



DADOS E METADADOS: **reflexões conceituais**

Felipe Augusto Arakaki (UnB)
Ana Carolina Simionato Arakaki (UFSCar)

Introdução



Em áreas com escopo interdisciplinar:

- **conceitos similares** são usados com **terminologias diferentes**
- O termo dado é usado em diversos contextos, como Linked Data, Big Data, e-Science, Ciência de dados e da Web de dados.
- Na área de Ciência da Informação, o termo dado obteve um novo contexto, influenciado ao conjunto de atualizações terminológicas e conceituais do termo metadado.

Introdução



Diante das diversas definições de dado e metadado, e da própria evolução dos conceitos relacionados a esses dois termos:

- Baker et al. (2011) ressalta a necessidade de **discutir os conceitos** que são utilizados e aproximados entre as áreas de **Ciência da Informação e Ciência da Computação**.
- Furner (2019) aponta a necessidade de **discutir e comparar o conceito de metadados** com outros termos, como exemplo: **dados, documento, informação e os dados do próprio registro**.

Objetivos



O objetivo deste trabalho consiste em discutir e relacionar os conceitos dos termos dado e metadado, no intuito de debater a proximidade dos dois conceitos na Ciência da Informação.

Procedimentos Metodológicos



Caracterizada por uma metodologia de **análise exploratória** para identificar os elementos conceituais, a partir da literatura científica da área de Ciência da Informação, os **conceitos de dados e metadados**.

Para a elucidação dos termos, utilizou-se o aporte teórico do **Perspectivismo** definido por Peterson (1996) e contextualizado para Ciência da Informação por Santos e Vidotti (2009).

Procedimentos Metodológicos



“De acordo com Peterson (1996), a **adequação de uma perspectiva é sempre relativa**, mas ela não se estabelece de modo arbitrário. Na escolha de uma perspectiva **deve-se considerar a sua adequação ao usuário** a uma situação determinada, a um processo, a uma ontologia e a uma meta específica. É a análise de cada um destes componentes que se pretende desenvolver no estudo dos processos que atuam nas diretrizes, modelagens e estruturas de sistemas para atendimento de necessidades de sujeitos em ambientes informacionais específicos.”
(SANTOS; VIDOTTI, 2009).

Resultados e discussões



Ao longo dos anos, autores como Méndez Rodríguez (2002); Alves (2010); Alves e Santos (2013); Joudrey, Taylor e Wisser (2018); Pomerantz (2015); Gilliland (1998, 2008, 2016); Zeng e Qin (2008, 2016), têm discutido e estruturado definições sobre o termo metadados na perspectiva da Ciência da Informação.

Dentre esses autores, o que é possível de identificar como ponto em comum é que **o conceito de metadados está atribuído a uma informação estruturada para as ações de identificação, descoberta, seleção, uso, acesso e gerenciamento.**

O termo metadado é conceituado por meio das “[...] informações de valor agregado que criam para **organizar, descrever, rastrear e melhorar o acesso a objetos de informação e itens físicos e coleções, relacionados a esses objetos**”. (GILLILAND, 2016, p. 02, tradução nossa).

Os metadados inicialmente foram identificados pela expressão ‘dados sobre dados’, cunhada na década de 60 para se referir a um conjunto de declarações sobre os dados (POMERANTZ, 2015). **Mas é perceptível que os metadados fazem parte da rotina de diversas comunidades profissionais que projetam, criam, descrevem, preservam e usam sistemas e recursos informacionais** (GILLILAND, 2016).

A partir da análise terminológica realizada por Arakaki (2019, p. 80-81), as tipologias são identificadas como:

- Metadados administrativos;
- Metadados de autenticação;
- Metadados de preservação;
- Metadados de proveniência;
- Metadados técnicos;
- Metametadata;
- Metadados descritivos;
- Metadados de direitos;
- Metadados de acesso e uso;
- Metadados estruturais;
- *Markup languages*.

Os dados são destacados inicialmente pelo conceito atribuído por Santos e Sant'Ana (2013, p. 205) que definem que dado como “[...] uma unidade de conteúdo necessariamente relacionada a determinado contexto e composta pela tríade entidade, atributo e valor, de tal forma que, mesmo que não esteja explícito o detalhamento sobre contexto do conteúdo, ele deverá estar disponível de modo implícito no utilizador, permitindo, portanto, sua plena interpretação.”

