



Encontros Bibli: revista eletrônica de biblioteconomia e
ciência da informação

ISSN: 1518-2924

Universidade Federal de Santa Catarina

Rautenberg, Sandro; Souza, Lucélia de; Hild, Tony Alexander
CONJUNTOS DE DADOS PARA PESQUISAS CIENTOMÉTRICAS:
RECUPERANDO QUALIS, SJR, SNIP A PARTIR DA WEB DE DADOS

Encontros Bibli: revista eletrônica de biblioteconomia
e ciência da informação, vol. 26, e80155, 2021

Universidade Federal de Santa Catarina

DOI: <https://doi.org/10.5007/1518-2924.2021.e80155>

Disponível em: <https://www.redalyc.org/articulo.oa?id=14768130026>

- Como citar este artigo
- Número completo
- Mais informações do artigo
- Site da revista em redalyc.org

UAEM redalyc.org

Sistema de Informação Científica Redalyc

Rede de Revistas Científicas da América Latina e do Caribe, Espanha e Portugal

Sem fins lucrativos acadêmica projeto, desenvolvido no âmbito da iniciativa
acesso aberto



Encontros Bibli

CONJUNTOS DE DADOS PARA PESQUISAS CIENTOMÉTRICAS: RECUPERANDO QUALIS, SJR, SNIP A PARTIR DA WEB DE DADOS

Datasets for Scientometric Research: Retrieving QUALIS, SJR, SNIP from the Web of Data

Sandro Rautenberg

Doutor em Engenharia e Gestão do Conhecimento
Universidade Estadual do Centro-Oeste
Guarapuava, PR, Brasil
srautenberg@unicentro.br

<https://orcid.org/0000-0002-2375-9365> 

Lucélia de Souza

Doutora em Ciência da Computação
Universidade Estadual do Centro-Oeste
Guarapuava, PR, Brasil
lucelia@unicentro.br

<https://orcid.org/0000-0002-9190-7411> 

Tony Alexander Hild

Mestre em Informática
Universidade Estadual do Centro-Oeste
Guarapuava, PR, Brasil
thild@unicentro.br

<http://orcid.org/0000-0002-0369-1842> 

A lista completa com informações dos autores está no final do artigo ●

RESUMO

Discorre sobre a publicação de conjuntos de dados cientométricos como dados abertos conectados. Os dados originais são extraídos de fontes primárias via navegadores de Internet e tratados de acordo com a Curadoria Digital de Dados e o Ciclo de Vida de Dados Conectados. Os dados referem-se ao histórico dos índices Qualis, *SCImago Journal Rank* e *Source Normalized Impact per Paper*, os quais são integrados para privilegiar acesso, reúso, interoperabilidade e processabilidade de recursos para com outras pesquisas bibliométricas/cientométricas. A integração dos dados decorre da evolução do trabalho “RAUTENBERG, S.; HILD, T. A.; SOUZA, L. de. Curadoria Digital de Dados e Web de Dados: mantendo Dados Abertos Conectados para estudos bibliométricos e cientométricos. Em *Questão*, v. 24, p. 29-47, Edição Especial 6 EBBC, 2018”, disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/86519>, DOI: <http://dx.doi.org/10.19132/1808-5245240.29-47>. Dinamicamente, os conjuntos de dados resultantes são compartilhados na Web de Dados, via um servidor OpenLink Virtuoso acessado no endpoint <http://lod.unicentro.br>.

PALAVRAS-CHAVE: Preservação de Dados. Dados Abertos Conectados. Cientometria. Bases de Dados Científicos. Web de Dados.

ABSTRACT

It discusses the publication of scientometric datasets as linked open data. The original data are extracted from primary sources via Internet browsers and processed by involving practices of the Digital Data Curation and Linked Data Lifecycle. The data refer to the Qualis, *SCImago Journal Rank* and *Source Normalized Impact per Paper* historical indexes, which are integrated to promote access, reuse, interoperability and processability of resources with other bibliometric or scientometric researches. The integration of those datasets evolves a prior paper named “RAUTENBERG, S.; HILD, T. A.; SOUZA, L. de. Curadoria Digital de dados e Web de Dados: mantendo Dados Abertos Conectados para estudos bibliométricos e cientométricos. Em *Questão*, v. 24, p. 29-47, Edição Especial 6 EBBC, 2018”, available at: <https://seer.ufrgs.br/EmQuestao/article/view/86519>, DOI: <http://dx.doi.org/10.19132/1808-5245240.29-47>. The datasets are maintained on the Web of Data, by an OpenLink Virtuoso server at the endpoint <http://lod.unicentro.br>.

