

Estudo do conteúdo de arquivos de metadados de Base de Dados científicas em relação a norma de referência bibliográfica ABNT NBR 6083:2018

Isabella C. S. de Araújo^(1,2), Dra. Marli F.G. Hernandez¹, Dra. Marbilia P Sergio²
isabellacs_tec@outlook.com, marli@ft.unicamp.br, marbilia.sergio@cti.gov.br

¹Faculdade de Tecnologia (FT)
UNICAMP - da Universidade Estadual de Campinas

²Núcleo de qualificação de Software (NQS)
CTI/MCTI Renato Archer – Campinas/SP

Resumo

Inúmeras são as bases de dados científicas que permitem a busca por publicações científicas e que disponibilizam funções para baixar arquivos de metadados em várias extensões. Dentre as várias utilidades destes arquivos está a de fornecer informações para a construção de referência bibliográfica do trabalho em desenvolvimento. Está pesquisa se propõem a esclarecer o conteúdo destes arquivos com o objetivo validar o grau de completeza para atender o padrão ABNT de artigos de congressos e de journals. Foram criadas, baixadas e analisadas as diversas extensões oferecidas por quatro bases dos dados científicas e se realizou uma verificação quanto a existência de informações necessárias para atender a norma ABNT NBR 6023:2018. Concluiu-se que a BD WoS é a que disponibiliza informações mais completa em todas as extensões e que os arquivos da ACM atendem apenas ao padrão APA, padrão internacional que exige menor quantidade de informações numa referência. Trabalhos futuros devem acrescentar a validação de outros tipos de conteúdo (publicações) disponibilizados por estas bases de dados e se prevê criar um algoritmo que leia, armazene e gere referências de publicações selecionadas pelo pesquisador, no padrão ABNT.

Palavras chave: metadados; informações bibliográficas; bibliotecas digitais; arquivos de exportação; ABNT NBR 6023.

1 Introdução

Referência bibliográfica de envolve um conjunto de informações de uma publicação ao serem referenciadas pelo autor do trabalho. Através das informações de uma referência pode-se encontrar o trabalho citado, conferindo credibilidade e transparência para uma argumentação de pesquisa.

A disponibilidade da internet ampliou o leque de fontes de revisão bibliográfica assim como novas maneiras de coletar, organizar, acessar e referenciar as publicações identificadas como relevantes para a pesquisa. As bibliotecas digitais (BDs) inovaram com a disponibilização de metadados das publicações em forma de arquivos passíveis de serem baixados tratados por diversos aplicativos que se propõem auxiliar pesquisadores nas revisões literárias, inclusive gerando as referências bibliográficas a serem utilizadas nos relatórios da pesquisa[1].

No entanto, não existe um padrão ou formato único de arquivos de metadados gerados pelas bases de dados [2]. As bases de dados digitais possuem informações variadas das publicações por elas mantidas e apresentam sintaxe própria na sua identificação. Por outro lado, congressos e Journals solicitam formatações próprias na descrição de referências. No Brasil, temos a norma da ABNT NBR 6023:2018 que define padrões para composição de referências em trabalhos científicos

e acadêmicos. Neste contexto, a questão a ser respondida por esta pesquisa é a de identificar e analisar arquivos de metadados disponibilizados por algumas bibliotecas digitais (BDs) e pontuar a aderência destas informações ao padrão nacional.

O restante deste artigo está organizado da seguinte forma: A Seção 2 introduz o conceito de BDs científicas, conceitua metadados e descreve a norma da ABNT sobre referências bibliográficas; A Seção 3 apresenta a metodologia, a organização e conceitos relevantes para o entendimento da pesquisa; A Seção 4 discute o que foi levantado; e, por último, a Seção 5 descreve as lições aprendidas, volta a questão da pesquisa verificando se foi respondida e conclui o trabalho.

2 Bibliotecas digitais e padrão ABNT de referência bibliográfica

Esta seção descreve as bibliotecas digitais (BDs), o que são metadados e apresenta a norma ABNT que padroniza referências bibliográficas.

2.1 Bibliotecas Digitais (BDs)

De acordo com a federação de BDs, uma biblioteca digital (BD) pode ser definida como:

“Organizações que fornecem recursos, para selecionar, estruturar, oferecer acesso intelectual, interpretar, distribuir, preservar a integridade e garantir a persistência ao longo do tempo de coleções de obras digitais para que sejam prontamente e economicamente disponível para uso por uma comunidade definida ou conjunto de comunidades [2]”.

