10° EBBC
Encontro Brasileiro de Bibliometria e Cientometria

UFPR
UNIVERSIDADE FEDERAL DO PARANÁ

LEVANTAMENTO DE INSTITUIÇÕES COM REPOSITÓRIOS E BIBLIOTECAS DIGITAIS PRESENTES NO DEPOSITA E VERIFICAÇÃO DA DUPLICIDADE DE REGISTROS NOS PORTAIS AGREGADORES BDTD E OASISBR

Millena Cordeiro Matos de Lima

https://orcid.org/0009-0008-6033-8018
millenalima@ibict.br
Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)

Blena Estevam dos Santos

https://orcid.org/0009-0006-7329-9236
blenasantos@ibict.br
Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)

Gabriel Silveira Marques

https://orcid.org/0000-0001-5886-6294
gabrielmarques@ibict.br
Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)

Resumo:
O estudo objetiva realizar o levantamento das instituições vinculadas aos registros incluídos no Deposita e que ao longo do tempo essas instituições criaram seus próprios repositórios, o que resultou em um cenário de duplicidade de registros nos portais agregadores BDTD e Oasisbr. Verificou-se 131 instituições com duplicidade de registros nas bases mencionadas, consequentemente torna-se um desafio para a curadoria, qualidade dos dados e recuperação da informação. Conclui-se que a duplicidade compromete a organização e recuperação da informação, evidenciando a necessidade de curadoria, bem como ampla divulgação dos portais visando integrar instituições que possuem seus repositórios, mas não se integraram.

Palavras-Chave: Deposita. Repositórios digitais. Bibliotecas digitais. Portais agregadores. Ciência Aberta.

EIXO TEMÁTICO:
Bases e Fontes de Dados

MODALIDADE:
Resumo expandido

DOI: 10.22477/x.ebbc.997

COMO CITAR ESTE ARTIGO:

LIMA, Millena Cordeiro Matos de; SANTOS, Blena Estevam dos; MARQUES, Gabriel Silveira. Levantamento de instituições com repositórios e bibliotecas digitais presentes no Deposita e verificação da duplicidade de registros nos portais agregadores BDTD e Oasisbr. *In*: EBBC - Encontro Brasileiro de Bibliometria e Cientometria. 10. **Anais [...]**. Curitiba, 2026. DOI: 10.22477/x.ebbc.997. Disponível em: https://doi.org/10.22477/x.ebbc.997.

CC BY

1 INTRODUÇÃO

A partir do surgimento do Movimento de Acesso Aberto (MAA), observa-se o crescimento no debate sobre recomendações voltadas à criação de repositórios digitais e o incentivo ao depósito das pesquisas científicas em Acesso Aberto (AA). Dentro desse contexto, fortalecido pela Ciência Aberta (CA) na atualidade, destaca-se o papel dos Repositórios Institucionais (RI) e os seus impactos no aumento da visibilidade, do uso e dos resultados das pesquisas disponibilizadas em acesso livre (Leite, 2009). Entre as instituições que apoiam o MAA e a CA, o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), apresenta-se como pioneiro, tornando-se referência em ações e projetos voltados à criação e manutenção de infraestruturas para consolidação e fortalecimento do MAA e da CA no cenário brasileiro (Ibict, [s.d.]).

Dentre as iniciativas do Ibict, destaca-se o Repositório Comum do Brasil, mais conhecido como Deposita, que foi lançado no ano de 2017 como uma base de dados centralizada em reunir e disponibilizar as produções científicas dos pesquisadores brasileiros ligados às instituições que ainda não possuem seus próprios repositórios ou que tenham trabalhos defendidos em instituições estrangeiras. O Deposita é um repositório digital que tem como objetivo ampliar o acesso e visibilidade das produções científicas dos pesquisadores brasileiros, como também incentivar a criação de RI nas instituições de ensino e pesquisa. Vale ressaltar que a inclusão dos documentos no Deposita é realizada mediante o autodepósito, ou seja, os próprios autores realizam o depósito do trabalho no repositório (Santos; Lima; Morais, 2025). Logo, o Ibict se apresenta como um meio de apoio e suporte para implementação de repositórios digitais em AA.

