado a um maior número de autores com projetos contemplados e encontra-se em sintonia com o destaque em ocorrências absolutas que aparece no Gráfico 1. Além disso, devido a múltiplas conexões de muitos nós avermelhados (autores), a rede revela a

existência de um grande número de autores com projetos associados a diversos eixos temáticos.

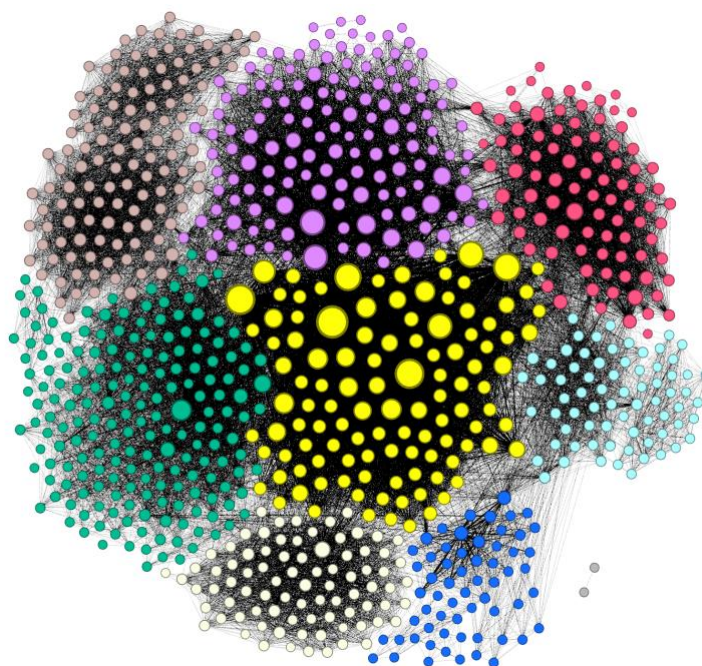
**Figura 1 - Rede bipartida com destaque para eixos com maior incidência de inscrições de autores**



Fonte: Elaboração própria (2024).

A rede monopartida de autores contemplados, oriunda de uma projeção bipartida, é mostrada na Figura 2.

**Figura 2 - Rede monopartida de autores contemplados**



Fonte: Elaboração própria (2024).

Para a obtenção dos nove grupos apresentados, destacados em cores diferentes, foi aplicado um algoritmo de clusterização e um algoritmo de distribuição disponíveis no *software* Gephi. O tamanho dos nós foi apresentado de forma proporcional ao grau ponderado dos mesmos.

Em consonância com a rede da Figura 2, a Tabela 1 apresenta a ocorrência dos eixos temáticos nos oito principais grupos de autores contemplados mostrados na Figura 2. Para a obtenção de um resultado mais sistematizado, foram considerados apenas os eixos temáticos cujo percentual de ocorrência de contemplados é maior do que 10% sobre o restante dos eixos temáticos do mesmo grupo.

Os agrupamentos resultantes na rede da Figura 2, analisados por inspeção visual com auxílio da Tabela 1, revelam alguns padrões de incidência de autores que participam regularmente de projetos em eixos temáticos, tais como "Documentário", "Projetos (Audiovisual)", "Curta" e "Finalização", no grupo de cor rosa; "Circulação Teatro e Dança", "Residência Teatro e Dança", "Circo", "Infantil" e "Lona de circo", no grupo cor azul escuro. Por outro lado, o grupo de cor amarela, destaca o eixo "Cineclube" de forma isolada sugerindo uma interpretação de que seus autores (maiores graus ponderados da rede) tendem a permanecer no mesmo eixo.

**Tabela 1 - Distribuição dos eixos temáticos nos oito principais grupos da rede da Figura 2**

|                          |     |                              |     |
|--------------------------|-----|------------------------------|-----|
| <b>Grupo cor amarela</b> |     | <b>Grupo cor azul escuro</b> |     |
| Cineclube                | 96% | Circulação Teatro e Dança    | 26% |
| <b>Grupo cor branca</b>  |     | Residência Teatro e Dança    | 20% |
| Shows                    | 61% | Circo                        | 18% |
| Setorial (Literatura)    | 22% | Infantil                     | 18% |
| <b>Grupo cor rosa</b>    |     | Lona de circo                | 15% |
| Documentário             | 26% | <b>Grupo cor verde</b>       |     |
| Projetos (Audiovisual)   | 26% | Pequeno Porte                | 26% |
| Curta                    | 16% | Núcleo                       | 14% |
| Finalização              | 13% | Leitura                      | 12% |
| <b>Grupo cor bege</b>    |     | Diversidade                  | 12% |
| Grupo Folclórico         | 54% | <b>Grupo cor azul claro</b>  |     |

|                           |     |
|---------------------------|-----|
| Mestre Armojo             | 43% |
| <b>Grupo cor vermelha</b> |     |
| Inéditas                  | 64% |

|          |     |
|----------|-----|
| Produção | 49% |
| Bolsa    | 37% |

Fonte: Dados da pesquisa (2024).

A análise exploratória de dados (Castro; Ferrari, 2017) é empregada para examinar, resumir e visualizar os dados para entender suas características principais, identificar padrões, detectar *outliers* e formular hipóteses preliminares, utilizando-se de gráficos, medidas-resumo e análises de distribuição (histogramas, redes de informação etc.) Os Gráficos 1 e 2 e as Redes (Figuras 1 e 2) mostram que a análise exploratória de dados pode ser feita a partir da sumarização realizada em categorias pré-concebidas (com o uso de vocabulário controlado nas bases de dados). Desse modo, é possível perceber que, o que distingue dado ou um conjunto de dados de informação é o conhecimento que ela propicia ao tomador de decisões (Hjørland, 2018).

