

Análise da representação da informação por meio de metadados: estudo comparativo do padrão de metadados do Governo Federal em portal de dados abertos de uma instituição federal

Analysis of the representation of information via metadata: a comparative study of Federal Government metadata patterns in a federal institution open data portal

Bárbara Carvalho Diniz

Doutoranda em Ciência da Informação pela Universidade Federal da Paraíba – UFPB, João Pessoa, PB, Brasil.
Professora substituta no curso de Arquivologia da Universidade Estadual da Paraíba – UEPB, João Pessoa, PB, Brasil.

ORCID: <https://orcid.org/0000-0002-0519-6589>

E-mail: dinizbarbara6@gmail.com

Guilherme Ataíde Dias

Doutor em Ciência da Informação pela Universidade de São Paulo – USP, São Paulo, SP, Brasil.
Professor do Departamento de Ciência da Informação Universidade Federal da Paraíba – UFPB, Brasil.

ORCID: <https://orcid.org/0000-0001-6576-0017>

E-mail: guilhermeataide@gmail.com

Resumo

Como forma de fortalecer a transparência pública e incentivar a participação dos cidadãos na fiscalização dos serviços e ações estatais, o Brasil aprovou leis e decretos para incentivar a cultura da transparência. Uma dessas iniciativas é o Decreto Federal nº 8.777/2016, que institui a política de dados abertos no Poder Executivo Federal. Para que os conjuntos de dados disponíveis em portais dos órgãos federais sejam encontrados e utilizados para os mais diversos fins, é necessário que estes sejam descritos por metadados. Os metadados contidos em portais do Governo Federal devem seguir o padrão de metadados do governo eletrônico. Sendo assim, o objetivo geral deste artigo é analisar os conjuntos de dados abertos da Universidade Federal do Rio Grande do Norte (UFRN) a partir do padrão de metadados do governo eletrônico. Dentre as instituições em nível universitário pesquisadas, a UFRN apresentou um número expressivo e atualizado de conjuntos de dados. Apresenta metodologia de uma pesquisa descritiva-documental, quali-quantitativa, em que os dados coletados foram analisados a partir da técnica de análise de conteúdo. Da amostra coletada dos conjuntos de dados, em grande parte os itens obrigatórios estavam presentes, como também se fez uso de elementos opcionais no intuito de enriquecer a descrição do conjunto de dados. No entanto, apesar das descrições serem concisas e atualizadas, alguns itens necessitam de mais precisão no que se refere à proveniência e publicador específico.

Palavras-chave: dados abertos; metadados; padrão de metadados do governo eletrônico.

Abstract

To strengthen public transparency and encourage citizen participation in monitoring state services and actions, Brazil has passed laws and decrees to encourage a culture of transparency. Federal Decree No. 8,777/2016 is one such initiatives, establishing an open data policy in the Federal Executive Branch. Finding the datasets available in federal agency portals and using it for several purposes requires their description by metadata. The metadata in Federal Government portals must follow the e-government metadata standard. Thus, the general objective of this study is to analyze Universidade Federal do Rio Grande do Norte open data sets (UFRN) based on the e-government metadata standard. Among the surveyed university-level institutions, UFRN showed an expressive and updated number of datasets. It offers a descriptive-documentary, quali-quantitative research methodology, in which the collected data were analyzed via content analysis. The collected data set sample had most mandatory items, as well as optional ones, used to enrich our data set description. However, despite its concision and up-to-dateness, some items require greater provenance and publisher precision.

Keywords: open data; metadata; e-government metadata standard.

1. Introdução

A partir da Constituição Federal do Brasil de 1988, o país passou por diversas transformações, principalmente no que se refere aos seus princípios e objetivos, em especial na transparência estatal. Em relação a este aspecto, vemos de forma clara o princípio administrativo da Publicidade, “[...] que faz com que sejam obrigatórios a divulgação e o fornecimento de informações de todos os atos praticados pela Administração Pública” (PINTO, 2008, p. 134). Além disso, o inciso XXXIII do artigo 37 da Constituição Federal, assegura que todos os cidadãos têm direito de receber informação dos órgãos públicos, seja de seu interesse particular ou de interesse coletivo, prestadas no prazo da Lei, sob risco de pena de responsabilidade ao órgão que não fornecer, com ressalva das informações de caráter sigiloso (BRASIL, 2020).

Estando cientes dos princípios e prerrogativas constitucionais, outras leis que impactam a transparência pública também foram promulgadas, a exemplo da Lei dos Arquivos (Lei nº 8.159 de 8 de janeiro de 1991) que dispõe sobre a política nacional de arquivos públicos e privados (BRASIL, 1991); a Lei de Acesso à Informação (Lei nº 12.527, de 18 de novembro de 2011) que regulamenta o acesso à informação de âmbito público (BRASIL, 2011); e o Decreto Federal nº 8.777, de 11 de maio de 2016, que institui a política de dados abertos do Poder Executivo Federal (BRASIL, 2016).

Quanto ao Decreto que institui a política de dados abertos, percebe-se que esta é uma das políticas mais recentes que no que se refere à transparência pública, estando ainda em processo de implantação em diversos órgãos federais. Tanto para dados, informações ou documentos é necessário que haja gestão, incluindo etapas de produção, recebimento, organização, recuperação e análise. Uma etapa primordial para que se possa alcançar o principal objetivo de tais prerrogativas legais, ou seja o acesso à informação, trata-se da Representação da Informação.

