Análise focada no padrão MTD3-BR sob a luz dos Princípios FAIR

Keli Rodrigues do Amaral Benin[a]

1. Introdução

"O movimento de Acesso Aberto tem como principal objetivo fomentar a leitura e obtenção de conhecimento para todos. Além disso, busca estimular o diálogo entre cientistas de maneira mais rápida e eficaz, uma vez que todos terão acesso àquela publicação" (PERIODICOS DE MINAS, 2021). Para isso o movimento de Acesso Aberto à informação científica criou duas estratégias: a Via Verde (autoarquivamento em repositórios eletrônicos, repositórios institucionais) e a Via Dourada (periódicos científicos de acesso livre).

Nessa mesma linha de movimento de Acesso Aberto, a Biblioteca Digital de Teses e Dissertações (BDTD), do Instituto Brasileiro de Informação em Ciência, Tecnologia e Inovação (IBICT), foi o primeiro sistema a utilizar os conceitos de Arquivos Abertos.

A BDTD/IBICT tem como objetivo de "integra e dissemina, em um só portal de busca, os textos completos das teses e dissertações defendidas nas instituições brasileiras de ensino e pesquisa" (BDTD, 2021).

a Especialista Informação, Conhecimentos e Sociedade pela Universidade Estadual de Londrina – UEL. E-mail: keliamaral@hotmail.com. ORCID: https://orcid.org/0000-0002-2564-2531.

No processo de criação da BDTD/IBICT vislumbrou-se a criação de um padrão de metadados que é fundamental para que se garanta a interoperabilidade entre os sistemas de informação, o Padrão Brasileiro de Metadados para Teses e Dissertações (MTD-BR).

Os padrões de metadados apresentam-se como importante ferramenta na organização dos dados em contextos específicos, conforme defendem Nogueira, Vaz e Souza (2010), a utilização de metadados para descrição de objetos contribui para minimizar as dificuldades inerentes às atividades de localização, descoberta, documentação, seleção e recuperação de informação.

Ainda dentro do movimento de Acesso Aberto, vem de encontro os Princípios FAIR, conjuntos preciso e mensurável de qualidades que uma boa publicação de dados deve apresenta. Com a adoção dos Princípios FAIR nos dados, metadados e infraestruturas traz os benefícios como:

- Ajudar os colegas e o próprio futuro a compreender o projeto de pesquisa e os dados;
- Facilitar o compartilhamento de dados e colaborações;
- Aumentar a visibilidade da pesquisa e pode levar a mais citações;
- Melhorar a transparência, confiabilidade e reprodutibilidade da pesquisa;
- Prevenir a perda de dados;
- Maximizar o potencial dos ativos de dados; e
- Maximizar o impacto da pesquisa.

O foco principal da recomendação para os Princípios FAIR é para ambientes de gestão de dados de pesquisa, contudo, considerando que o FAIR se aplica a dados, metadados e infraestruturas, pode-se recomendadas essas boas práticas nas Bibliotecas Digitais de Teses e Dissertações e nos Repositórios Institucionais, que envolvem dados, metadados e infraestruturas.

Moraes e Oliveira (2010) apresentam a Biblioteca Digital Brasileira de Teses e Dissertações como uma das principais fontes de consulta para teses e dissertações brasileiras. Nesse sentido surgiu a seguinte questão, a Biblioteca Digital de Teses e Dissertações (BDTD/IBICT), tem adotado os Princípios FAIR no MTD3-BR? O presente artigo teve como objetivo principal analisar quais metadados do MTD3-BR, já estão adequados nos Princípios FAIR.

2. Metadados e Padrões de Metadados

Os metadados são essenciais para a gestão e a recuperação da informação, por serem elementos usados para identificar, descrever e representar o conteúdo de um recurso informacional, além de descrever a localização do recurso e viabilizar sua busca, acesso e recuperação na Web.