KEYWORDS: Data Preservation. Linked Open Data. Scientometrics. Scientific Datasets. Web of Data.

1 APRESENTAÇÃO

Os avanços da Internet permitem o surgimento de novas formas de produção e exploração de dados, informações e conhecimento. Essa realidade sustenta o movimento

de liberdade de distribuição e uso de recursos digitais, despontando na Web de Dados a plataforma global para publicar dados abertos conectados (AUER, 2014). Resumidamente, a Web de Dados possibilita a exploração de dados abertos conectados, permitindo o reuso, a compreensão, a interligação, a descoberta, a confiança, o acesso, a interoperabilidade e a processabilidade de recursos digitais (W3C, 2020a). Ressalta-se que o acolhimento desses benefícios ocorre com a implantação de *endpoints* (W3C, 2020b), os quais se caracterizam como um serviço web que implementa um protocolo específico de consulta para recuperar recursos digitais em formatos específicos (como por exemplo, CSV¹, TSV², XLS³, RDF⁴, entre outros).

Diante esse contexto, apresenta-se o *endpoint* <http://lod.unicentro.br/sparql> como um ponto de acesso da Web de Dados e a evolução de seus conjuntos de dados abertos conectados para pesquisas do domínio da Ciência da Informação. Especificamente, como pergunta de pesquisa, este artigo de dados explicita “Como recuperar simultaneamente os conjuntos de dados Qualis, SJR⁵ e SNIP⁶ para promover pesquisas bibliométricas e/ou cientométricas?” Ao atender a essa questão, tal esforço permite:

1. a recuperação agregada dos índices compartilhados, uma vez que a referida integração pode se tornar importante na condução de pesquisas relevantes;
2. a colaboração com a comunidade científica em estudos bibliométricos e/ou cientométricos, ao minimizar os esforços para obter e recorrentemente criticar os recursos digitais; e
3. o acesso aberto e perene de pesquisadores aos recursos digitais compartilhados, permitindo a exploração desses recursos digitais e até sustentar possíveis relacionamentos destes com outros conjuntos de dados da Ciência da Informação.

2 MÉTODOS E INSTRUMENTOS

Neste trabalho são considerados os conjuntos de dados dos estratos e índices cientométricos:

- a) **Qualis** – coletado nos últimos 15 anos, principalmente, a partir da Plataforma Sucupira (SUCUPIRA, 2020);