De acordo com Lima e Alvares (2012), representar é um ato de simbolizar por meio de palavras, figuras, esquemas, mímicas etc., na intenção de substituir uma ideia, um objeto ou um fato. Ainda de acordo com os autores, na Ciência da Informação, o principal objetivo da Representação da Informação é “[...] recuperar objetos informacionais, que são as informações registradas nos mais variados suportes – textos, imagens, registros sonoros, representações cartográficas e páginas web, entre outros” (LIMA; ALVARES, 2012, p. 35).

Ainda referente à política de dados abertos do Governo Federal brasileiro, entende-se que para proporcionar a organização, acesso e uso dos conjuntos de dados é preciso representá-los da melhor maneira possível, de forma padronizada e objetiva por meio de metadados. Estando cientes da relevância do uso de metadados, o Governo Federal criou no ano de 2014, o documento Padrão de Metadados do Governo Eletrônico (e-PMG), que busca padronizar o uso dos metadados em seus *sites* (BRASIL, 2014). Compreendemos que os metadados, que descrevem os conjuntos de dados abertos disponíveis em *sites* governamentais, são um recurso primordial para a encontrabilidade¹ das informações a serem disponibilizadas para a sociedade.

A principal justificativa desta proposta é ressaltar a relevância da transparência pública nas universidades públicas do país, pois diante das contingências impostas pelo Governo Federal, como cortes de verbas e questionamento acerca de sua importância na sociedade brasileira², é mais que necessário demonstrar e levar o conhecimento produzido, bem como efetuar ações universitárias para mais perto da população, além de ser mais transparente em relação aos seus atos e prestação de contas por meio dos portais de transparência, a exemplo dos dados abertos. A UFRN foi escolhida para esta análise não por acaso. Diante das nove universidades nordestinas que possuem sede nas capitais de seus Estados, onde na maioria das vezes concentram sua administração e grande número de cursos, a UFRN é a que possui maior número de conjuntos de dados disponíveis e atualizados no Portal de Dados Abertos do Governo Federal.

Este trabalho tem como **objetivo geral**: analisar os conjuntos de dados abertos da UFRN a partir do padrão de metadados do governo eletrônico. Já seus **objetivos específicos** pautam-se em: compreender conceitos sobre metadados, bem como aspectos de sua qualidade; conhecer os metadados do documento e-PMG; coletar metadados dos conjuntos de dados abertos disponíveis no portal de dados abertos da UFRN e; comparar os metadados coletados com o padrão e-PMG.

O artigo está organizado nas seguintes seções: Introdução, Representação da Informação e metadados, Padrão de metadados do Governo Eletrônico (e-PMG),

¹ “O termo encontrabilidade, traduzido do inglês *findability*, foi definido por Peter Morville no livro *Ambient Findability*, trazendo uma perspectiva técnica no contexto em que podemos encontrar uma informação em qualquer lugar e momento” (FERNANDES; VECHIATO, 2017, p. 3)

² “O contingenciamento de verbas no ensino superior tem sido defendido com o argumento de que as universidades são caras. De fato, essas instituições requerem vultosos recursos para cumprir suas funções [...] cabe a ela, cada vez mais, prestar contas, ser transparente e fazer contínua avaliação de suas ações, sempre à luz de sua autonomia” (BERTOTTI, 2019, s. p.).

Procedimentos Metodológicos, Análise dos metadados do portal de Dados Abertos da UFRN e Considerações finais.

2. Representação da Informação e metadados

Desde o início da história da sociedade, seja em clãs, tribos, aldeias, até chegar nas formas de organizações sociais mais modernas, o ser humano busca registrar o que aprende. A representação faz parte desse processo, pois diante do conhecimento e informações registradas existem maneiras de organizar o que se produz para serem recuperados e utilizados para os mais diversos fins.

Assim, “[...] a representação da informação consiste em uma prática de enunciação das propriedades de um objeto informacional (unidade de informação organizável) ou das relações desse objeto com outros que o identificam, através de operações que visam a sua descrição” (ALBUQUERQUE; GAUDÊNCIO; SANTOS, 2019, p. 17). Ainda de acordo com os autores, os objetos informacionais podem ser representados/descritos por meio de suas características físicas, como também pelo seu conteúdo temático.

Na concepção de Novellino (1996), a representação da informação é a substituição de uma entidade longa, a exemplo de um texto de um documento, por uma descrição breve, resumida, com principal intuito de demonstrar a essência do objeto informacional. Acrescenta que para representar informação é preciso: 1) analisar o assunto do documento e colocar o resultado desta análise numa expressão linguística; e 2) atribuir conceitos ao documento/objeto analisado. Sendo assim, a representação da informação proporciona a um objeto informacional uma representação clara, sucinta, de rápido entendimento, para que não seja necessário acessar o documento na íntegra a fim de saber do que se trata (VIGNOLLI; ALMEIDA; CATARINO, 2014).

Tendo em vista o crescimento e a consolidação das tecnologias da informação e comunicação, documentos das mais variadas espécies e conteúdos encontram-se disponíveis em formato digital. A imensidão de informações encontradas na *web* demanda que para ser encontrada e utilizada, é essencial representar. Para tais objetos em formato digital, surgiu o conceito de metadados, sendo entendidos como:

[...] campos utilizados para a descrição de recursos de informação em sistemas digitais. Geralmente, os sistemas que utilizam os metadados recuperam as informações por meio da descrição realizada em seus campos, e não em texto completo [...] (MEDEIROS, 2010, p. 76).

No Quadro 1 são descritos alguns conceitos sobre metadados.