Os autores Souza, Vendrúsculo e Melo (2000, p. 93) definem metadados como "[…] dado sobre dado. É a catalogação do dado ou descrição do recurso eletrônico". Da mesma forma, Takahashi (2000, p. 172) conceitua metadados como: "[…] dados a respeito de outros dados, ou seja, qualquer dado usado para auxiliar na identificação, descrição e localização de informações. Trata-se,

em outras palavras, de dados estruturados que descrevem as características de um recurso de informação". Neste sentido, metadado refere-se a uma estrutura descritiva da informação sobre outro dado.

Em outras palavras, metadados é um "Conjunto de elementos que descrevem as informações contidas em um recurso, com o objetivo de possibilitar sua busca e recuperação" (GRÁCIO, 2002, p. 23).

"A padronização dos metadados facilita o acesso e a recuperação da informação, assegurando aos usuários a fácil movimentação de dados entre diferentes sistemas e plataformas computacionais" (ALVES e CAFÉ, 2010, p. 185).

Alves (2005, p. 115) esclarece que:

> […] formatos de metadados, também chamados de padrões de metadados, são estruturas padronizadas para a representação do conteúdo informacional que será representado pelo conjunto de dados-atributos (metadados). Em outras palavras, os formatos ou padrões de metadados podem ser considerados como formas de representação de um item documentário (ALVES, 2005, p.115).

"A implantação da Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), que teve as primeiras discussões sobre sua criação a partir de 2001 e seu lançamento oficial em dezembro de 2002, se deu em nível nacional, através do IBICT, com base no *Open Archives*" (MORAES; OLIVEIRA, 2010, p. 74).

No processo de criação da BDTD, o IBICT elaborou e implantou o Padrão Brasileiro de Metadados para Teses e Dissertações (MTD-BR). Desde sua criação, o MTD-BR teve três versões, sendo que a primeira versão lançada em 2002 continha 71 metadados, sendo 38 obrigatórios e 33 opcionais. A segunda versão (MTD2-BR) foi desenvolvida em 2005 e possui oitenta metadados sendo 31 obrigatórios e 49 opcionais. A terceira versão (MTD3-BR) lançada em 2014 tem, portanto, ao todo 38 metadados, sendo 18 metadados obrigatórios e 20 metadados opcionais. A nova proposta do Padrão Brasileiro de Metadados da BDTD foi desenvolvido com base no esquema de metadados do Dublin Core com adaptações para a realidade brasileira.

3. Princípios FAIR

Os princípios FAIR vêm com o movimento atual em direção a e-Science e Ciência Aberta, envolvendo principalmente a necessidade do gerenciamento de dados de pesquisa. O FAIR é um conjunto de princípios orientadores para o bom gerenciamento de dados, com o objetivo de tornar os dados de pesquisa Encontráveis, Acessíveis, Interoperáveis e Reutilizáveis (Quadro 1) (WILKINSON et. al, 2016).

Os princípios se referem a três tipos de entidades: (i) Dados (ou qualquer objeto digital produzido em pesquisa: dados de pesquisa no sentido mais estrito, código, software, apresentações, etc.); (ii) Metadados (informações sobre esse objeto digital); e (iii) Infraestrutura.

Quadro 1 – Diretrizes dos princípios FAIR

Princípio	Especificações
Encontrável (*findable*)	F.1 (meta) dados devem ter identificadores globais, únicos e persistentes.
	F.2 dados devem ser descritos utilizando metadados ricos (impacta diretamente R1).
	F.3 metadados devem incluir clara e explicitamente os identificadores dos dados que descrevem.
	F.4 (meta) dados devem ser registrados ou indexados em mecanismos de busca
Acessível (*acessible*)	A.1 (Meta) dados devem ser recuperáveis pelos seus identificadores usando protocolo de comunicação padronizado.
	A.1.1 O protocolo deve ser aberto, gratuito e universalmente implementável.
	A.1.2 O protocolo deve permitir procedimentos de autenticação e autorização, quando necessário.
	A.2 Metadados devem ser acessíveis, mesmo quando os dados não estiverem mais disponíveis.
Interoperáveis (*interoperable*)	I.1 (Meta) dados devem ser representados por meio de uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento.
	I.2 (Meta) dados devem usar vocabulários de acordo com os princípios FAIR.
	I.3 (Meta) dados devem incluir referências qualificadas para outros (Meta) dados.
Reusáveis (*reusable*)	R.1 (Meta) dados são descritos com uma pluralidade de atributos precisos e relevantes.
	R.1.1 (Meta) dados devem ser disponibilizados com licenças de uso claras e acessíveis.
	R.1.2 (Meta) dados devem estar associados à sua proveniência.
	R.1.3 (Meta) dados devem estar alinhados com padrões relevantes do seu domínio.