Os elementos de uma BD segundo a Associação de Bibliotecas de Pesquisa [3] são:

- As coleções de bibliotecas digitais não se restringem a substituir documentos físicos; elas também englobam artefatos digitais que não podem ser representados ou distribuídos em formatos impressos.
- As bibliotecas digitais não constituem uma entidade única.
- Criar e manter uma biblioteca digital requer tecnologia integrar de diversas fontes.
- As conexões entre as múltiplas bibliotecas digitais e serviços de informação são invisíveis para os usuários finais.
- Proporcionar acesso universal a bibliotecas digitais e serviços de informação é um objetivo essencial.

AS BDs podem estar focadas em áreas de conhecimento específicos como área médica, área tecnologia, área de engenharia ou área de humanas, dentre outras. Este trabalho definiu quatro bases de dados a serem analisadas. Duas de abrangência genérica, isto é atende a várias áreas do conhecimento e duas com abrangência específica com foco na área tecnológica, a saber:

Scopus: Esta BD científica genérica disponibiliza um banco de dados de resumos e citações de literatura acadêmica vinculada em uma ampla variedade de disciplinas [4] e conta com uma curadoria especializada para inclusão de novas fontes.

Web of Science: Também conhecida como WoS, é uma plataforma online, com foco genérico, suportada pela Clarivate Analytics. Dá acesso a uma ampla coleção de literatura acadêmica. Com uma vasta cobertura de revistas científicas, periódicos, conferências e outras fontes de pesquisa de alta qualidade [5].

IEEE Xplore: A BD IEEE Xplore fornece acesso via Web para documentos de texto completo de publicações relevantes na área de engenharia elétrica, ciência da computação e eletrônica [6].

ACM Digital Library: É uma plataforma disponibilizada pela Association for Computing Machinery (ACM) com foco em computação. Ela disponibiliza o acesso a coleção de todas as

publicações da ACM, incluindo periódicos, anais de conferências, revistas técnicas, boletins informativos e livros [7].

Estas bases de dados foram escolhidas pela sua relevância para a ciência e por disponibilizarem funções para baixar os metadados. As quatro bases de dados são fortemente indicadas, em especial na área tecnológica.

2.2 Metadados

Metadados, literalmente significa “dados sobre dados”, são dados estruturados sobre qualquer coisa que possa ser nomeada, como páginas da Web, livros, artigos de periódicos, imagens, músicas, produtos, processos, pessoas (e suas atividades), dados de pesquisa, conceitos e serviços [8]. Fazendo uma analogia, a um sistema bibliotecário, quando buscamos um livro, podemos ter informações em um catálogo de livro com informação de nome, autor, editora, ano da publicação e assim por diante. Neste exemplo, o metadados dos livros existentes na biblioteca é o catálogo digital em forma de arquivo de metadados.

O padrão Dublin Core é amplamente aceito pela comunidade científica quando se trata de descrever conteúdo de bibliotecas (o que inclui as digitais). Profissionais de diversas instituições especializadas em biblioteconomia e ciência da informação elaboraram o Dublin Core estabelecendo o conjunto de quinze elementos genérico: Criador, Colaborador, Editor, Título, Data, Idioma, Formato, Assunto, Descrição, Identificador, Relação, Fonte, Tipo, Cobertura e Direitos [8]. A instituição DCMI afirma que:

“Dublin Core foi desenvolvido inicialmente para facilitar a descoberta de informações em uma Web em crescimento explosivo, incorporando metadados simples, semelhantes a catálogos de cartões, em suas páginas... prosseguida e aperfeiçoada através de uma série de workshops e conferências para alcançar uma interoperabilidade aproximada entre línguas e disciplinas através de um núcleo de semântica partilhada [8]”.

2.3 Normas ABNT NBR 6023 – Informação e documentação – Referências - Elaboração

No Brasil, os trabalhos acadêmicos devem seguir os padrões da ABNT. Existe um conjunto de normas da ABNT que são indispensáveis para sua correta elaboração. Para referenciar os conteúdos citados no trabalho tem-se a norma ABNT NBR 6023 – Informação e documentação – Referências – Elaboração” para garantir a qualidade e confiabilidade as informações publicadas e proteger os direitos autorais dos autores [9]. Esta norma define as informações que devem estar contidas nos diversos conteúdos de referências variando conforme as características. A tabela 1, apresenta exemplos de referências para alguns dos diferentes tipos de conteúdo.