Quadro 1 – Conceitos sobre metadados

Conceito	Autor
“dados associados com objetos que desoneram os usuários potenciais de ter conhecimento completo antecipado da existência e características desses objetos”	Dempsey e Heery, (1997 p. 5, tradução nossa)
“[...] um conjunto de dados, chamados ‘elementos’, cujo número varia de acordo com o padrão adotado, e que descreve o recurso, possibilitando a um usuário ou a um mecanismo de busca acessar e recuperar esse recurso.”	Grácio (2012, p. 121)
“Os metadados explicitam os diferentes aspectos do recurso que descreve: sua estrutura, conteúdo, qualidade, contexto, origem, propriedade e condição. E auxiliam na organização, favorecem a interatividade, validam as identificações e asseguram a preservação e principalmente, otimizam o fluxo informacional melhorando o acesso aos dados e a localização dos recursos informacionais”.	Santos, Simionato e Arakaki (2014, p. 150-151)
“Sejam criados automaticamente ou manualmente, os metadados são o meio pelo qual outros podem descobrir a existência dos objetos que representam. Eles também podem ser os meios pelos quais os dados são combinados, comparados e diferenciados. Os metadados podem ser cumulativos, com representações adicionadas, como dados usados e reutilizados em novos contextos, migrados para novas tecnologias e formatos, com curadoria por longos períodos de tempo”.	Borgman (2015, p. 67-68, tradução nossa)
“Metadados são informações estruturadas que descrevem características dos recursos informacionais para diferentes propósitos, como identificação, gestão, descoberta, recuperação, interoperabilidade, descrição, entre outros. A literatura da área costuma classificar os metadados em cinco tipos básicos: administrativos, descritivos, preservação, técnicos e uso, os quais são utilizados para a descrição de recursos informacionais em geral.”	Brandt e Vidotti (2019, p. 2)

Fonte: Elaborado pelos autores

Entendemos que os metadados auxiliam o processo de representar e recuperar informações em meios digitais, seja em repositórios, *sites*, bancos de dados, entre outros meios de armazenamento digital. Para que possam de fato ser úteis para representar e encontrar o objeto informacional desejado é preciso que sigam determinado padrão, para que assim o fluxo e acesso ocorram de forma ágil e efetiva. Neste estudo adota-se a definição de metadados de Brandt e Vidotti (2019) devido à sua completude do conceito e definição.

Mesmo reconhecendo a importância do uso dos metadados, existem preocupações em relação à sua qualidade, principalmente em assegurar se o valor atribuído ao metadado está

descrito da maneira correta (BETANCOURT; ROCHA, 2012). Sendo assim, a norma ISO/IEC 25012/2012 especifica um modelo de qualidade de dados³ de maneira multidimensional, que estabelece características de qualidade de dados, que devem ser levadas em conta ao avaliar as propriedades de um determinado produto de dados.

Quadro 2 – Qualidade do produto de dados

Qualidade de dados inerentes e dependentes do sistema	Qualidade de dados inerentes ao sistema	Qualidade de dados dependentes do sistema
Acessibilidade: os dados podem ser acessados em um contexto específico, em particular por pessoas que possuem limitações e necessitam que as tecnologias apoiem uma configuração especial.	Exatidão: os dados representam corretamente o verdadeiro valor do atributo, descrito a partir de um conceito ou de um evento específico. Divide-se em exatidão sintática e semântica.	Disponibilidade: os dados possuem atributos que permitem ser obtidos por usuários e/ou aplicações autorizadas em um contexto de uso específico.
Conformidade: os dados têm atributos que atendem a regulamentos ou padrões atuais, bem como regras semelhantes relacionada à qualidade de dados num contexto específico.	Completo: os dados associados a uma entidade possuem valores para todos os atributos e entidades relacionadas a determinado contexto de uso.	Portabilidade: permitem que sejam instalados, substituídos ou eliminados de um sistema para outro, preservando o nível de qualidade em um determinado contexto de uso.
Confidencialidade: os dados têm atributos que asseguram que só pessoas autorizadas pelo sistema podem acessá-los.	Consistência: os dados estão livres de contradição. São consistentes com outros dados em um contexto de uso específico.	Recuperabilidade: os dados permitem manter e preservar um nível específico de operações e qualidade, inclusive em casos de falha, num contexto específico.
Eficiência: os dados têm atributos que podem ser processados e fornecidos com os níveis de desempenho esperados usando quantidades e tipos de recursos apropriados em um contexto de uso específico.	Credibilidade: os dados possuem atributos que os consideram verdadeiros e confiáveis em um contexto de uso específico. A credibilidade inclui o conceito de autenticidade.	
Precisão: os dados possuem atributos que são precisos ou fornecem uma visão em determinado contexto específico.	Atualidade: os dados são atualizados em contexto específico.	
Rastreabilidade: possuem atributos que proporcionam um caminho de acesso auditável aos dados ou a qualquer mudança realizada nos dados.		
Compreensível: os dados podem ser lidos e interpretados por usuários. São expressos utilizando linguagens, símbolos e unidades apropriadas em um contexto. Certas informações sobre compreensão dos dados são expressas por meio de metadados.		

Fonte: Elaborado pelos autores com base em ISO 25000 (2012)

³ A qualidade do produto de dados pode ser entendida como o grau em que os dados satisfazem determinados requisitos, definidos por uma instituição a que pertence o produto de dados (ISO 25012/2012).

Nesse contexto, os metadados representam as características de um objeto digital, descrevendo de forma sucinta e objetiva o seu conteúdo e formato. É preciso lembrar que sem representação, encontrar um conteúdo ou objeto torna-se caótico. Desta forma, além de contextualizar do que se trata um objeto digital, ou conjunto de dados, os metadados devem ser utilizados como ferramenta para buscar e recuperar informações, tornando a navegação em um *site*, repositório ou banco de dados mais rápida, precisa e satisfatória.