Neste sentido, o objetivo do trabalho é realizar um levantamento das instituições

que estão presentes no Deposita, mas que ao longo dos anos já implementaram os seus próprios RI ou bibliotecas digitais, o que contrapõe à sua função inicial: conter apenas trabalhos de instituições que não tenham seus próprios RI. A partir desse levantamento é possível identificar a ocorrência de duplicidade de registros e as questões que influenciam na qualidade, organização e recuperação da informação em portais agregadores nacionais, como a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) e o Portal Brasileiro de Publicações e Dados Científicos em Acesso Aberto (Oasisbr), ambos de iniciativa do Ibict. A problemática quanto às duplicidades surgiu a partir de buscas nos portais do Ibict que resultaram em títulos duplicados, levando a questionamentos sobre os impactos na recuperação informacional nas bases de dados.

No presente trabalho, assume-se que as duplicidades de registros são documentos com a mesma tipologia, por exemplo, ambos registros de teses com mesma autoria, data, fonte e título. Assim, a duplicidade é percebida como algo concreto e efetivo, visto que, os registros são exatamente iguais. Porém, em casos de tipologias diferentes e ambientes diferentes, por exemplo, trabalhos com o mesmo título e autoria, sendo uma dissertação presente no RI e o outro, um artigo disponível no portal de revista, não foram considerados na análise deste estudo como duplicidade.

A problemática da duplicidade de registros afeta todos os ambientes informacionais. Para Oliveira *et al.* (2010), a repetição de registros se apresenta como um problema para as redes cooperativas de bibliotecas, interferindo na qualidade das informações ofertadas, compreendendo os registros com duplicidades como aqueles que possuem o mesmo conteúdo (exatamente iguais) e que também possuem conteúdo aproximado. Tal perspectiva de aproximação de conteúdo não é seguida neste trabalho, pois, entre os dados analisados, observa-se a presença de

trabalhos com mesmo título, autoria e conteúdos similares, mas as tipologias são submetidas em fontes diferentes.

Perante isso, entende-se que a duplicidade "representa sobrecarga de informação para os usuários do catálogo, adicionando ruído e reduzindo o número de documentos relevantes no resultado de uma pesquisa" (Oliveira *et al.*, 2010, p. 10), o mesmo pode ser observado para RI, BDTD e principalmente para os portais agregadores que operam por cooperação de diferentes fontes, sistemas e grandes volumes informacionais.

Nesse contexto, a duplicidade dos registros em diferentes bases e portais agregadores configura-se como um desafio para a qualidade dos dados e a recuperação da informação. A partir disso, entende-se que a presente pesquisa está alinhada às discussões apresentadas e os eixos temáticos do Encontro Brasileiro de Bibliometria e Cientometria (EBBC).

2 PROCEDIMENTOS METODOLÓGICOS

Consiste em uma pesquisa de abordagem quantitativa e descritiva, com procedimentos de levantamento e análise documental, voltada à identificação de instituições com produções depositadas no Deposita e que já possuem RI ou bibliotecas digitais próprias, constatadas a partir de verificação manual na BDTD, Oasisbr e fontes digitais das instituições. A coleta de dados foi realizada no período de 16 de abril de 2025 a 11 de fevereiro de 2026, diretamente no Deposita, de forma manual, por meio da busca facetada com o filtro de "instituições".

Vale destacar que os dados obtidos podem não representar os dados que constam atualmente no Deposita, pois há um processo de curadoria, de acordo com o manual de depósito de obras. Nesse manual, consta que os documentos já de-

positados no Deposita, caso não estejam mais alinhados com a política do repositório podem ser mantidos ou excluídos, a depender do critério da equipe gestora (Ibict, 2025a).

A busca realizada no Deposita resultou em uma lista com todas as instituições (321) vinculadas aos registros disponíveis na plataforma. Em seguida, procedeu-se à verificação da existência de RI ou bibliotecas digitais dessas instituições, a partir de consultas aos sites institucionais, aos portais agregadores e pesquisas no Google, resultando em 131 instituições que possuem RI ou biblioteca digital, as quais formam o núcleo de análise deste estudo. Os dados coletados foram organizados em uma planilha do Google, na qual foram sistematizadas as seguintes informações: o nome da instituição que tem RI ou biblioteca digital; se o RI ou biblioteca digital está sendo coletado pela BDTD ou Oasisbr; a quantidade de itens depositados de cada instituição; o link da biblioteca digital ou RI; e algumas observações, por exemplo: se o repositório está disponível, se todos os registros estão no repositório da instituição.

