

# PADRONIZAÇÃO DE METADADOS ADOTADOS NO PORTAL AGREGADOR DE CONTEÚDO CIENTÍFICO OASISBR

---

*GABRIEL SILVEIRA MARQUES  
BLENA ESTEVAM DOS SANTOS  
MILLENA CORDEIRO MATOS DE LIMA  
CÁSSIO TEIXEIRA DE MORAIS  
MARCEL GARCIA DE SOUZA*



This work is licensed under a  
[Creative Commons Attribution  
4.0 International License](https://creativecommons.org/licenses/by/4.0/).



MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÃO



Porto, 01 de outubro de 2024.

# Autores

---



**Gabriel Silveira Marques**  
(gabrielmarques@ibict.br)



**Cássio Teixeira de Moraes**  
(cassiomoraes@ibict.br)



**Blena Estevam dos Santos**  
(blenasantos@ibict.br)



**Millena Cordeiro Matos de Lima**  
(millena lima@ibict.br)



**Marcel Garcia de Souza**  
(marcelsouza@ibict.br)

# I - Contexto

---

- ➡ Fontes de informação;
- ➡ Portais agregadores;
- ➡ Oasisbr;
- ➡ Padronização de metadados;
- ➡ Interoperabilidade.

# II - Pesquisa

---

- ➡ Objetivos e justificativa;
- ➡ Metodologia;
- ➡ Resultados;
- ➡ Considerações finais;
- ➡ Referências.

# Fontes de informação

---

**Repositórios digitais e diretórios de produção científica** constituem pilares fundamentais para o armazenamento, recuperação e disseminação do conhecimento científico. Eles abarcam um amplo espectro de produções, incluindo **teses, dissertações, monografias, artigos, conferências, livros, capítulos de livros, relatórios, preprints e conjuntos de dados**, os quais são cruciais tanto para o avanço da ciência quanto para a democratização do **acesso ao conhecimento**.

# Portais agregadores

As **plataformas agregadoras** de conteúdo são ferramentas valiosas para disponibilização de **conteúdo científico**. Iniciativas como a **Biblioteca Digital Brasileira de Teses e Dissertações (BDTD)** e o **Portal Brasileiro de Publicações e Dados Científicos em Acesso Aberto (Oasisbr)**, desenvolvidos pelo **Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)**, procuram oferecer essa alternativa reunindo em um só diretório inúmeras fontes de conteúdo acadêmico.

Conheça o Oasisbr:  
<https://oasisbr.ibict.br/vufind/>



Conheça a BDTD:  
<https://bdttd.ibict.br/vufind/>



# Oasisbr

---

Especificamente o Oasisbr, possui a **função de agregar** e disponibilizar gratuitamente uma vasta gama de documentos científicos de **fontes brasileiras**, promovendo maior **visibilidade, acesso e intercâmbio** à pesquisas e pesquisadores brasileiros, além de **estimular a criação** e abertura de plataformas científicas brasileiras cientificamente adequadas, atuando como provedor nacional de **informação de qualidade** (Ibict, 2024).

# Oasisbr

---

- 12 tipos de documentos científicos;
- Mais de 4,5 milhões de documentos;
- Acesso Aberto;
- Mais de 1500 fontes.

# Padronização de metadados

---

Os metadados representam um **componente fundamental** na Ciência da Informação (CI), desempenhando **um papel vital** na organização, no **acesso e na preservação** dos dados digitais.



# Interoperabilidade

---

Sistemas como **OJS**, **DSpace** e **DataVerse** surgiram como **soluções inovadoras** para facilitar a **gestão e a disseminação** da produção científica. A **interoperabilidade**, nesse contexto, refere-se à **capacidade de diferentes softwares** e organizações trabalharem juntos para proporcionar um **acesso mais amplo e integrado** à informação científica (Paganine et al., 2020).

- Padrão Dublin Core;
- Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH);
- Heterogeneidade dos metadados.

## II - Pesquisa

---

- ➡ Objetivos e justificativa;
- ➡ Metodologia;
- ➡ Resultados;
- ➡ Considerações finais;
- ➡ Referências.

# Objetivos

---

- O objetivo norteador é analisar os padrões de metadados utilizados no Oasisbr, utilizando a plataforma de busca corporativa de código aberto baseada em java, o Solr;
- Identificar as principais lacunas e inconsistências presentes e possíveis problemáticas com a qualidade dos metadados das fontes inseridas no Oasisbr;
- A partir do mapeamento, almeja-se fornecer recomendações para a adoção de práticas mais uniformes na gestão de metadados, contribuindo assim com a melhora da interoperabilidade entre os sistemas e, consequentemente, na eficácia da recuperação de informações científicas.

# Justificativa

---

- Portal de iniciativa de país lusófono;
- Apresenta uma vasta tipologia de documentos;
- Trata-se de uma fonte de informação científica de acesso aberto.