Fonte: Go FAIR apud Wilkinson et al. (2016, tradução nossa).

Os Princípios não são necessariamente sobre dados abertos. É possível trabalhar de forma FAIR com dados que não se

destinam à disponibilidade pública. Também não são regras ou normas. A adoção dos Princípios FAIR será frequentemente uma adaptação gradual das rotinas de trabalho, mas também pode ser um enorme salto, onde um tipo de infraestrutura é substituído por outro. Caberá às diferentes áreas de pesquisa e comunidades de pesquisa fazer com que os Princípios do FAIR funcionem em seus respectivos contextos.

Além disso, a modularidade dos Princípios e sua distinção entre dados e metadados suportam explicitamente uma ampla gama de circunstâncias especiais. Um exemplo envolve a publicação de objetos de pesquisa sem dados. Fluxos de trabalho analíticos, por exemplo, são um componente crítico do ecossistema acadêmico e sua publicação formal é necessária para alcançar transparência e reprodutibilidade científica. Os Princípios FAIR podem igualmente ser aplicados a esses ativos não relacionados a dados, que precisam ser identificados, descritos, descobertos e reutilizados da mesma maneira que os dados (WILKINSON et al, 2016, tradução nossa)

Quando implementado, será uma gestão e administração mais rigorosas desses valiosos recursos digitais, para o benefício de toda a comunidade acadêmica.

4. Procedimentos metodológicos

A pesquisa se define como exploratória e descritiva, a fim de realizar estudos sobre o Padrão Brasileiro de Metadados para Teses e Dissertações e os Princípios FAIR.

Gil (2008) cita algumas fontes de interesse para a realização de pesquisas científicas, como publicações em obras de referência, teses e dissertações, periódicos científicos e anais de encontros. Para exploração dessas fontes, o presente estudo utilizou-se as seguintes bases de dados eletrônicas: Brapci, Repositório Institucional UNESP, Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), Periódicos Capes e Scielo.

5. Resultados e discussões

Segundo GO FAIR (2021), os Princípios FAIR podem ser aplicados nos metadados, dados e infraestrutura de suporte, a "maioria dos requisitos de encontrável e acessível podem ser alcançados no nível de metadados. Mas a interoperabilidade e a reutilização exigem mais esforços no nível dos dados". Então, o presente artigo focou analisar apenas os requisitos dos Princípios FAIR 'Encontrável' e 'Acessível'.

O Quadro 2 mostra todos o Padrão MTD3-BR e seus respectivos atributos, foi realizada a análise de cada metadado, destacando os quais estão adequados nos Princípios FAIR dos requisitos Encontráveis e Acessíveis.

Quadro 2 – Padrão Brasileiro de Metadados da Biblioteca Digital Brasileira de Teses e Dissertações (MTD3-BR) e Princípios FAIR

Metadado	Valor	Princípios FAIR
dc.creator	Autor	F.2. Metadado enriquecido
dc.creator.ID	ID do autor	F.2 Metadado enriquecido F.3. ID em metadados