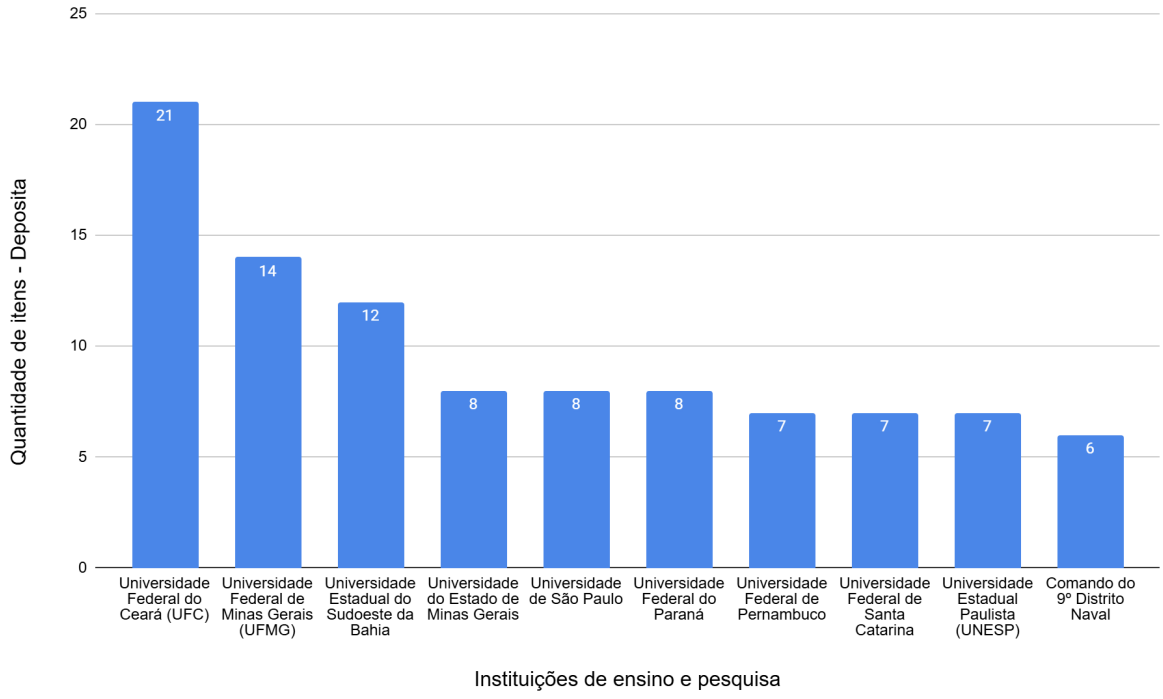
A partir dessas informações, é possível identificar a ocorrência de duplicidade de registros entre as bases analisadas, por meio de verificação cruzada nos portais do Ibict, com conferência de títulos presentes no Deposita e validação nos RI das instituições. Os resultados foram apresentados por meio de gráficos, possibilitando a visualização das instituições com maior incidência de registros no Deposita, que já possuem RI ou bibliotecas digitais de forma a promover a discussão dos impactos dessa duplicação para a organização, qualidade e recuperação da informação em bases de dados científicas nacionais.

3 DISCUSSÃO E RESULTADOS

Após o levantamento dos dados torna-se possível averiguar quais as instituições

que já possuem seus próprios repositórios e, simultaneamente, registros no Deposita, como também o quantitativo de registros vinculados às instituições e a participação nos portais agregadores BDTD e/ou Oasisbr, além de outras observações verificadas durante o levantamento. Dessa forma, no Gráfico 1 são apresentadas apenas as 10 instituições identificadas com os maiores quantitativos de registros. Esse corte torna-se necessário, pois entre 131[1] instituições e 369 registros encontrados, aproximadamente 84 instituições possuem apenas 1 ou 2 registros, contudo 3 instituições possuem mais de 10 registros cada uma.

Gráfico 1 - As 10 instituições com mais trabalhos disponíveis no Deposita e que possuem RI ou biblioteca digital



Fonte: dados da pesquisa (2026).