3. Padrão de metadados do governo eletrônico e-PMG

O Governo Federal brasileiro lançou no ano de 2014, o documento “Padrão de Metadados do Governo Eletrônico (e-PMG)”. Este documento, em sua versão 1.1, consiste em explicar e orientar o uso de elementos descritivos: metadados para uso na descrição dos recursos informacionais nas plataformas e *sites* do Governo Federal (BRASIL, 2014).

Moreira et al. (2017) afirmam que é por meio de um sistema de recuperação da informação que se torna possível realizar pesquisas no Portal Brasileiro de Dados abertos, podendo ser encontrados dados referentes a segurança pública, saúde, gastos do governo, processo eleitoral, educação etc.

De acordo com o documento, o padrão de metadados e-PMG foi desenvolvido com base no padrão Dublin Core Metadata Initiative (DCMI). Além dos 15 elementos presentes nos metadados tipo Dublin Core, foram acrescentados mais cinco, necessários para o contexto do Governo Brasileiro. Importante ressaltar que o padrão e-PMG pode descrever os recursos informacionais, e serviços disponíveis na internet; adota qualificadores adicionais e diferentes critérios em relação a obrigatoriedade, ou não, de certos elementos (BRASIL, 2014).

O objetivo do e-PMG é assegurar que as pessoas que pesquisam as informações do governo brasileiro na *Web* tenham acesso rápido e eficiente a descrições dos recursos. Os elementos e-PMG têm o propósito de facilitar as pessoas localizar os recursos que precisam, mesmo sem possuir conhecimento detalhado da localização ou da entidade responsável pelos mesmos (BRASIL, 2014, p. 4).

No que se refere aos 20 elementos do e-PMG, o Quadro 3 busca listar e descrever todos os elementos, além de dizer se são obrigatórios ou não.

Quadro 3 – Elementos dos metadados e-PMG

Elemento	Descrição	Obrigatoriedade
Abrangência	Descreve a informação presente no recurso, se o recurso estiver relacionado a um determinado tempo ou lugar específico.	Obrigatório se aplicável
Assunto	Termos/palavras-chaves que representam/identificam o conteúdo do recurso. Busca identificar de forma sucinta a temática relacionada ao recurso.	Obrigatório para categoria; Opcional para palavras-chave e pessoa
Colaborador	Pessoa ou organização que contribuiu para a elaboração do recurso, seja em sua elaboração ou divulgação.	Opcional
Contexto jurídico-administrativo	Normas/legislações que autorizam e regulamentam o funcionamento e desenvolvimento das funções e atividades do órgão produtor ou disseminador do recurso.	Opcional
Criador	Pessoa ou organização responsável pela criação do recurso.	Obrigatório
Data	Data associada ao recurso em seu ciclo de vida.	Obrigatório
Descrição	Resumo e explicação breve sobre o conteúdo do recurso.	Opcional
Destinação	Decisão sobre a destinação final dos documentos: guarda permanente ou eliminação.	Obrigatório apenas para documentos arquivísticos.
Destinatário	Pessoa ou organização para quem a informação ou documento foi encaminhado.	Obrigatório se aplicável
Direitos	Informações sobre as restrições e permissões sobre o acesso e uso dos recursos.	Opcional
Fonte	Identifica o recurso que deu origem ao mais novo recurso.	Opcional
Formato	Refere-se aos formatos do recurso, a exemplo de <i>software</i> e <i>hardware</i> necessários para apresentar um objeto.	Opcional
Identificador	Identifica de forma única e persistente um recurso informacional, a exemplo de um código de classificação arquivística.	Obrigatório para identificador de recurso e documentos arquivísticos; Obrigatório se aplicável à versão.
Idioma	Idioma que está expresso no conteúdo do recurso.	Opcional
Localização	Local onde está armazenado o recurso, seja físico ou digital.	Opcional
Preservação	Informação para apoiar estratégias de acesso, uso e preservação do recurso a longo prazo.	Opcional
Publicador	Organização ou pessoa responsável por divulgar o recurso disponível.	Obrigatório se aplicável
Relação	Referência a outro recurso relacionado.	Opcional
Tipo	Gênero ou natureza documental do recurso.	Obrigatório se aplicável
Título	Palavra ou frase que nomeia um recurso, podendo corresponder a sua denominação oficial/legal.	Obrigatório; Opcional para título alternativo

Fonte: Elaborado pelos autores com base em Brasil (2014).

No Quadro 3 o padrão de metadados e-PMG possui 20 elementos, sendo cinco obrigatórios, cinco obrigatórios se aplicáveis e dez opcionais. Quanto mais o objeto/recurso digital for descrito, mais fácil será encontrá-lo diante de uma gama de conjuntos de dados. Salienta-se que é imprescindível que os metadados estejam atualizados, como também correspondam as necessidades do público que o busca.

4. Procedimentos metodológicos

O percurso metodológico representado neste trabalho é do tipo descritivo-documental. Desta forma, descreveu-se alguns dos conjuntos de dados coletados no Portal de Dados Abertos da UFRN, objeto de análise desta pesquisa. É documental, já que se estudou os elementos do documento de Padrão de Metadados do Governo Federal (e-PMG) para serem utilizados como premissa de análise e comparação com os conjuntos de dados coletados.