Tabela 1 – Modelos de referência ABNT NBR 6023:2018 [10]

Conteúdo	EXEMPLO DE REFERÊNCIA
e-book e com DOI	RABELLO, L. S. Promoção da saúde: a construção social de um conceito em perspectiva comparada. Rio de Janeiro: Fiocruz, 2010. E-book. 220 p. DOI 10.7476/9788575413524. Disponível em: http://books.scielo.org/id/z7jxb/epub/rabello-9788575413524.epub . Acesso em: 8 jan. 2019.
Autores Cap. X livro	BACHEGA, K.; ACCETTURI, E. Transplantes de tecido ósseo no Brasil: uma história segura de sucesso da odontologia. In: SANTOS, P. S. S. et al. (org.). Odontologia em transplante de órgãos e tecidos. Curitiba: Editora CRV, 2018. cap. 7, p. 109-127.
tese	RIZZO, M. Stability of first and second premolars extraction space closure. 2018. 79 p. Tese (Doutorado em Ortodontia) - Faculdade de Odontologia de Bauru, Universidade de São Paulo, Bauru, 2018.
Dissertação com DOI	CABALLERO, J. T. Comparação da estabilidade dos arcos dentários em pacientes com e sem fissura labiopalatina após tratamento ortodôntico/reabilitador. 2018. 62 p. Dissertação (Mestrado em Reabilitação Oral) - Faculdade de Odontologia de Bauru, Universidade de São Paulo, 2018. DOI 10.11606/D.25.2018.tde-31082018-174817. Disponível em: http://www.teses.usp.br/teses/disponiveis/25/25146/tde-31082018-174817/pt-br.php . Acesso em: 8 jan. 2019
Artigos eletrônico	TOMAZELA, J. M. Duas pessoas morrem com sintomas de febre maculosa no interior de SP. Estadão, São Paulo, 03 jan. 2019. Disponível em: https://saude.estadao.com.br/noticias/geral,duas-pessoas-morrem-com-sintomas-de-febre-maculosa-no-interior-de-sp,70002666449 . Acesso em: 9 jan. 2019.
trabalho anais e resumos publicados em revistas	OYADOMARI, A. T. et al. Efeitos da terapia por laser de baixa potência no processo de reparo de defeitos ósseos preenchidos pelo osso bovino Bio-Oss® associados ao novo selante heterólogo de fibrina. In: SIMPÓSIO INTERNACIONAL DE INICIAÇÃO CIENTÍFICA DA UNIVERSIDADE DE SÃO PAULO, 25., 2017, Bauru. Resumos [...]. São Paulo: Universidade de São Paulo, 2017.
Normas técnicas	ABNT - Associação Brasileira de Normas Técnicas. ABNT NBR ISO 12836: odontologia: dispositivos para sistemas CAD/CAM para restaurações dentárias indiretas: métodos de ensaio para avaliação de exatidão. Rio de Janeiro: ABNT, 2017. 14 p.

Leis	BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Dispõe sobre a proteção de dados pessoais e altera a Lei nº 12.965, de 23 de abril de 2014 (Marco Civil da Internet). Diário Oficial da União: seção 1, Brasília, DF, ano 155, n. 157, p. 59-64, 15 ago. 2018.
VÍDEO youtube	OSIER, F. The key to a better malaria vaccine. [New York]: TED, 2018. 1 vídeo (7 min). Disponível em: https://www.youtube.com/watch?v=b2Jv8vC-m3g . Acesso em: 6 jan. 2019
TWITTER	ANVISA – Agência Nacional de Vigilância Sanitária. Uma nova promessa terapêutica em situações clínicas complexas e sem alternativas disponíveis começará a ter seus estudos desenvolvidos no Brasil [...]. Brasília, DF, 05 jan. 2019. Twitter: @anvisa_oficial. Disponível em: https://twitter.com/anvisa_oficial/status/1081597493614465025 . Acesso em: 6 jan. 2019.
BLOG	PACKER, A. L. et al. SciELO pós 20 anos: o futuro continua aberto. In: SCIELO - Scientific Electronic Library Online. SciELO em Perspectiva. São Paulo, 19 dez. 2018. Disponível em: https://blog.scielo.org/blog/2018/12/19/scielo-pos-20-anos-o-futuro-continua-aberto/#.XDYD31xKiUk . Acesso em: 9 jan. 2019.
PODCAST	DRAGÕES DE GARAGEM #137: vó Maria: vacinas e escolhas #semanadavacina. [Locução de]: Barbara Paes. [S.l.]: Dragões de Garagem, 14 ago. 2018. Podcast. Disponível em: http://dragoesdegaragem.com/podcast/dragoes-de-garagem-137-vo-maria-vacinas-e-escolhassemanadavacina/ . Acesso em: 9 jan. 2019.
Páginas na WEB	CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Plataforma Sucupira. Brasília, DF: CAPES, c2016. Disponível em: https://sucupira.capes.gov.br/sucupira/public/index.xhtml . Acesso em: 20 mar. 2019.