# Metodologia

---

- Pesquisa quantitativa e descritiva;
- Foi realizada uma análise de frequência dos campos de metadados para identificar os mais utilizados, bem como a existência de possíveis lacunas ou inconsistências na indexação dos documentos.
- A coleta dos dados ocorreu entre os dias 27 de fevereiro a 11 de março de 2024 e os relatórios encontram-se disponíveis na planilha de dados (Marques et. al., 2024).

# Metodologia

- A fase quantitativa envolveu a coleta de dados primários através do Solr, focando no total de padrões de metadados presentes nas diversas fontes do Oasisbr.

The screenshot displays the Solr Admin interface. On the left, the 'Request Handler (q)' is set to '/select'. The 'q' field contains a query, and the 'select' method is chosen. The 'sort' field is set to 'start\_time'. The 'start' and 'end' fields are set to 0 and 10, respectively. The 'wt' field is set to 'json'. The 'debugQuery' checkbox is checked. The 'Execute Query' button is visible. On the right, the JSON response is shown, containing metadata about a document. Red arrows 1, 2, and 3 point to the 'q' field, the 'select' method, and the 'response' field respectively.

```
{
  "responseHeader": {
    "status": 0,
    "QTime": 1488,
    "params": {
      "q": "...",
      "indent": "on",
      "qt": "lucene",
      "wt": "json",
      "_: "/>
```

# Resultados

- Tabela 1 - Frequência dos campos de metadados

METADADO	% NÃO COBERTURA	OBSERVAÇÃO
publishDate	0,00%	Data de publicação
spelling	0,00%	Resumo + dados referenciais
spellingShingle	0,00%	Título + primeiro autor + Palavras-chave
reponame_str	0,00%	Nome da fonte (revista ou repositório)
network_acronym_str	0,00%	Acronimo da instituição no Oasisbr
id	0,00%	ID no Oasisbr
title	0,02%	Título do documento
instname_str	0,04%	Título da instituição fonte
format	0,42%	Tipo de documento
author_facet	7,35%	Todos os autores
eu_rights_str_mv	17,78%	Direito autoral e o acesso livre
oai_identifier_str	18,09%	Link OAI do registro
dc.source.none.fl_str_mv	18,11%	Dados referenciais do documento (título da fonte; informações da edição e página; ISSN; link; instituição)
institution	18,12%	Sigla da instituição fonte
dc.identifier.uri.fl_str_mv	18,13%	Link do documento
dc.language.iso.fl_str_mv	18,91%	Idioma
description	31,24%	Resumo
identifier_str_mv	44,51%	DOI do documento (metadado que traz poucos registros mas são importantes)
dc.identifier.citation.fl_str_mv	74,39%	Referência (apenas para documentos em BDTD's)
dc.identifier.issn.pt_BR.fl_str_mv	98,53%	ISSN da revista (metadado que traz poucos registros mas são importantes)
dc.identifier.isbn.pt_BR.fl_str_mv	99,91%	ISBN do livro (metadado que traz poucos registros mas são importantes)

Fonte: dados da pesquisa (2024).

# Resultados

---

- Uso concentrado do padrão de metadados Dublin Core;
- Identificadores únicos dos itens com pouca cobertura, tal como ISBN, ISSN e DOI (99,91%, 98,53% e 44,51%);
- Referência do item (dc.identifier.citation.fl\_str\_mv), poucos registos apresentam tal informação (apenas 35,61%);
- “Title”, “format” e “instname\_str” apresentam a porcentagem de não preenchimento baixa;
- “Description” metadado para o resumo com 31,24% de não preenchimento;



# Resultados

---

- A seleção de **amostras** para esta análise foi realizada por meio de uma **busca estratificada**, assegurando **representatividade** das diversas **tipologias documentais** hospedadas no Oasisbr e possibilitando identificar **problemas** nos metadados utilizados.
- A interoperabilidade promove o **acesso simultâneo aos dados** presentes nos variados tipos de repositórios digitais, reduzindo o **tempo de resposta**, juntamente com a **ampliação e potencialização das buscas**. Para melhorar a interoperabilidade, torna-se necessário a **aplicação de padrões**, diretrizes e modelos comuns, como OAI-PMH, muito utilizado para o **compartilhamento dos metadados**.

# Considerações finais

---

- O portal Oasisbr disponibiliza uma **ampla variedade de fontes e produções científicas**, resultando em problemáticas para **interoperabilidade e usabilidade** do ambiente, a partir da **heterogeneidade dos metadados** das diversas fontes coletadas pelo portal. Uma **lacuna identificada** na pesquisa, que está relacionada a organização e acessibilidade dos acervos, é a **disparidade** nos padrões de metadados adotados e a **inconsistência** na sua aplicação.
- Ressaltamos a necessidade de uma **abordagem coordenada** para a gestão de metadados entre **plataformas agregadoras** de conteúdo científico, com o intuito de melhorar a interoperabilidade, a **eficiência na busca e recuperação de informações**, e por fim, a **visibilidade da produção científica**. Adotar práticas de **metadados consistentes** e baseadas em **padrões reconhecidos** internacionalmente pode, significativamente, contribuir para o **avanço da Ciência Aberta**, trazendo benefícios para pesquisadores, instituições e para a sociedade como um todo.