A pesquisa possui aspectos quali-quantitativos, pois além de quantificar os conjuntos de dados a serem escolhidos e analisados, também contabilizou a quantidade de elementos obrigatórios e opcionais utilizados nos metadados observados. Também é qualitativa, já que os conteúdos e descrições presentes nos metadados foram observados e analisados perante o documento e-PMG e outras fontes teóricas sobre metadados. É caracterizado também como estudo de caso, pois é um estudo detalhado sobre o uso de metadados no Portal de Dados Abertos da UFRN.

Ressalta-se que a escolha deste portal não foi feita de modo aleatório, considerou-se que a região brasileira com maior número de estados é o Nordeste, composta por nove (Bahia, Pernambuco, Paraíba, Alagoas, Ceará, Maranhão, Sergipe, Rio Grande do Norte e Piauí) e suas capitais abrigam as universidades federais mais antigas desses estados, com grande relevância social em sua região. Por isso, selecionou-se: Universidade Federal da Bahia (UFBA), Universidade Federal de Pernambuco (UFPE), Universidade Federal da Paraíba (UFPB), Universidade Federal de Alagoas (UFAL), Universidade Federal do Ceará (UFC), Universidade Federal do Maranhão (UFMA), Universidade Federal de Sergipe (UFS), Universidade Federal do Rio Grande do Norte (UFRN) e Universidade Federal do Piauí (UFPI).

Foram verificados se os portais de dados abertos dessas universidades existiam, se possuíam conjuntos de dados e se estavam atualizados. Dentre as nove universidades pesquisadas, verificou-se que a Universidade Federal do Rio Grande do Norte é a que possui um maior número de conjunto de dados (84), considerando a data da pesquisa, agosto de 2021, e com conjuntos atualizados. Por conta da numerosa quantidade de dados disponíveis não será possível analisar os metadados presentes em todos os conjuntos de dados, sendo assim estabeleceu-se uma amostra de 10% do total dos conjuntos de dados, ou seja, serão analisados os metadados de 8 conjuntos, contidos nos grupos de: Ensino, Pesquisa, Extensão e

Institucional. Dentre os grupos citados, serão analisados dois conjuntos de dados, dado a ordem de atualização dos dados.

A análise de conteúdo foi feita mediante comparação entre os elementos presentes nos conjuntos de dados e os elementos descritos no padrão de metadados e-PMG do Governo Federal brasileiro, bem como o padrão de qualidade de dados da ISO 25012/2012, verificando sua aplicabilidade e uso pelo portal de dados analisado.

Salienta-se que houve buscas sobre análise de dados abertos governamentais na base de dados BRAPCI, a qual é voltada para Ciência da Informação. As buscas foram feitas a partir das palavras “dados abertos” e “dados abertos governamentais”, “e-pmg” recuperando respectivamente, 106, 16 e 0 trabalhos, entre artigos científicos, trabalhos completos e resumos expandidos apresentados em anais de eventos. Dos trabalhos analisados, apenas um (1), intitulado “Metadados para descrição de *datasets* e recursos informacionais do ‘Portal Brasileiro de Dados Abertos’” de Moreira et al (2018), aborda sobre os metadados do padrão e-PMG e sua aplicação no Portal de Dados Abertos.

5. Análise dos metadados do Portal de Dados Abertos da Universidade Federal do Rio Grande do Norte

O Portal de Dados Abertos da Universidade Federal do Rio Grande do Norte (UFRN) é um ambiente informacional que busca disponibilizar dados de forma pública e transparente para todos os cidadãos que desejem acompanhar as ações e atividades da universidade. Além disso, os dados disponibilizados também podem ser utilizados para o desenvolvimento de interfaces de programação de aplicações (APIs), um princípio básico de utilização dos dados governamentais abertos.

Como abordado anteriormente, o portal possui 84 conjuntos de dados divididos em 12 grupos ou categorias, são elas: biblioteca; comunicados e documentos; contratos e convênios; despesas e orçamentos; ensino; extensão; institucional; patrimônio; materiais; processos; pesquisa e pessoas.

Figura 1 – Página inicial do Portal de Dados Abertos UFRN



Fonte: Portal de dados abertos UFRN (2021).

Dos 12 grupos, foram escolhidos quatro de maneira aleatória, são eles: Ensino, Pesquisa, Extensão e Institucional, e de cada um desses grupos, foram escolhidos dois conjuntos de dados, dando preferência para os mais atualizados, até a data da coleta de dados. Logo em seguida são apresentadas as análises divididas pelos grupos de dados.

Os dois conjuntos de dados mais atualizados do grupo ensino foram: docentes (Quadro 4) e diplomas (Quadro 5). Além desses conjuntos, o grupo de dados de ensino também contém conjunto de dados sobre turmas, discentes, empréstimo dos acervos da biblioteca, planos individuais docentes, programas de pós-graduação, docentes externos, matrículas em componentes, projetos de ensino, cursos da UFRN e mais nove conjuntos de dados.

Quadro 4 – Elementos de metadados do conjunto Docentes

Obrigatórios	Opcionais
Título: docentes	Descrição: relação dos docentes da UFRN.
Criador: suporte-UFRN	Formato: PDF e CSV
Data: 28 de setembro de 2016/ atualizado em junho de 2021.	Palavras-chave: docentes; professores; servidores
Assunto: Ensino	
Publicador:portaldadosabertos@info.ufrn.br	

Fonte: Portal de Dados Abertos UFRN (2021).

Quadro 5 – Elementos de metadados do conjunto Diplomas

Obrigatórios	Opcionais
Título: diplomas	Descrição: diplomas expedidos pela UFRN para os níveis de graduação, mestrado, doutorado e especialização
Criador: suporte-UFRN	Formato: PDF e CSV
Data: 2 de junho de 2021	Palavras-chave: diploma; doutorado; especialização; mestrado; graduação
Assunto: Ensino	
Publicador:portaldadosabertos@info.ufrn.br	

Fonte: Portal de dados abertos UFRN (2021).