Os tipos de conteúdo podem ser informações de livros, trabalhos acadêmicos (tese, dissertação ou TCC), artigos (congresso ou Journal), Jornal, patente, bula de remédio, normas técnicas, referências legislativas (constituição, leis federais, estaduais ou municipais, resoluções, portarias), documento iconográfico (figura, cartazes e desenhos), vídeos (Youtube, figshare), twitter, facebook, enciclopédia/dicionário, wikipédia, Blog, Podcast, e-mail, base de dados e página da Web.

Embora espera-se de um arquivo de metadados traga informações para geração de referências bibliográficas, deve se reconhecer que referências de conteúdos extraídos da internet não serão contemplados em metadados extraídos destas fontes. Este trabalho está focado na análise do conteúdo dos metadados em suas diversas extensões no que se refere ao atendimento a norma ABNT para referenciar artigos de congresso e Journal. A figura 1 mostra um exemplo do padrão ABNT para referencia-los.

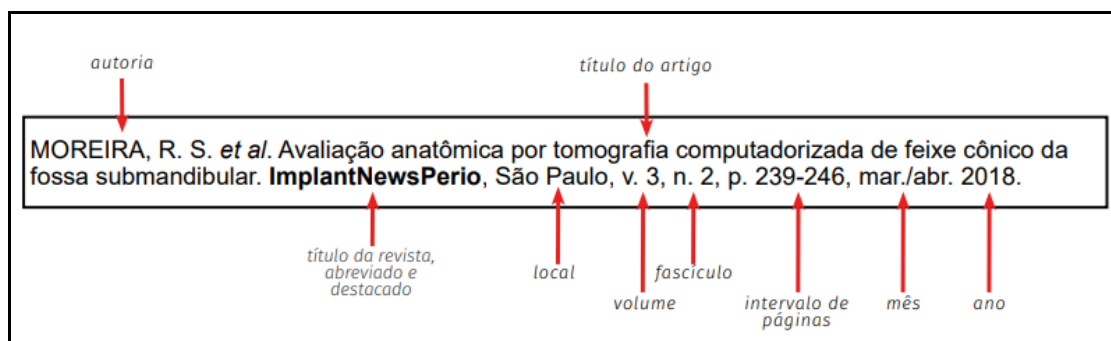


Figura 1 - exemplo de como referenciar artigos de periódico no padrão ABNT

2.4 Padrão APA – American Psychological Association

Conhecido pela sigla APA, o padrão **American Psychological Association** – ou padrão simples – é um estilo de formatação de referências muito utilizado. As referências de artigos de periódicos são abordadas na sétima edição dos manuais de estilo APA [11]. Na figura 2, temos um exemplo de referência APA retirado da web *site* da organização APA que coordena a norma [11].

Grady, J. S., Her, M., Moreno, G., Perez, C., & Yelinek, J. (2019). Emotions in storybooks: A comparison of storybooks that represent ethnic and racial groups in the United States. *Psychology of Popular Media Culture*, 8(3), 207–217. <https://doi.org/10.1037/ppm0000185>

Figura 2 - Exemplo estrutura de referência APA

Em resumo, nas publicações estudadas, os itens que temos apresentados no formato APA são: título, ano, editora, volume, edição, páginas, número de citações, DOI e URL.

3 Metodologia, Organização da pesquisa e levantamento realizados

Nessa seção, descreve-se as atividades desenvolvidas na pesquisa, sua organização e analisa por BD as informações diferentes extensões de metadados oferecidas por cada uma delas.

3.1 Metodologia e Organização da pesquisa

Foi adotado o método um estudo de caso onde são analisados o conteúdo dos arquivos de metadados estabelecendo como e quais informações são comuns e quanto ao atendimento ao padrão da ABNT no que se refere a conteúdo de artigos científicos.

Nesta pesquisa foram realizadas as seguintes atividades:

- Levantamento de padrões da ABNT para referência bibliográfica;
- Identificação das BDs científicas e seleção das quatro delas para análise de seus arquivos com metadados;
- Geração das extensões de arquivos metadados disponibilizados para download;
- Análise do conteúdo das diversas extensões de cada BD comparando a completude entre elas.
- Análise da adequação do conteúdo dos metadados das extensões mais completas de cada BD quanto ao atendimento aos requisitos da ABNT para artigos;
- Descrever os achados e redigir o artigo.

3.2 Levantamento - gerar e analisar arquivos de metadados

Para análise conteúdo dos arquivos de metadados, foi realizada uma busca nas BDS Scopus, Web of Science, IEEE Xplore e ACM Digital Library para selecionar artigos de congresso e de Journals para baixar seus metadados nas diferentes extensões.

3.3. Os tipos de extensões de arquivos de metadados disponibilizados pelas BDs

Cada BD disponibiliza alguns tipos de formatos de arquivos. Na tabela 2 encontra-se nove extensões de arquivos encontrados nas quatro BDs estudadas.