Metadado	Valor	Princípios FAIR
dc.creator.Lattes	Currículo Lattes do autor	F.2. Metadado enriquecido
dc.contributor.advisor1	Orientador1	F.2. Metadado enriquecido
dc.contributor.advisor1ID	ID do orientador1	F.2. Metadado enriquecido F.3. ID em metadados
dc.contributor.advisor1Lattes	ID Lattes do orientador1	F.2. Metadado enriquecido F.3. ID em metadados.
dc.contributor.advisor2	Orientador2	F.2. Metadado enriquecido
dc.contributor.advisor2ID	ID do orientador2	F.2. Metadado enriquecido F.3. ID em metadados
dc.contributor.advisor2Lattes	ID Lattes do orientador2	F.2 Metadado enriquecido F.3. ID em metadados.
dc.contributor.advisor-co1	Co-orientador1	F.2. Metadado enriquecido
dc.contributor.advisor-co1ID	ID do co-orientador1	F.2. Metadado enriquecido F.3. ID em metadados
dc.contributor.advisor-co1Lattes	ID Lattes do co-orientador1	F.2 Metadado enriquecido F.3. ID em metadados.
dc.contributor.advisor-co2	Co-orientador2	F.2. Metadado enriquecido
dc.contributor.advisor-co2ID	ID do co-orientador2	F.2. Metadado enriquecido F.3. ID em metadados
dc.contributor.advisor-co2Lattes	ID Lattes do co-orientador2	F.2 Metadado enriquecido F.3. ID em metadados.
dc.contributor.referee1	Membro da banca1	F.2. Metadado

Metadado	Valor	Princípios FAIR
		enriquecido
dc.contributor.referee1ID	ID do membro da banca1	F.2. Metadado enriquecido F.3. ID em metadados
dc.contributor.referee1Lattes	Lattes do membro da banca1	F.2 Metadado enriquecido F.3. ID em metadados.
dc.contributor.referee2	Membro da banca2	F.2. Metadado enriquecido
dc.contributor.referee2ID	ID do membro da banca2	F.2. Metadado enriquecido F.3. ID em metadados
dc.contributor.referee2Lattes	Lattes do membro da banca2	F.2 Metadado enriquecido F.3 ID em metadados.
dc.contributor.referee3	Membro da banca3	F.2. Metadado enriquecido
dc.contributor.referee3ID	ID do membro da banca3	F.2. Metadado enriquecido F.3. ID em metadados
dc.contributor.referee3Lattes	Lattes do membro da banca3	F.2 Metadado enriquecido F.3. ID em metadados.
dc.contributor.referee4	Membro da banca4	F.2. Metadado enriquecido
dc.contributor.referee4ID	ID do membro da banca4	F.2. Metadado enriquecido F.3. ID em metadados
dc.contributor.referee4Lattes	Lattes do membro da banca4	F.2 Metadado enriquecido F.3. ID em metadados.
dc.contributor.referee5	Membro da banca5	F.2. Metadado enriquecido
dc.contributor.referee5ID	ID do membro da banca5	F.2 Metadado enriquecido

Metadado	Valor	Princípios FAIR
		F.3. ID em metadados
dc.contributor.referee5Lattes	Lattes do membro da banca5	F.2 Metadado enriquecido F.3. ID em metadados.
dc.identifier.doi	Identificador permanente – DOI	F.1 Identificador permanente F.2 Metadado enriquecido A.1.Protocolo padronizados de comunicação. A.1.1 Protolocolo abertos e gratuitos.
dc.description.resumo	Resumo	F.2 Metadado enriquecido
dc.description.abstract	Abstract	F.2 Metadado enriquecido
dc.publisher	Nome da instituição por extenso	N/A
dc.publisher.country	País da instituição	N/A
dc.publisher.department	Nome da faculdade, instituto ou departamento por extenso	N/A
dc.publisher.program	Nome do Programa de Pós-Graduação por extenso	N/A
dc.publisher.initials	Sigla da instituição	N/A
dc.identifier.uri	Identificador persistente	F.1 Identificador permanente F.2 Metadado enriquecido A.1.Protocolo padronizados de comunicação. A.1.1 Protolocolo