Os dois conjuntos de dados mais recentes do grupo pesquisa foram bolsistas de iniciação científica (Quadro 6) e pesquisadores (Quadro 7). Tal conjunto contém dados sobre tudo que se refere à atividade de pesquisa da UFRN, a exemplo da relação de pesquisadores da universidade, bolsistas de iniciação científica, projetos de pesquisa, indicadores de pesquisa e grupos de pesquisa.

Quadro 6 – Elemento de metadados do conjunto Bolsistas de iniciação científica

Obrigatórios	Opcionais
Título: bolsistas de iniciação científica	Descrição: relação de bolsistas de iniciação científica de projetos de pesquisa da UFRN
Criador: suporte-UFRN	Formato: PDF e CSV.
Data: 13 de outubro de 2016/ atualizado em agosto de 2021.	Palavras-chave: bases de pesquisa; bolsa de iniciação científica; bolsa de pesquisa; bolsistas de iniciação científica; bolsistas de pesquisa; pesquisa; projetos; pesquisadores; projetos de pesquisa.
Assunto: pesquisa	
Publicador:portaldadosabertos@info.ufrn.br	

Fonte: Portal de Dados Abertos UFRN (2021).

Quadro 7– Elementos de metadados do conjunto pesquisadores

Obrigatórios	Opcionais
Título: Pesquisadores.	Descrição: relação de pesquisadores da UFRN.
Criador: suporte-UFRN	Formato: PDF e CSV.
Data: 11 de outubro de 2016/ atualizado em agosto de 2021.	Palavras-chave: pesquisadores; projetos de pesquisa.
Assunto: pesquisa	
Publicador:portaldadosabertos@info.ufrn.br	

Fonte: Portal de Dados Abertos UFRN (2021).

Os conjuntos de dados mais recentes da categoria extensão foram produtos de extensão (Quadro 8) e atividades de extensão (Quadro 9). Ressalta-se que existem apenas dois conjuntos de dados nessa categoria, que são os dois citados a seguir.

Quadro 8 – Elementos de metadados do conjunto produtos de extensão

Obrigatórios	Opcionais
Título: produtos de extensão	Descrição: relação de produtos de extensão gerados pela UFRN.
Criador: suporte-UFRN	Formato: PDF e CSV.
Data: 11 de outubro de 2016/ atualizado em agosto de 2021.	Palavras-chave: extensão; produtos de extensão.
Assunto: extensão	
Publicador:portaldadosabertos@info.ufrn.br	

Fonte: Portal de Dados Abertos UFRN (2021).

Quadro 9 – Elementos de metadados do conjunto atividades de extensão

Obrigatórios	Opcionais
Título: atividades de extensão	Descrição: relação das atividades de extensão da UFRN.
Criador: suporte-UFRN	Formato: PDF e CSV.
Data: 12 de setembro de 2016/ atualizado em agosto de 2021.	Palavras-chave: atividades; extensão; projetos.
Assunto: extensão	
Publicador:portaldadosabertos@info.ufrn.br	

Fonte: Portal de Dados Abertos UFRN (2021).

Os conjuntos de dados mais atualizados do grupo institucional são Quadro de referência dos servidores técnicos administrativos (Quadro 10) e avaliação de docência (Quadro 11). Outros conjuntos os quais pertencem a este grupo e que se referem as atividades institucionais da UFRN são unidades acadêmicas, concursos, telefones e unidades.

Quadro 10 – Elementos de metadados do conjunto Quadro de referência dos servidores técnicos administrativos

Obrigatórios	Opcionais
Título: Quadro de referência dos servidores técnicos administrativos.	Descrição: Quadro de Referência dos Servidores Técnico-Administrativos (QRSTA) da UFRN. O QRSTA é uma tabela com os quantitativos de lotação dos cargos dos níveis de classificação “C”, “D” e “E” integrantes do Plano de Carreira dos Cargos Técnico-Administrativos em Educação.
Criador: não especificado.	Formato: PDF e CSV.
Data: 25 de junho de 2020/ atualização em julho de 2021.	Palavras-chave: cargos; servidores; técnicos administrativos; vagas.
Assunto: institucional	
Publicador:portaldadosabertos@info.ufrn.br	

Fonte: Portal de Dados Abertos UFRN (2021).

Quadro 11 – Elementos de metadados do conjunto avaliação de docência

Obrigatórios	Opcionais
Título: Avaliação de docência	Descrição: relação de avaliações de docência da instituição.
Criador: suporte- UFRN	Formato: PDF e CSV.
Data: 23 de outubro de 2017/ atualização julho de 2021.	Palavras-chave: avaliação de docentes; avaliações; docente; indicadores.
Assunto: institucional	
Publicador:portaldadosabertos@info.ufrn.br	

Fonte: Portal de Dados Abertos UFRN (2021).

De modo geral, sete dos oito conjuntos de dados apresentaram todos os elementos obrigatórios com vistas ao padrão e-PMG, foram eles: **título, criador, data, categoria e publicador**. Teoricamente, são cinco elementos obrigatórios, no entanto o elemento “identificador” é obrigatório apenas quando descrever um recurso, ou seja, um objeto dentro do conjunto de dados, não sendo obrigatório nesta análise.

Os elementos título e assunto (referindo-se à categoria) foram descritos. Foram descritos: título descreve o nome do conjunto de dados e o assunto classifica quais assuntos os conjuntos de dados tratam ou se referem.