Tabela 2 – Extensões de arquivos de metadados por BDs

extensão	Web of Science	Scopus	IEEE Xplore	ACM Digital Library
BibTeX	X	X	X	X
RIS	X	X	X	
Plain Text	X	X	X	
CSV	X	X		
EndNote	X			X
Refworks	X		X	
HTML	X			
ACM Ref				X
Tab-Delimited	X			

A BD que mais oferecem arquivos de metadados é a WoS com oito extensões diferentes. Os três formatos de maior incidência entre as BDs são *BibTex*, *RIS* e *Plain Text*.

O entendimento do conteúdo dos arquivos de cada BD envolveu a análise dos campos de cada artigo contido nas extensões, assim como sobre o foco da extensão.

A extensão **BibTex** se destacou por estar disponível em quase todas as BDs, não sendo disponibilizada apenas pela ACM Ref. Há uma tendência de que novas publicações sejam construídas no Latex, o qual utiliza esta extensão para receber e incluir referências num trabalho. O formato **RIS** se baseia na norma ANSI/NISO Z39.29-2005, um padrão internacional também conhecido como "*ANSI/NISO Bibliographic References*". Já o **Plain Text**, por ser um arquivo sem

formatação, traz a propriedade de ser compatível com uma variedade de aplicativos podendo ser aberto em praticamente qualquer editor de texto.

O **CSV** tem um formato de tabela e aderente, por exemplo a planilha Excel. O formato **EndNote** tem sua relevância dada sua compatibilidade com o software de gerenciamento de referências de mesmo nome, o EndNote.

As demais extensões não foram analisadas neste trabalho por motivos diversos. O conteúdo do formato **RefWorks**, disponibilizado pelas bibliotecas Web of Science e IEEE Xplore, só é acessível por meio da autenticação do software pago RefWorks¹, exigindo o uso do software para ter acesso aos seus campos. Os formatos **HTML**, **ACM Ref** e **Tab-Delimited** são disponibilizados por uma única BD, impossibilitando comparações deste estudo.

3.4. Informações contidas nas extensões

Todos os arquivos de metadados baixados continham informações de três artigos. As diferentes extensões de arquivo apresentam diferentes sintaxes para as informações de artigos. Mesmo na análise do conteúdo de cada arquivo identificou-se que entre os artigos de uma mesma BD inexistia um padrão.

A seguir são apresentadas por extensão as análises dos metadados por BDs. Analise comparando as informações entre as BDs quanto atenderem a norma da ABNT encontra-se no item 5 Discussão. Assim como se os campos observados apresentam ou não as informações necessárias e uniformes.

3.4.1 BibTeX

De acordo com o site organizacional do formato, arquivos BibTeX contém uma lista de referências que são comumente usadas em documentos acadêmicos. A figura 2 apresenta um exemplo de um arquivo BibTeX.

```
@article{
  Author = {foo},
  Title = {titleBIBTEXing},
  Journal = {journalBIBTEXing},
  ...
}
```

Figura 3- Exemplo estrutura arquivo BibTeX

Tem-se que:

O BibTeX introduzido por Oren Patashnik há 20 anos, tem sido amplamente utilizado. Como o nome sugere, foi planejado para ser usado em combinação com o sistema de composição LaTeX², mas tornou-se possível, por exemplo, incluir bibliografias BibTeX em documentos do Word usando ferramentas de terceiros. Embora existam ferramentas que fornecem uma interface mais conveniente para tratar arquivos BibTeX, seu formato de texto simples permite que seja criado ou modificado utilizando outros editores de texto. [12]

Quanto a padronização dos campos da extensão BibTeX de cada BD, observou-se que:

- Dos vinte e três campos encontrados no arquivo BibTeX, gerado pelo **Scopus**, apenas vinte campos são comuns entre os artigos;

¹ RefWorks é um pacote de software de gerenciamento de referência comercial baseado na web que armazena e cita suas fontes.

² O LaTeX é um sistema de composição tipográfica de qualidade, que inclui recursos destinados à produção de documentos técnicos e científicos.

- De trinta e nove campos encontrados no arquivo BibTeX, gerado pelo **Web of Science**, apenas trinta e cinco são comuns entre os artigos;
- Os treze campos encontrados no arquivo BibTeX, gerado pelo **IEEE Xplore**, são comuns entre os artigos;
- De sete campos encontrados no arquivo BibTeX, gerado pelo **ACM Digital**, seis campos são comuns entre os artigos;

No arquivo BiBtex da biblioteca Scopus apresenta divergência em três campos, pois não existiam em todos metadados dos artigos.