Observa-se no Gráfico 1, o destaque da Universidade Federal do Ceará (UFC) com o maior quantitativo analisado, 21 registros, seguida pela Universidade Federal de Minas Gerais (UFMG) com 14 registros. Em ambos os casos, verificou-se que parte

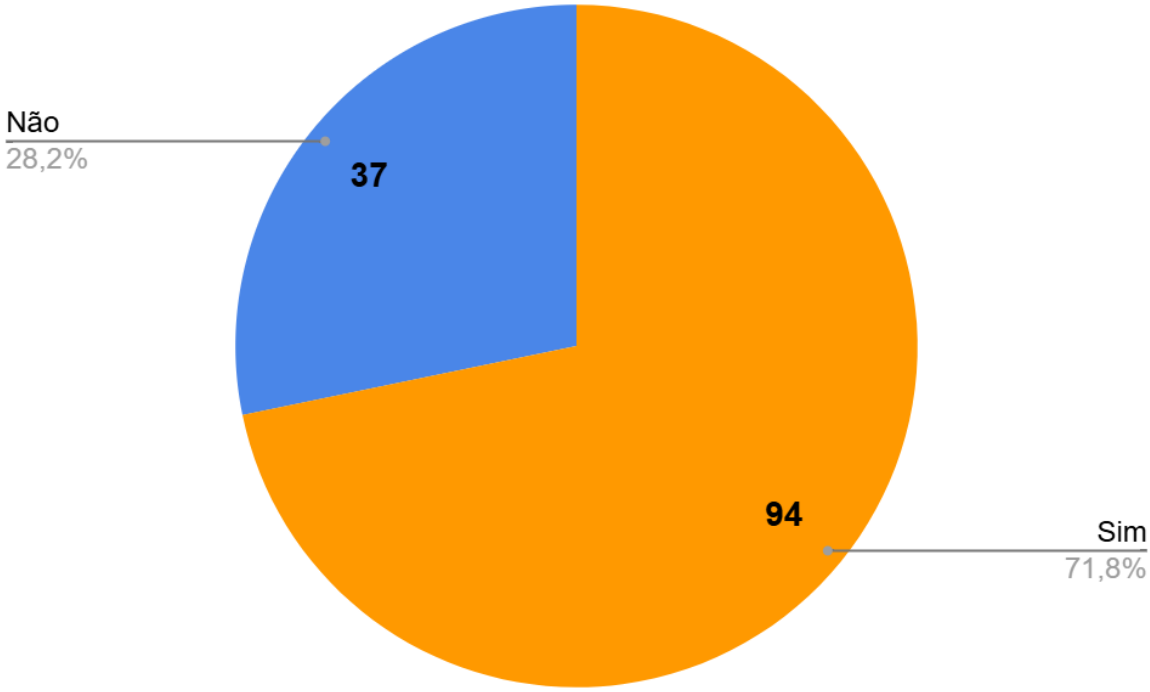
[1] No conjunto de dados, que será disponibilizado em repositório de dados posteriormente, a Universidade de São Paulo aparece duas vezes, a primeira com 8 itens e a segunda ocorrência com 2 itens, pois a primeira ocorrência corresponde ao RI e à biblioteca digital de teses e dissertações que são coletados pela BDTD e Oasisbr; e a segunda ocorrência corresponde à biblioteca digital de monografias da instituição, a qual não é coletada pelo Oasisbr nem pela BDTD, por esse motivo, são 131 instituições no total.

dos registros presentes no Deposita não foram localizados nos respectivos RI.

Além disso, nota-se que as demais instituições apresentam variações graduais nos quantitativos, sem discrepâncias significativas como observado a partir da Universidade Estadual do Sudoeste da Bahia (UESB) (12 registros), seguida da Universidade de São Paulo (USP) (8 registros) e a Universidade Federal de Pernambuco (UFPE) (7 registros). Entende-se que esse cenário pode não estar relacionado ao período de criação dos RI ou bibliotecas das instituições, pois, por exemplo, o RI da UFC foi implementado por volta de 2011 e o da USP foi inaugurado no ano de 2012 (UFC, 2025; USP, [202-]). E tais instituições apresentam um grande distanciamento quanto aos registros presentes no Deposita.

Diante disso, pode-se inferir que no caso da USP o RI possa ser mais disseminado entre a sua comunidade acadêmica e o depósito seja mais incentivado, diferente da realidade de outras instituições. Pois é necessário considerar o contexto de cada instituição, incluindo sua política de depósito, aspectos culturais, níveis de conscientização sobre o AA, entre outros fatores, para compreender melhor o impacto da duplicidade dos registros.

Gráfico 2 - Instituições com trabalhos depositados no Deposita e presença em agregadores nacionais: **BDTD e Oasisbr**



Fonte: dados da pesquisa (2026).