O elemento obrigatório “data” aparece descrito em todos os conjuntos analisados, além disso, tal elemento possui a data de atualização do recurso. O elemento que descreve a data de atualização e a frequência que o recurso deve ser atualizado faz total diferença, pois a partir disso infere-se que existe controle e maior precisão do conteúdo presentes nos conjuntos de dados. Sendo assim, este elemento cumpre o critério de qualidade **atualidade**: dados atualizados para uso em contexto específico (ISO 25012, 2012).

O elemento publicador no Portal de Dados Abertos indica quem é o responsável por disponibilizar o recurso no portal. Este elemento está preenchido em todos os conjuntos, no entanto aparece um endereço de *e-mail* e não identifica o nome do órgão, diretoria, coordenação ou pessoa que o fez, transparecendo falta de exatidão, também encontrada em outro elemento (criador).

No que remete ao elemento criador, apenas o conjunto “Referência dos servidores técnicos administrativos” (Quadro 10) não especifica quem é seu criador, ferindo o princípio da **completude**, pois um conjunto de dados necessita estar descrito de maneira completa para que possa fazer sentido em seu contexto de uso (ISO 25012, 2012). Ressalta-se que por mais que o elemento criador esteja descrito em quase todos os conjuntos (7/8), ele se refere muito

mais a quem o publicou (suporte UFRN), do que de fato a quem o produziu. Isto é, indica apenas que foi criado pelo órgão maior: UFRN, mas sem especificar uma pessoa, diretoria, departamento etc. De certa forma, a descrição inespecífica de quem criou o conjunto de dados, contradiz uma característica importante da qualidade de dados, a **exatidão**, em que se deve representar corretamente o verdadeiro valor de um dado ou atributo (ISO 25012, 2012).

Os elementos opcionais presentes em todos os oito conjuntos foram: descrição, formato e palavras-chave. Sobre o elemento descrição, todas são sucintas, detalham de forma mais específica do que se trata o conjunto de dados, atingindo o objetivo do elemento. Tal elemento parece dotar o conjunto de dados com a qualidade da **precisão**: ou seja, oferece uma visão geral do que se trata o conjunto de dados (ISO 25012, 2012).

O elemento formato indica o formato de arquivo. No caso do conjunto de dados analisados é informado que os recursos estão no formato CSV e PDF.

Encerrando os elementos opcionais, as palavras-chave estão presentes em todos os conjuntos de dados, sendo uma maneira de classificar e enquadrar os conjuntos em uma temática, além de serem uma forma de buscar e recuperar, apoiando a pesquisa direta.

6. Considerações finais

A partir das análises sobre o Portal de Dados Abertos, levando em consideração o padrão de metadados e-PMG, obteve-se algumas conclusões. A maioria dos metadados obrigatórios (título, criador, data, assunto e publicador) estão sendo utilizados, correspondendo às mínimas expectativas. No entanto, os elementos criador e publicador pecam pela generalização e não especificam quem de fato o gerou (proveniência) ou o publicou, todos se resumem a UFRN, sendo apenas a representação maior de organismos menores que a compõem (coordenações, pró-reitorias, diretorias, departamentos etc.). Assim é necessária maior precisão na descrição de tais itens. Acentua-se que o elemento data está em perfeita conformidade com o esperado, além de atualizado, dando mais credibilidade ao uso dos dados.

Os elementos opcionais complementam e facilitam o entendimento sobre os conjuntos de dados e o uso de alguns deles, a exemplo das palavras-chaves, contribui no processo de busca e recuperação dos itens.

Infere-se que a principal limitação desta pesquisa se trata da limitação da análise total dos conjuntos de dados, que por sua quantidade (84 conjuntos), tornou-se inviável para este tipo de trabalho. Para estudos futuros, sugere-se a análise dos conjuntos de dados a partir dos princípios FAIR, para a análise se os dados são encontráveis (*Findable*), acessíveis (*Accessible*), interoperáveis (*Interoperable*) e reusáveis (*Reusable*), investigando de forma mais minuciosa se os dados disponíveis terão um maior potencial de uso, visto que o grande objetivo dos dados abertos é a transparência e reuso pelos cidadãos.

Este trabalho teve como propósito demonstrar a importância dos dados abertos para a sociedade e como estes podem ser relevantes para o acesso à informação das atividades e funções da universidade pública brasileira, principalmente no que se refere à prestação de contas.

Por fim, salienta-se que o Portal de Dados Abertos da UFRN se destaca por publicar de maneira constante e aberta seus dados, contribuindo para a transparência pública de suas ações e sendo exemplo para outras instituições, que, por diversos motivos, ainda não possuem seu portal ou não os mantêm atualizado. Espera-se que, por mais que os metadados dos conjuntos de dados não estejam totalmente adequados aos padrões, ao longo do tempo possam ser aperfeiçoados conforme uma adequada representação da informação.

Referências

ALBUQUERQUE, M. E. B.; GAUDÊNCIO, S. M.; SANTOS, R. F. Reflexões teóricas em representação da informação. *In*: ALBUQUERQUE, M. E. B.; MARTINS, G. K.; MOTA, D. A. R. (Org.). **Organização e representação da informação e do conhecimento**: intersecções teórico-sociais. João Pessoa: Editora UFPB, 2019. p. 13-28.

BERTOTTI, M. Os ataques à universidade pública. **Jornal da USP**, São Paulo, 14 maio 2019. Disponível em: <https://jornal.usp.br/artigos/os-ataques-a-universidade-publica/>. Acesso em: 11 ago. 2021.