# Referências

---

APACHE. Apache Solr Reference Guide. Delaware, EUA: Solr.apache.org, 2022. Disponível em: <https://solr.apache.org/guide/latest/index.html>. Acesso em: 15 jul. 2024.

FORMENTON, Danilo. Identificação de padrões de metadados para preservação digital. 2015. 102 p. Dissertação (Mestrado em Ciência, Tecnologia e Sociedade) - Universidade Federal de São Carlos, São Paulo, São Carlos, 2015. Disponível em: <https://repositorio.ufscar.br/handle/ufscar/7221?show=full#:~:text=Como%20resultado%20apresenta-se%20um%20quadro%20referencial%20te%C3%B3rico%20t%C3%A9cnico,relevantes%20para%20a%20modelagem%20de%20reposit%C3%B3rios%20institucionais%20digitais>. Acesso em: 15 ago. 2024.

GRÁCIO, José Carlos Abbud. Preservação digital na gestão da informação: um modelo processual para as instituições de ensino superior. São Paulo: Cultura Acadêmica, 2012. Disponível em: <https://repositorio.unesp.br/items/5d654eb0-a979-49f8-ac7d-a0a57febcd0c/full>. Acesso em: 15 ago. 2024.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA - IBICT. Gerenciar metadados. Ibict, [s. l.], 18 set. 2013. Disponível em: [https://wiki.ibict.br/index.php/Gerenciar\\_metadados](https://wiki.ibict.br/index.php/Gerenciar_metadados). Acesso em: 12 jul. 2024.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA - IBICT. Sobre OASISBR. Brasília: Ibict. Disponível em: <https://oasisbr.ibict.br/vufind/about/home>. Acesso em: 15 jul. 2024.

MARTINS, D. L.; DEMACHKI, É.; OLIVEIRA, L. F. R. Agregador de Repositórios Científicos em Artes - coleta de dados e interoperabilidade entre repositórios. Mouseion: Revista do Museu e Arquivo Histórico La Salle, [s. l.], n. 43, 2022. Disponível em: <https://www.revistas.unilasalle.edu.br/index.php/Mouseion/article/view/11090>. Acesso em: 12 jul. 2024.

MARQUES, G. S.; ESTEVAM, B. S.; LIMA, M. C. M.; MORAES, C. T.; SOUZA, M. G. Investigação da padronização de metadados adotados no Oasisbr. 2024. DOI: <https://doi.org/10.48472/aleia/THFG9S>. Disponível em: <https://aleia.ibict.br/dataset.xhtml?persistentId=doi:10.48472/aleia/THFG9S>. Acesso em: 15 jul. 2024.

OLIVEIRA, R. R. de; CARVALHO, C. L. de. Implementação de interoperabilidade entre repositórios digitais por meio do protocolo OAI-PMH. 2009. 56 f. Dissertação (Mestrado em Ciência da Computação) – Instituto de Informática, Universidade Federal de Goiás, Goiânia, 2009. Disponível em: [https://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF\\_003-09.pdf](https://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_003-09.pdf). Acesso em: 12 jul. 2024.

OPENDOAR. The Directory of Open Access Repositories. 2024. Disponível em: <http://www.opendoar.org/>. Acesso em: 12 jul. 2024.

PAGANINE, L. N.; CARVALHO SEGUNDO, W. L. R.; MOREIRA, J. L. R.; SAYÃO, L. F.; BARRETO NETO, V. C.; CIUFFO, L. N.; FELICÍSSIMO, C. H.; DIAS, G. N. Padrões de interoperabilidade para repositórios de dados de pesquisa. Brasília, DF: Ibict, 2023. Disponível em: <https://livroaberto.ibict.br/bitstream/123456789/1085/2/Padrões%20de%20interoperabilidade%20para%20repositórios%20de%20dados%20de%20pesquisa%20OGP.pdf>. Acesso em: 14 jul. 2024.

## Título: Padronização de metadados adotados no portal agregador de conteúdo científico Oasisbr

---

Agradeço a  
atenção!

Contatos:

[gabrielmarques@ibict.br](mailto:gabrielmarques@ibict.br)  
[blenasantos@ibict.br](mailto:blenasantos@ibict.br)  
[millenalima@ibict.br](mailto:millenalima@ibict.br)  
[cassiomorais@ibict.br](mailto:cassiomorais@ibict.br)  
[marcelsouza@ibict.br](mailto:marcelsouza@ibict.br)



This work is licensed under a  
[Creative Commons Attribution  
4.0 International License](https://creativecommons.org/licenses/by/4.0/).



MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÃO



Porto, 01 de outubro de 2024.