Diante do gráfico 2, percebe-se que grande parte das instituições presentes no Deposita e que possuem seus próprios RI ou bibliotecas digitais, são coletadas tanto pela BDTD quanto pelo Oasisbr (71,8%). Isso pode ocasionar duplicação nos registros desses portais, pois, os mesmos coletam o Deposita e podem coletar os RI e bibliotecas digitais próprias das instituições. A duplicidade de registros pode interferir diretamente na recuperação e qualidade das informações de tais bases de dados, dificultando a limpeza de dados e a produção de novas pesquisas científicas. Constatou-se também que alguns itens dessas instituições não estão no RI ou biblioteca digital, a exemplo, a UFC com artigos, livros e capítulos de livros e TCC que não constam no RI e estão presentes no Deposita. Isso pode ocorrer por diversos motivos, incluindo a política da instituição não aceitar o depósito das tipologias textuais mencionadas, ou a inclusão ser realizada por meio do autodepósito e o autor deixar de realizar, entre outros fatores relacionados também à cultura acadêmica de cada instituição.

No caso da UFC, de acordo com a sua política, o depósito é realizado pela comunidade acadêmica (autodepósito) e exige obrigatoriamente o depósito de TCC de graduação e especialização, teses e dissertações, já os artigos, livros e capítulos de livros são permitidos, mas não é obrigatório o depósito (UFC, 2023). A partir disso, entende-se que os documentos desta instituição presentes no Deposita não são obrigatórios no RI. Logo, é possível inferir que tais documentos não são vistos como prioridade para inclusão no RI e o Deposita se apresenta como uma via mais rápida para disponibilização e divulgação desses trabalhos.

Quanto às instituições que não estão presentes nos portais agregadores (28,2%), considera-se de suma importância que tais instituições se integrem à BDTD e ao Oasisbr, visto o papel estratégico desses portais na promoção da CA, na visibilidade da produção científica e no fortalecimento da interoperabilidade entre sistemas de informação científica (Ibict, 2025b). Entretanto, a ausência dessas instituições pode estar ligada às dificuldades técnicas e estruturais, como a inexistência ou configuração inadequada no protocolo de interoperabilidade, essencial na troca de informações, no caso dos portais do ibict sugere-se o *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH). Ademais, a falta de padrões de metadados exigidos pelos portais, além de limitações institucionais relacionadas à gestão, manutenção e curadoria dos repositórios (Marques *et al*., 2025).

Cabe destacar, exemplificando a questão de duplicidade percebida nos portais do ibict, o caso da Pontifícia Universidade Católica do Paraná (PUC-PR), onde uma dissertação de mesma autoria, ano e título que está presente no Deposita, foi identificada repetida na BDTD e no Oasisbr, pois tais portais também coletam a biblioteca digital de teses e dissertações da instituição, que já possui o registro da dissertação percebida, podendo ocasionar dúvidas aos usuários em relação ao acesso que pode ser inválido ou equivocado, gerando uma sensação de desor-

dem e atrapalhando a curadoria nas bases de dados.

Esse cenário de duplicidades promovido pelas instituições que são coletadas pelo Oasisbr e BDTD e que ao mesmo tempo possuem registros no Deposita, é uma questão importante na realidade dos portais do Ibict, já que, utilizam o mesmo sistema de coleta. Porém, nos portais do Ibict também é realizada uma curadoria após os depósitos e as coletas, para atualização das informações de modo a garantir a sua qualidade e a recuperação. Tal cenário de registros duplicados também é observado na plataforma Sucupira da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) por inconsistência de metadados, apesar da plataforma orientar sobre a verificação de duplicidades, demandado pesquisas futuras mais aprofundadas de investigação na plataforma.

4 CONSIDERAÇÕES FINAIS

Os resultados evidenciam um cenário de duplicidade de parte significativa dos registros analisados, caracterizada pela repetição de documentos com metadados equivalentes em múltiplas bases (RI, portais agregadores e Deposita). Tal ocorrência, pode comprometer a integração, padronização, organização, qualidade e recuperação da informação, o que evidencia a necessidade de capacitação, apoio técnico e ampla divulgação dos portais do Ibict em workshops e eventos científicos com o intuito de ampliar a participação das instituições em portais agregadores nacionais e alinhamento com os princípios da CA em prol de uma ciência mais colaborativa no país.

Diante disso, entende-se que a duplicidade de registros em portais e bases agregadoras assim como em RI e bibliotecas digitais, dificulta o processo de busca e recuperação da informação, demandando mais esforço para organização e curadoria nos ambientes informacionais. Consequentemente limita e cria barreiras no

processo da pesquisa científica, ao se promover ambientes poluídos pelo excesso de informações, afetando a qualidade da informação disponibilizada, com muita informação irrelevante, e afeta o uso efetivo dos ambientes informacionais, que deveriam minimizar a sobrecarga informacional. Tal cenário não favorece a otimização da busca nos portais, tornando-os em ambientes não atrativos para a comunidade científica e acadêmica, sendo necessário avaliar os registros para que possam ser excluídos ou mantidos do Deposita.