No arquivo BiBtex da biblioteca WoS destaca-se pela quantidade de campos presentes. Além dos campos típicos de um artigo, outros dados são disponibilizados.

No arquivo BiBtex da biblioteca IEEE Xplore encontram-se informações mais assertivas se comparadas as bibliotecas Scopus e WoS.

Por último, a ACM Digital Library apresenta a menor quantidade de campos. O foco da biblioteca é permitir que, com os dados encontrados nos metadados disponíveis, os pesquisadores possam criar uma referência de padrão simples.

3.4.2 RIS

RIS (*Research Information Systems*) utiliza a sintaxe da norma ANSI/NISO Z39.29-2005 - *Bibliographic References*", um padrão internacional. Estabelece um conjunto de códigos de dois caracteres para identificar as diferentes informações bibliográficas. A Tabela 3 expõem todas as siglas e seus significados encontrados nos metadados de todas as três BDs.

Tabela 3 – Tabela de siglas ANSI/NISO Bibliographic References

SIGLAS ANSI/NISO Bibliographic References					
TY	Tipo	UR	URL	Y1	Ano primeira publicação
AU	Autor	AD	Address	M3	Miscellaneous 3
TI	Título	AB	Abstract	DB	Database
T2	Título secundário	KW	Keywords	FU	Núm. suplemento da revista
C2	Qt. de citações	PB	Publisher	FX	Texto de referência livre
C3	Afiliações	SN	Serial Number	CR	Referências citadas
C7	identificação publicação	LA	Idioma	PA	Local
PY	Ano publicação	JA	Nome do jornal	PI	Editora
VL	Volume	JI	Nome jornal em itálico	PU	Nome da editora
IS	Issue	JO	Periódico	DA	Data da publicação
SP	Start Page	N1	Notes 1	WE	Endereço na web
EP	End Page	J2	Secondary Journal Title	AN	Núm. de acesso do artigo
DO	DOI	J9	Nome revista abreviado	ER	End of Record

Vale salientar que algumas dessas siglas são personalizadas pelas BDs, isto é, siglas de campos podendo variar de acordo com as BDs. Outra observação importante é em relação ao campo "ER", comum a todas as BDs e que indica o fim de um registro bibliográfico.

Entre os metadados das bibliotecas SCOPUS, WoS e IEEE Xplore, tem-se entre os trinta e nove campos encontrados, apenas treze comuns, são eles: TY, AU, TI, T2, PY, VL, IS, DO, KW, SN, SP, EP e AB. Cada biblioteca também possui siglas personalizadas que podem conter informações específicas ou diferentes siglas com as mesmas informações. Da análise, salienta-se que nem todos os campos são presentes nas três publicações estudadas de cada BDs.

3.4.4 Plain Text

O termo, traduzindo para o português, significa formato "Texto Simples", diz respeito a um tipo de arquivo que contém apenas texto sem formatação especial. Esse formato traz uma representação direta do metadados de um artigo, sem incluir elementos como fontes, tamanhos, cores, entre outros atributos de formatação. Devido ao seu estilo de arquivo ser sem formatação,

ele é compatível com uma grande variedade de aplicativos e sistemas operacionais, podendo ser aberto em praticamente qualquer editor de texto.

Por não possuir uma estrutura própria, cada BD estabelece um tipo de tag para representar os campos dos metadados. No levantamento dos campos dos arquivos de extensão Plain Text, observou-se:

- As BDs Scopus e IEEE Xplore possuem, além dos campos que informam os metadados, um trecho que contém uma referência no padrão American Psychological Association (APA);
- A BD WoS aderiu as siglas ANSI/ISO *Bibliographic Reference* observada no formato RIS e possui os mesmos campos que o formato EndNote (olhar o tópico 3.4.5 EndNote);
- A BD Scopus, além da referência em padrão APA, possui vinte campos no arquivo Plain Text. Destes, apenas quatorze são comuns entre os artigos do mesmo arquivo.
- A BD IEEE Xplore disponibiliza no arquivo Plain Text, além da referência em padrão APA, mais três campos. Todos os campos são comuns entre os arquivos.

3.4.3 CSV

CSV (*Comma-Separated Values*) é um formato de arquivo utilizado para armazenar e trocar dados tabulares de forma simples e eficiente. Neste formato, os valores são separados por vírgulas. A vírgula, o ponto e vírgula e a tabulação são tipos de pontuação usados como delimitadores. Devido sua estrutura simples, além da possibilidade de leitura e análise por vários softwares, o CSV pode ser utilizado para armazenar referências bibliográfica de forma efetiva.