A partir da definição de Santos e Sant'Ana (2014), observa-se que o dado é composto pela **tríade entidade, atributo e valor e apesar de implícito, o dado sempre está atrelado a um contexto.**

Furner (2016) faz um levantamento do conceito de dados ao longo dos séculos e discute a partir da perspectiva histórica a mudança do conceito.

- **Abordagem extensional:** busca caracterizar coisas ou tipos de coisas que se enquadram no conceito de “dados”;
- **Abordagem intencional:** identifica as propriedades que algo deve ter se for ser tratado como dados;
- **Abordagem classificatória:** reconhece um conceito individual como “dados”; pode ter, ou possuir múltiplos sentidos, e que esses sentidos podem ser categorizados de acordo com similaridades em função e contexto.
- **Abordagem histórica:** que, ao invés de ou além delas, conduzir análises lógicas e / ou computacionais das propriedades necessárias dos conceitos, permitir que os autores considerem o desenvolvimento culturalmente específico dos significados de termos como dados ao longo do tempo.

A partir dessas abordagens (extensional, intensional, classificatória e histórica), Furner (2016) destaca as diversas interpretações que os dados podem assumir.

- A interpretação clássica: origem do termo do latim;
- **A interpretação documental: dados como metadados;**
- A interpretação eclesiástica: dados como dons de Deus;
- A interpretação geométrica: dados como premissas geométricas.
- A interpretação matemática: dados como premissas matemáticas.
- A interpretação epistêmica: dados como evidência.
- **A interpretação informacional: dados como valores de atributo.**
- **A interpretação computacional: dados como bits.**
- A interpretação diafórica: os dados são atribuídos como realidade objetiva, aparências subjetivas, observações, idéias, significados, ou mesmo, expressões linguísticas de observações individuais.

A literatura aponta que em muitos casos, dados e metadados **são tratados como sinônimos.**

Segundo Wickett et al. (2013) “[...] os componentes de dados e metadados estão entrelaçados: nenhuma distinção estrutural permite uma discriminação imediata entre dados e metadados.”

Jeffery et al. (2014) distinguem que “[...] para o pesquisador, o registro da biblioteca são metadados para descobrir um livro ou artigo de interesse. Para o bibliotecário, o registro pode ser utilizado como dados para analisar a completude relativa das coleções por assunto, por editora, por ano, etc.”.

Considerações Finais



Diante do levantamento e das discussões apresentadas no texto, observa-se que os conceitos dados e metadados apresentam relações de **significados semelhantes e/ou diferentes**, variada pela atribuição de cada autor.

O desenvolvimento tecnológico, atrelado às discussões e formalização da *Web Semântica*, *Ciência dos dados*, *Big data*, entre outros conceitos, **levantaram a importância dos dados e dos metadados, como temática para estudos da área de Ciência da Informação.**

Considerações Finais



Uma melhor conceituação desses termos para o domínio da área de Ciência da Informação é necessária, visto a variedade de definições utilizadas. Isso refletirá em uma fundamentação teórica mais consolidada e maior clareza à pesquisa e ao ensino.

Considera-se um aprofundamento do domínio terminológico para os trabalhos futuros, com a finalidade de **fomentar as discussões sobre a temática e contextualizar e delinear posicionamentos e perspectivas divergentes.**