Metadado	Valor	Princípios FAIR
		abertos e gratuitos.
dc.subject	Palavras-chave que descrevem o assunto do documento	F.2 Metadado enriquecido
dc.subject.cnpq	Área(s) do conhecimento do documento (Tabela CNPq)	N/A
dc.thumbnail.url	URL Arquivo (imagem)	F.2 Metadados enriquecidos A.1.Protocolo padronizados de comunicação. A.1.1 Protolocolo abertos e gratuitos.
dc.description.sponsorship	Agência de fomento	N/A
dc.type	Tipo do documento: tese ou dissertação	N/A
dc.title	O título e subtítulo da tese ou dissertação na língua do documento	F.2 Metadado enriquecido
dc.title.alternative	Título alternativo do documento e/ou traduzido em outro idioma	F.2 Metadado enriquecido
dc.rights	Direitos de acesso ao documento	A.1.2 Autentificação A.2. Metadados sempre disponíveis.
dc.date.issued	Data de defesa	N/A
dc.date.available	Data de embargo	N/A
dc.language	Idioma	N/A
dc.relation.references	Referências citadas na tese ou dissertação	F.2 Metadado enriquecido
dc.identifier.citation	Citação do documento	F.2 Metadado enriquecido
dc.format	Formato do documento	

Metadado	Valor	Princípios FAIR
dc.relation	Relação, quando for o caso, com outros materiais produzidos para a elaboração da tese/dissertação ou comentários do próprio autor sobre a tese/dissertação	N/A

Fonte: Autora adaptado de BDTD.

A partir da análise de todos os elementos do padrão MTD3-BR e seus respectivos atributos apuraram-se os requisitos Encontrável (F.1): Identificador único persistente tanto para o conjunto de dados quanto para os metadados, no padrão MTD3-BR é possível inserir o identificador persistente para o conjunto de dados, como também para alguns metadados, como exemplo: Identificador permanente – DOI, ID Lattes do autor, orientador, co-orientador e dos membros da banca e outros.

Já para os requisitos Encontrável (F.2): Os metadados ricos (metadados completos), tornando metadados de qualidade e simétricos, o padrão MTD3-BR traz vários campos que podem enriquecer o conjunto de dados, mas alguns metadados não alcança a sua completude. Os requisitos Encontrável (F.3): ID em metadados, alguns campos permitem que insira o identificador persistente da descrição dos metadados, como exemplo já citado no F.1.

Os requisitos Encontrável (F.4): Meta (dados) registrados ou indexados em mecanismos de busca, aqui fala sobre o componente da infraestrutura, no caso da MTD3-BR, são as

Bibliotecas Digitais de Teses e Dissertações e os Repositórios Instrucionais.

Analisando os requisitos Acessível (A.1): Protocolo padronizados de comunicação, esse requisito se concentra como os dados e metadados poderão ser recuperados por seus identificadores permanentes. Mostrando a sua importância para configurar todo o ambiente e o serviço de dados FAIR, esse requisito tem ligação com o requisito F.1. Os requisitos Acessível (A.1.1) é uma continuidade do (A.1): Protocolo abertos e gratuitos, no caso da Biblioteca Digital de Teses e Dissertações, o acesso sem custo e aberto, qualquer pessoa com um computador e uma conexão com a Internet pode acessar pelo menos os metadados. Os requisitos Acessível (A.1.2) também é uma continuidade do requisito (A.1): Autenticação, quando necessário, muitos casos podem haver restrições de acesso aos dados e/ou metadados um mecanismo de autenticação e autorização para o acesso deve ser liberado pelo protocolo de comunicação.

E finalizando, os requisitos Acessível o (A.2): Metadados sempre disponíveis, mesmo quando os dados não tiverem mais disponíveis ou nunca estiveram disponíveis. Esse requisito está relacionado com o requisito F.4, o componente da infraestrutura e o direito de acesso.

6. Considerações finais

A análise do padrão MTD3-BR, permitiu observar que a existência de vários metadados que já estão adequados aos

Princípios FAIR, favorecendo o Acesso Aberto, a recuperação da informação e o compartilhamento da informação.