Na análise dos arquivos das bibliotecas Scopus e WoS, únicas bibliotecas que exportam arquivos com extensão CSV, tem-se que este formato exige padronização de campos, ou seja, todos os metadados conterão os mesmos campos. Se a informação não existir, um valor vazio entre as vírgulas reservada o local do conteúdo. Enquanto a Scopus possui quarenta e seis campos de metadados, a Web of Science oferece um range de setenta e dois campos para armazenar informações.

3.4.5 EndNote

EndNote é uma extensão utilizada em documentos acadêmicos e outros tipos de trabalhos escritos. Esse tipo de arquivo é compatível com o software de gerenciamento de referências de mesmo nome, EndNote. Tal software permite que a organização, o armazenamento e a formatação das citações e referências, facilitando a inclusão dessas informações em seus trabalhos acadêmicos, como artigos, teses e dissertações.

Para as BDs estudadas, apenas as bibliotecas Web of Science e ACM Digital Library exportam arquivos com esta extensão. A Web of Science exporta este arquivo com extensão “ciw” e a ACM usa a extensão “enw”. Ambas as extensões podem ser importadas para a plataforma End Note. A tabela 4 apresenta as siglas do padrão ANSI/NISO geradas pelo WoS em arquivo com extensão lida pelo EndNote.

Tabela 4 – Tabela de siglas ANSI/NISO Bibliographic References

	Web of Science End Note												
Technology	FN	VR	PT	AU	AF	TI	SO	LA	DT	ID	C1	C3	RP
	EM	RI	OI	FU	FX	CR	NR	TC	Z9	U1	U2	PU	PI
	PA	SN	J9	JI	PD	PY	VL	IS	BP	EP	DI	PG	WC
	WE	SC	GA	UT	PM	OA	ER	DA	EF	EI	HC	HP	AB
Cloud	FN	VR	PT	AU	AF	TI	SO	LA	DT	ID	C1	C3	RP
	EM	RI	OI	FU	FX	CR	NR	TC	Z9	U1	U2	PU	PI

	PA	SN	J9	JI	PD	PY	VL	IS	BP	EP	DI	PG	WC
	WE	SC	GA	UT	PM	OA	ER	DA	EF	EI			
Internet	FN	VR	PT	AU	AF	TI	SO	LA	DT	ID	C1	C3	RP
	EM	RI	OI	FU	FX	CR	NR	TC	Z9	U1	U2	PU	PI
	PA	SN	J9	JI	PD	PY	VL	IS	BP	EP	DI	PG	WC
	WE	SC	GA	UT	PM	OA	ER	DA	EF	HC	HP	AB	

Em relação aos dados encontrados, Web of Science apresenta o mesmo conteúdo encontrado no formato Plain Text de todas as publicações analisadas. A estrutura segue usando a norma *ANSI/ISO Bibliographic References*. Na tabela 5, são relacionados os códigos encontrados no arquivo de metadados da ACM.

Tabela 5 – Sumário ACM Digital Library EndNote

Sigla	Descrição
%O	Tipo
%T	Título
%@	Código Identificação
%I	Editor
%D	Ano
%C	Local

Os arquivos EndNote disponibilizados da ACM, tem códigos de campo que representam partes diferentes de uma referência bibliográfica e que são aceitas pelo software EndNote. Os códigos podem ser observados na tabela 5.

5 Discussão

Os resultados obtidos traçam um panorama das observações e análises das bibliotecas digitais estudadas nesta pesquisa, isto é, a bibliotecas Scopus, a Web of Science, a IEEE Xplore e a ACM Digital Library. Durante a análise, a questão levantada foi sobre a completeza dos metadados para descrever uma referência bibliográfica. Levando em consideração que as normas de citação podem variar, foi considerado a norma ABNT NBR 6023:2018 para referenciar conteúdos publicados em congressos e journals. A tabela 6 relaciona as descrições de uma referência de artigo de Journals e de conferência.

Tabela 6 – Informações obrigatórias segundo norma ABNT NBR 6023:2018 [10]

Tipo	Descrição	Tipo	Descrição
Registros de Conferência	<ul style="list-style-type: none"> • Autor(es) da apresentação • Título da apresentação. • In: Título dos "Proceedings" • Nome do evento • Local do evento • Ano do evento • Pág da apresentação nos "Proceedings" 	Artigo	<ul style="list-style-type: none"> • Autor(es) • Título do artigo • Título da revista • Local de publicação da revista • Volume da revista (número) • Página inicial e final do artigo • Ano de publicação

Os itens dessa tabela têm-se os dados que é preciso para se formatar uma referência segundo a norma brasileira.