Referências

- ALVES, R. C. V. Metadados como elementos do processo de catalogação. 2010. 132 f. f. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Marília/SP, 2010. Disponível em: <http://repositorio.unesp.br/handle/11449/103361>. Acesso em: 8 set. 2019.
- ALVES, R. C. V.; SANTOS, P. L. V. A. da C. Metadados no domínio bibliográfico. Rio de Janeiro: Intertexto, 2013.
- ARAKAKI, F. A. Metadados administrativos e a proveniência dos dados: modelo baseado na família PROV. 2019. 139 f. Tese (Doutorado) - Doutorado em Ciência da Informação, Universidade Estadual Paulista “Júlio Mesquita Filho”, Marília, 2019. Disponível em: <https://repositorio.unesp.br/handle/11449/180490>. Acesso em: 8 set. 2019.
- BAKER, T. et al. Library Linked Data Incubator Group Final Report. W3C Incubator Group Report, 2011. Disponível em: <http://www.w3.org/2005/Incubator/llid/XGR-llid-20111025/>. Acesso em: 8 set. 2019.
- FURNER, J. “Data”: The data. In: Kelly M., Bielby J. (eds) Information Cultures in the Digital Age. Springer VS, Wiesbaden, 2016.
- FURNER, J. Definitions of “Metadata”: A Brief Survey of International Standards. Journal of the Association for Information Science and Technology, 2019. doi:10.1002/asi.24295. Acesso em: 8 set. 2019.
- GILLILAND, A. J. Setting the Stage. In: BACA, Murtha (Org.). Introd. Metadata. 3. ed. Los Angeles: Getty Research Institute, 2016. Disponível em: <http://www.getty.edu/publications/intrometadata/>. Acesso em: 8 set. 2019.
- HAYNES, D. Metadata for information management and retrieval. [S.l.]: Facet Publishing, 2004.
- HYVÖNEN, E. Publishing and Using Cultural Heritage Linked Data on the Semantic Web. EUA: Morgan & Claypool Publishers, 2012.
- ISAAC, A. et al. Library Linked Data Incubator Group: Datasets, Value Vocabularies, and Metadata Element Sets: W3C Incubator Group Report 25 October 2011. W3C, 2011. Disponível em: http://www.w3.org/2005/Incubator/llid/XGR-llid-vocabdataset20111025/#Published_Datasets. Acesso em: 8 set. 2019.
- III Workshop de Informação, Dados e Tecnologia, UNB, Brasil
- JEFFERY, K. et al. A 3-Layer Model for Metadata. INTERNATIONAL CONFERENCE ON DUBLIN CORE AND METADATA APPLICATION, 13., Portugal, Anais... DCMI, EUA. 2014. Disponível em: <http://dcevents.dublincore.org/IntConf/dc-2013/paper/view/199/199>. Acesso em: 8 set. 2019.
- JOUDREY, D. N.; TAYLOR, A. G.; WISSER, K. M. The organization of information. 4. ed. Santa Barbara, California: Libraries Unlimited, 2018.
- MÉNDEZ RODRÍGUEZ, E. Metadatos y recuperación de información. Gijón, Asturias: Ediciones Trea, 2002.
- PETERSON, D. (Org.). Forms of representation: an interdisciplinary theme for cognitive science. Wiltshire: Cromwell Press, 1996. 208 p.
- POMERANTZ, J. Metadata. Cambridge, Massachusetts ; London, England: The MIT Press, 2015.
- SANTOS, P. L. V. A. da C.; SANTANA, R. C. G. Dado e Granularidade na perspectiva da Informação e Tecnologia: uma interpretação pela Ciência da Informação. Ciência da Informação, [S.l.], v. 42, n. 2, jan. 2013. ISSN 1518-8353. Disponível em: <http://revista.ibict.br/index.php/ciinf/article/view/228>. Acesso em: 8 set. 2019.
- SANTOS, P. L. V. A. da C.; SIMIONATO, A. C.; ARAKAKI, F. A. Definição de metadados para recursos informacionais: apresentação da metodologia BEAM. Informação & Informação, Londrina, v. 19, n. 1, p. 146-163, fev. 2014. ISSN 1981-8920. Disponível em: <http://repositorio.unesp.br/handle/11449/114736>. Acesso em: 8 set. 2019.
- SANTOS, P. L. A. C.; VIDOTTI, S. A. B. G. Perspectivismo e tecnologias de informação e comunicação: acréscimos à Ciência da Informação. DataGramaZero: revista de Ciência da Informação, Rio de Janeiro, v. 10, n. 3, 2009.
- WICKETT, K. M. et al. Identifying content and levels of representation in scientific data. Proceedings Of The American Society For Information Science And Technology, [s.l.], v. 49, n. 1, p.1-10, 2013. Wiley. <http://dx.doi.org/10.1002/meet.14504901199>. Acesso em: 8 set. 2019.
- ZENG, M. L. QIN, J. Metadata. New York: Neal-Schuman Publishers, 2008.
- ZENG, M. L. QIN, J. Metadata. 2. ed. London: facet publishing, 2016.