REFERÊNCIAS

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Manual de depósito de obras no Deposita - Repositório Comum do Brasil**. Brasília, DF: IBICT, 2025a. Disponível em: https://deposita.ibict.br/assets/docs/Deposita_Publica%C3%A7%C3%B5es_manual_de_dep%C3%B3sito_de_obras.pdf. Acesso em: 23 fev. 2026.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Missão**. Brasília: IBICT, [s.d.]. Disponível em: https://www.gov.br/ibict/pt-br/acesso-a-informacao/institucional/missao-1. Acesso em: 26 fev. 2026.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Sobre**. Brasília, DF: Ibict, 2025b. Disponível em: https://oasisbr.ibict.br/vufind/about/home. Acesso em: 25 fev. 2026.

LEITE, Fernando César Lima. **Como gerenciar e ampliar a visibilidade da informação científica brasileira**: repositórios institucionais de acesso aberto. Brasília: Ibict, 2009. 120 p. Disponível em: https://livroaberto.ibict.br/handle/123456789/775. Acesso em: 10 fev. 2026.

MARQUES, Gabriel Silveira; SANTOS, Blena Estevam dos; LIMA, Millena Cordeiro Matos de; MORAIS, Cássio Teixeira de; SOUZA, Marcel Garcia de. Padronização de metadados adotados no portal agregador de conteúdo científico Oasisbr. **Ciência da Informação**, Brasília, DF, v. 54, n. 2, p. 1-11, 2025. Disponível em: https://revista.ibict.br/ciinf/article/view/7223/7344. Acesso em: 25 fev. 2026.

OLIVEIRA, Zita Prates de; FREITAS, Lais Borges; PAVÃO, Caterina Groposo; COSTA, Janise Silva Borges da; SCHMIDT, Margarida Cecília; HOROWITZ, Zaida; SAATKAMP, Carla Metzler; NOTARIO, Bruce Lucien Santos. Gerência de registros duplos em base de dados bibliográfica. *In*: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 15.; SEMINÁRIO INTERNACIONAL DE BIBLIOTECAS DIGITAIS, 2., 2010, Rio de Janeiro. **Anais** [...]. Rio de Janeiro: UFRJ, 2010. Disponível em: http://

repositorio.febab.org.br/items/show/5219. Acesso em: 29 abr. 2026.

SANTOS, Blena Estevam dos; LIMA, Millena Cordeiro Matos de; MORAIS, Cássio Teixeira de. Repositório Comum do Brasil (Deposita). *In*: AMARO, Bianca; CAMPOS, Fhillipe de Freitas; BARCELOS, Janinne (org.). **Infraestruturas de Ciência e de Acesso Aberto no Brasil**: iniciativas do Instituto Brasileiro de Informação em Ciência e Tecnologia. Brasília, DF: Ibict, 2025. Cap. 5, p. 107-118. DOI: 10.22477/9788570132543.cap5. Disponível em: https://omp-editora.prd.ibict.br/index.php/edibict/catalog/view/358/506/2656. Acesso em: 13 fev. 2026.

UNIVERSIDADE FEDERAL DO CEARÁ (UFC). Repositório Institucional. **Biblioteca UFC**. Fortaleza, 2025. Disponível em: https://biblioteca.ufc.br/pt/servicos-e-produtos/repositorio-institucional/. Acesso em: 25 fev. 2026.

UNIVERSIDADE FEDERAL DO CEARÁ (UFC). **Resolução nº 05, de 09 de fevereiro de 2023**. Dispõe sobre a Política de Informação Técnico-Científica do Repositório Institucional da Universidade Federal do Ceará. Fortaleza: UFC, 9 fev. 2023. Disponível em: https://repositorio.ufc.br/documentos/UFC_CONSUNI_Resolucao_05_2023_pdfa.pdf. Acesso em: 25 fev. 2026.

UNIVERSIDADE DE SÃO PAULO (USP). Sobre o Repositório. **Repositório da Produção USP**. São Paulo, [202-]. Disponível em: https://repositorio.usp.br/sobre.php. Acesso em: 25 fev. 2026.