As bibliotecas Scopus e Web of Science são as únicas que disponibilizam todas as extensões estudadas sendo: BibTeX, RIS, CSV, Plain Text e EndNote. Estes arquivos de metadados contém todos os dados necessários para se formar uma referência bibliográfica nas normas ABNT. De fato, essas duas bibliotecas possuem uma quantidade dados considerável e se apresentam como as bibliotecas mais completas.

Já a biblioteca ACM Digital Library, disponibiliza o formato BibTex e EndNote, constatou-se falta de informação para o formato brasileiro. Por se tratarem de títulos de tipo registros de conferência, ficaram faltando os dados dos nomes dos autores, título dos "Proceedings" e as

páginas da apresentação. Essa carência de informações pode ser atribuída ao fato de que, para as publicações analisadas, a BD visa atender a norma de referências bibliográficas APA cujo formato é simplificado.

Outro indicativo notado durante as análises foi a contemplação de algumas das propostas de como deve ser a apresentação dos metadados de acordo com o padrão Dublin Core. A flexibilidade, interoperabilidade semântica e simplicidade dos elementos encontrados permitiu à capacidade de interpretação dos metadados sem grandes dificuldades.

Destacamos que a biblioteca Web of Science, em alguns tipos de documentos, conseguiu aderir as quinze categorias de metadados propostas pelo padrão Dublin Core.

6 Conclusão

Esse artigo apresentou um estudo abrangente das informações presentes nos arquivos de metadados exportados pelas bibliotecas digitais Scopus, Web of Science, IEEE Xplore e ACM Digital Library e permitem a elaboração de referência bibliográfica de um artigo. As observações e análise dos metadados encontrados permitiu um entendimento da relação entre os tipos de extensões disponibilizados por cada biblioteca.

Além disso, o estudo investigou a viabilidade de se estabelecer paralelos entre as informações contidas nos metadados e as referências a serem criadas no padrão brasileiro ABNT apontando outros padrões igualmente relevantes para a ciência. Ao conectar essas abordagens, o artigo buscou identificar potenciais sinergias e divergências entre os sistemas de referência contribuindo para uma compreensão das diversidades e interoperabilidade de referências bibliográficas.

O próximo passo para esse estudo é ampliar a análise para outros tipos de conteúdo a serem referenciados, como livro, por exemplo. Assim como, a criação de um algoritmo com a capacidade de ler os arquivos de metadados a automatizar a construção de uma referência em formato ABNT, ou até mesmo em outros formatos. Além, de ampliar a possibilidade de formatação de outros conteúdos.

6 Agradecimentos

Agradecemos também ao CNPq e ao CTI a oportunidade de participar do Programa PIBIC.

Referências

- [1] KITCHENHAM, Barbara. 2004. Procedures for performing systematic reviews. Keele, UK, Keele University, v. 33, n. p. 1-26, 2004.
- [2] IFLA (Federação Internacional de Associações e Instituições de Bibliotecas). Bibliotecas Digitais: Recursos de Metadados. Disponível em: <http://www.ifla.org/II/metadata.htm>. Acesso em: 17 jul. 2023.
- [3] Associação de Bibliotecas de Pesquisa. Definição e finalidades de uma biblioteca digital. 1995. Disponível em: <http://www.arl.org/sunsite/definition.html>. Acesso em: 17 jul. 2023.
- [4] Elsevier. (2021). Scopus. Disponível em: <https://www.scopus.com/>. Acesso em: 17 jul. 2023.
- [5] Clarivate Analytics. Web of Science. Disponível em: <http://www.webofscience.com/>. Acesso em: 17 jul. 2023.
- [6] Institute of Electrical and Electronics Engineers (IEEE). IEEE Xplore. Disponível em: <https://ieeexplore.ieee.org/>. Acesso em: 17 jul. 2023.

- [7] Association for Computing Machinery (ACM). ACM Digital Library. Disponível em: <https://dl.acm.org/>. Acesso em: 17 jul. 2023.
- [8] Dublin Core Metadata Initiative. Dublin Core Specifications. Disponível em: <https://www.dublincore.org/specifications/dublin-core/>. Acesso em: 17 jul. 2023.
- [9] ABNT - Associação Brasileira de Normas Técnicas. Sistema Integrado de Bibliotecas UFSCar. Disponível em: <https://www.sibi.ufscar.br/>. Acesso em: 17 jul. 2023.
- [10] ABNT. NBR 6023:2018 - Informação e documentação - Referências - Elaboração. Rio de Janeiro, 2002.
- [11] APA (American Psychological Association). Página Inicial da American Psychological Association. Disponível em: <https://www.apa.org/>. Acesso em: 19 de julho de 2023.
- [12] BibTeX.org. Sobre o BibTeX. Disponível em: <https://www.bibtex.org/About/>. Acesso em: 19 de julho de 2023.