Logo, na base de dados de Projetos, por exemplo, nota-se que o dado “identificação dos projetos”, representado pelo metadado “ID\_Projeto”, é classificado numa categoria oriunda da taxonomia Categoria-Eixo. Por meio desta forma de modelagem de dados, foi possível gerar visualizações de informações importantes acerca das categorias culturais e eixos temáticos envolvidos nos editais, projetos e grupos de contemplados. O gestor de cultura da Instituição poderá, a partir de conhecimentos e habilidades técnicas, tomar decisões a partir das informações geradas frente às bases de dados, incluindo elaboração de estudos do perfil dos contemplados participantes de editais; extração de *insights* valiosos sobre o perfil do público, sua preferência em ações culturais, entre outras. Enfim, essas informações podem ser usadas pelos gestores para melhorar a experiência com os editais e tipos de projetos, além da possibilidade de criar editais personalizados e direcionar melhor recursos públicos e estratégias de captação de produtos e serviços culturais para o Estado.

Do mesmo modo, a partir da relação entre os dados “ano dos projetos” e “identificação dos projetos” (Gráfico 2), tornou-se possível projetar a distribuição dos projetos ao longo dos períodos na trajetória do Funcultura. Assim, o gestor do Fundo também poderá, a partir desse potencial indicador, tomar decisões de investir em

programas e políticas públicas para setores culturais que se sobressaíram nos últimos anos; além de identificar tendências de interesse dos contemplados frente aos produtos desenvolvidos anualmente a partir dos projetos e seus eixos temáticos. Esses *insights* podem orientar o desenvolvimento de curadorias digitais na plataforma Midiateca, com trajetos narrativos temáticos do Fundo que estimulem a participação de artistas capixabas nos editais ofertados pelo governo do Estado.

### 3 CONSIDERAÇÕES FINAIS

O trabalho apresentou o percurso da construção das bases de dados padronizadas e qualificadas para o Fundo de Cultura e, utilizando-se de análise exploratória, identificou agrupamentos coesos nos autores contemplados nos editais por meio de similaridades nos eixos temáticos de seus respectivos projetos.

O estudo reforçou a importância dos padrões de documentação para qualificar os dados adotando modelagem de metadados e vocabulários controlados, estruturando e padronizando as bases de dados.

A análise de redes de informação revelou relações entre autores, editais e temas, evidenciando padrões de participação e distribuição de recursos. Além disso, a análise exploratória de dados permitiu identificar tendências para subsidiar políticas públicas culturais. Assim, a pesquisa contribui para a gestão da informação cultural em ambientes digitais, indicando caminhos para aprimoramentos na organização de acervos e na democratização do acesso à cultura.

### REFERÊNCIAS

BAKER, Thomas; NILSSON, Mikael; JOHNSTON, Pete. **The singapore framework for dublin core application profiles**. 2008. Disponível em: <http://dublincore.org/documents/singapore-framework/>. Acesso em: 04 mar. 2025.

CASTRO, Leandro Nunes de; FERRARI, Daniel Gomes. **Introdução à mineração de dados**. São Paulo: Saraiva Educação SA, 2017.

ESPÍRITO SANTO. Instituto Jones dos Santos Neves. **Análise executiva do FUNCULTURA/ES**. Vitória, ES, 2023. 2. ed. 61 p.; il. tab. (Relatório de Pesquisa).

GILLILAND, Anne J. Setting the Stage. *In*: BACA, Murta. (ed.). **Introduction to metadata**. 3. ed. Los Angeles: Getty Research Institute, 2016. Disponível em: <https://www.getty.edu/publications/intrometadata/setting-the-stage/>. Acesso em: 04 mar. 2025.

HIGGINS, Silvio Salej; RIBEIRO, Antonio Carlos Andrade. **Análise de redes em Ciências Sociais**. Brasília: Enap, 2018. Disponível em: [https://repositorio.enap.gov.br/bitstream/1/3337/1/Livro\\_Analise%20de%20Redes%20em%20Ci%C3%Aancias%20Sociais.pdf](https://repositorio.enap.gov.br/bitstream/1/3337/1/Livro_Analise%20de%20Redes%20em%20Ci%C3%Aancias%20Sociais.pdf). Acesso em: 04 mar. 2025

HJØRLAND, Birger. Data (with big data and database semantics). **Knowledge Organization**, v. 45, n. 8, p. 643-652, 2018. Disponível em: <https://www.isko.org/cyclo/data>. Acesso em: 04 mar. 2025

LEMOS, Daniela Lucas da Silva; MARTINS, Dalton Lopes; MACEDO, Dirceu Flávio. Plataforma de agregação midiateca capixaba: projeto cultura digital nas instituições de memória do Espírito Santo. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 24., 2024, Vitória. **Anais [...]**. Vitória: ANCIB, 2024. Disponível em: <https://enancib.ancib.org/index.php/enancib/xxivenancib/paper/viewFile/2565/1516>. Acesso em: 04 de mar. 2025.

WASSERMAN, Stanley; FAUST, Katherine. **Social network analysis: methods and applications**. Cambridge, England; New York: Cambridge University Press, 1994.



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

#### Como citar este trabalho:

LEMOS, Daniela Lucas da Silva; CRISTOVÃO, Henrique Monteiro. Bases de dados qualificadas para a descoberta de conhecimentos úteis à gestão de acervos culturais. *In*: WORKSHOP DE INFORMAÇÃO, DADOS E TECNOLOGIA, 8., Marília, SP. **Anais [...]**. Marília, SP: Universidade de Marília, 2025. DOI: <https://doi.org/10.22477/viii.widat.238>.