BENTACOURT, S. S. M. P.; ROCHA, R. P. Metadados de qualidade e visibilidade na comunicação científica. **Encontros bibli**, Florianópolis, v. 17, n. esp. 2 – III SBCC, p. 82-101, 2012. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2012v17nesp2p82/23571>. Acesso em: 11 ago. 2021.

BORGMAN, C. L. **Big data, little data, no data**: scholarship in the networked world. Cambridge: MIT press, 2015.

BRANDT, M. B.; VIDOTTI, S. A. B. G. Metadados de negócio: representação da informação dos processos de trabalho. **Transinformação**, Campinas, v. 31, p. 1-11, 2019. Disponível em: <https://www.scielo.br/j/tinf/a/pmPqm9WwH6LKPKp9FWLCHsC/?format=pdf&lang=pt>. Acesso em: 11 ago. 2021.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. 56. ed. Brasília: Câmara dos Deputados: Edições Câmara, 2020.

BRASIL. **Lei 8.159, de 08 de janeiro de 1991**. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. Brasília, DF: Presidência da República, [1991]. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/L8159.htm. Acesso em: 11 ago. 2021.

BRASIL. **Lei 12.527, de 18 de novembro de 2011**. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei no 8.112, de 11 de dezembro de 1990; revoga a Lei no 11.111, de 5 de maio de 2005, e dispositivos da Lei no 8.159, de 8 de janeiro de 1991; e dá outras providências. Brasília, DF: Presidência da República, [2011]. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm. Acesso em: 11 ago. 2021.

BRASIL. **Padrão de metadados do governo eletrônico e-PMG**. Versão 1.1. Brasília: Comitê Executivo de Governo Eletrônico, julho de 2014. Disponível em: https://www.gov.br/governodigital/pt-br/governanca-de-dados/PMGVersao1_1.pdf. Acesso em: 11 ago. 2011.

BRASIL. **Decreto nº. 8.777, de 11 de maio de 2016**. Institui a política de dados abertos do poder executivo federal. Brasília, DF: Presidência da República, [2016]. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/d8777.htm. Acesso em: 11 ago. 2021.

DADOS ABERTOS UFRN. **Dados abertos UFRN**. 2021. Disponível em: <https://dados.ufrn.br>. Acesso em: 11 ago. 2021.

DEMPSEY, L.; HEERY, R. **A review of metadata**: a survey of current resource description formats. Work Package 3 of telematics for research Project DESIRE (RE1004), 1997.

FERNANDES, W. M.; VECHIATO, F. L. Encontrabilidade da informação no repositório institucional da UNESP: uma avaliação com dispositivos móveis. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 18., 2017, Marília. **Anais...** Marília: FFC: ANCIB, 2017. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/104997>. Acesso em: 28 set. 2021.

GRÁCIO, J. C. A. **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/bitstream/handle/11449/113727/ISBN9788579833335.pdf?sequence=1&isAllowed=y>. Acesso em: 11 ago. 2021

ISO 25000. **ISO 25012**: calidad de produto de datos. [S. l.], 2012. Disponível em: <https://iso25000.com/index.php/normas-iso-25000/iso-25012>. Acesso em: 11 ago. 2021.

LIMA, J. L. O.; ALVARES, L. Organização e representação da informação e do conhecimento. In: ALVARES, L. (Org.). **Organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações**. São Paulo: B4 editora, 2012. p. 21-48.

MEDEIROS, G.M. **Organização da informação em repositórios digitais**: implicações do auto-arquivamento na representação da informação. 2010. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Santa Catarina – UFSC, Florianópolis, 2010. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/94615/285680.pdf?sequence=1&isAllowed=y>. Acesso em: 11 ago. 2021

MOREIRA, F.M. *et. al.* Metadados para descrição de datasets e recursos informacionais do “portal brasileiro de dados abertos”. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 22, n. 3, p. 158-185, jul./set. 2017. Disponível em: <https://www.scielo.br/j/pci/a/nsNf68fmh3y4tNnh3XpjCZG/?format=pdf&lang=pt>. Acesso em: 11 ago. 2021.

NOVELLINO, M.S.F. Instrumentos e metodologias de representação da informação. **Informação & Informação**, Londrina, v. 1, n. 2, p. 37-45, jul./dez. 1996. Disponível em: <https://brapci.inf.br/index.php/res/download/45628>. Acesso em: 11 ago. 2021.

PINTO, A.G.G. Os princípios mais relevantes do direito administrativo. **Revista da EMERJ**, Rio de Janeiro, v. 11, n. 42, p. 130-141, 2008. Disponível em: https://www.emerj.tjrj.jus.br/revistaemerj_online/edicoes/revista42/Revista42_130.pdf. Acesso em: 11 ago. 2021.

SANTOS, P. L. V. A. C.; SIMIONATO, A. C; ARARAKI, F. A. Definição de metadados para recursos informacionais: apresentação da metodologia BEAM. **Informação & Informação**, Londrina, v. 19, n. 1, p. 146-163, jan./abr.2014. Disponível em: <https://brapci.inf.br/index.php/res/download/44884>. Acesso em: 11 ago. 2021.

VIGNOLLI, R. G; ALMEIDA, P. O. P; CATARINI, M. E. Folksonomias como ferramenta da organização e representação da informação. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 12, n. 2, p. 120-135, maio/ago. 2014. Disponível em: <https://brapci.inf.br/index.php/res/download/114314>. Acesso em: 19 ago. 2021.

Artigo submetido em: 16 nov. 2011
Artigo aceito em: 20 out. 2022