Também foi observado que a maioria dos metadados descritivos no padrão MTD3-BR estão nos requisitos Encontráveis (F.2) que tem uma importante ligação com os requisitos Reusáveis (R.1), portanto, quanto mais rico e completo forem as informações fornecidas nos metadados, mais fácil ficará do dado ser encontrado e aumentar a reutilização e a citação.

Percebeu-se que alguns casos os requisitos dos Princípios FAIR são bastantes repetitivos, assim reforçando alguns itens importante para a construção do ambiente e serviço de dados FAIR.

O presente trabalho analisou apenas os requisitos Encontrável e Acessível, porém ainda faltam mais estudos aprofundado para aprimorar mais metadados que atendem as boas práticas de gestão de dados nas Bibliotecas Digitais de Teses e Dissertações e Repositórios Institucionais.

Referências

ALVES, R. C. V. **Web semântica:** uma análise focada no uso de metadados. 2005. Dissertação (Mestrado em Ciência da Informação) - Universidade Estadual Paulista, Marília, 2005. Disponível em: http://hdl.handle.net/11449/93690. Acesso em: 05 jun. 2021.

ALVES, J. C.; CAFÉ, L. M. A. Análise focada em metadados sob a luz do padrão MTD-BR. **Em Questão**, Porto Alegre, v. 16, n.

2, p. 179-202, jul./dez. 2010. Disponível em: https://seer.ufrgs.br/EmQuestao/article/view/12930. Acesso em: 08 jun. 2021.

BDTD – Biblioteca Digital Brasileira de Teses e Dissertações. **O que é a BDTD**. São Paulo, 2021. Disponível em: https://bdtd.ibict.br/vufind/Content/whatIs. Acesso em: 07 jun. 2021.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.

GO FAIR. **FAIR Principles**. 2021. Disponível em: https://www.go-fair.org/fair-principles/. Acesso em: 08 jun. 2021.

GRÁCIO, J. C. A. **Metadados para a descrição de recursos da Internet: o padrão Dublin Core, aplicações e a questão da interoperabilidade.** 2002. Dissertação (Mestrado em Ciência da Informação) - Universidade Estadual Paulista, Marília, 2002. Disponível em: http://hdl.handle.net/11449/93722. Acesso em: 05 jun. 2021.

MORAES, A. F.; OLIVEIRA, T. M. de. Experiências relacionadas ao levantamento de teses e dissertações. **Informação & Sociedade: Estudos**, João Pessoa, v.20, n.1, p. 73-81, jan./abr. 2010. Disponível em: https://www.arca.fiocruz.br/handle/icict/1990. Acesso em: 07 jun. 2021.

NOGUEIRA, E. D. A.; VAZ, M. S. M. G.; SOUZA, L. Um Padrão de Metadados para Descrição de Imagens Astronômicas do Tipo FITS. **Revista Ciências Exatas e Naturais**, v.12 nº 1, jan./jun. 2010. Acesso em: http://ri.uepg.br/riuepg/bitstream/handle/123456789/636/ARTIGO_Padr%C3%A3oMetadados.pdf?sequence=1. Acesso em: 07 jun. 2021.

PERIÓDICOS DE MINAS. **Entenda o que é Acesso Aberto**. 2021. Disponível em: https://www.periodicosdeminas.ufmg.br/entenda-o-que-e-acesso-aberto/. Acesso em: 08 jun. 2021.

SOUZA, M. I. F.; VENDRUSCULO, L. G.; MELO, G. C. Metadados para a descrição de recursos de informação eletrônica: utilização do padrão Dublin Core. **Ciência da Informação**, Brasília, v.29, n.1, p.93-102, jan./abr. 2000. Disponível em: https://doi.org/10.1590/S0100-19652000000100010. Acesso em: 06 jun. 2021.

TAKAHASHI, T. (Org.). **Sociedade da Informação no Brasil:** o livro verde. Brasília: Ministério da Ciência e Tecnologia, 2000.

WILKINSON, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. **Scientific Data**. Nature. 2016.