¹ *Comma Separated Values*

² *Tab Separated Values*

³ *Microsoft eXcel Spreadsheet*

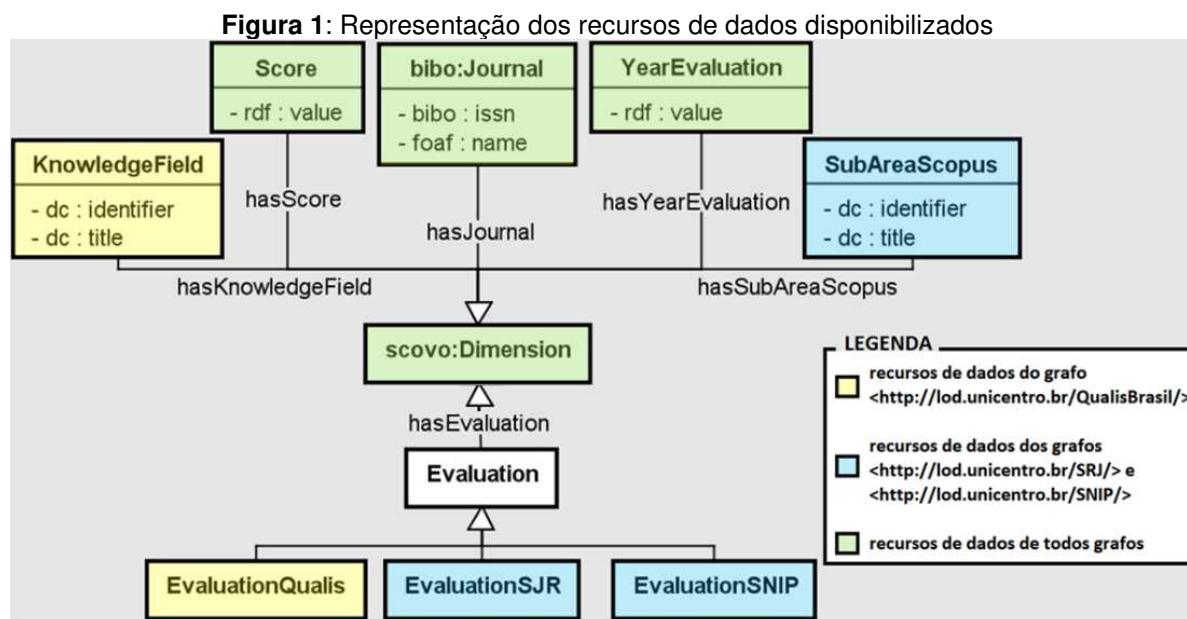
⁴ *Resource Description Framework*

⁵ *SCImago Journal Rank*

⁶ *Source Normalized Impact per Paper*

- b) **SJR** – coletado do Portal *Journal SCImago & Country Rank* (SJR, 2020); e
c) **SNIP** - recuperado no Portal *Journal Metrics* (JOURNAL METRICS, 2020);

Ressalta-se que os conjuntos de dados enumerados anteriormente são recuperados anualmente a partir de suas fontes originais e armazenados na base de dados de um sistema legado. Posteriormente, é realizada a socialização dos índices cientométricos na Web de Dados, como dados abertos conectados. Na Figura 1 é ilustrado como os recursos de dados são compartilhados.



Fonte: adaptado de (RAUTENBERG; HILD; SOUZA, 2018).

Dada a riqueza de detalhes, os apontamentos específicos sobre a modelagem de dados dos índices cientométricos e seu processo técnico-metodológico de transformação para dados abertos conectados são apresentados em Rautenberg et. al. (2017a) e Rautenberg et. al. (2017b), respectivamente.

Conforme o trabalho seminal (RAUTENBERG; HILD; SOUZA, 2018), para custodiar os conjuntos de dados abertos conectados Qualis, SJR e SNIP, são aplicadas cinco atividades da Curadoria Digital de Dados (DIGITAL CURATION CENTER, 2020). Representadas na Figura 2, interdisciplinarmente, essas atividades são mediadas por quatro fases do Ciclo de Vida de Dados Conectados (AUER, 2014), conforme descritas na sequência.

Figura 2: Representação do procedimento metodológico



Fonte: adaptado de (RAUTENBERG; HILD; SOUZA, 2018).

As atividades de Seleção e Ingestão são realizadas simultaneamente durante a fase de Extração do Ciclo de Vida de Dados Conectados. A Seleção visa recuperar os dados de interesse de fontes originais para preservação a longo prazo. Com os dados recuperados, a Ingestão transfere os dados para uma área de armazenamento, um sistema legado.

A atividade de Transformação é o ato de criar dados, como por exemplo, ao migrar os dados originais para outro formato ou enriquecer os dados com semântica. No procedimento metodológico, os dados são recuperados do sistema legado, em formato CSV. Posteriormente, os dados são convertidos para RDF, utilizando os vocabulários ou ontologias para enriquecer a semântica dos recursos a serem compartilhados.

Na Curadoria Digital, a atividade de Armazenamento dos recursos digitais deve ser realizada de forma segura, respeitando os padrões relevantes. Neste sentido, para privilegiar o reúso, os recursos digitais resultantes da atividade de Transformação são armazenados em um *endpoint* da Web de Dados.

A atividade de Acesso, Uso e Reúso engloba as ações para garantir que os recursos digitais sejam acessíveis pelos consumidores de dados de forma ininterrupta. No Ciclo de Vida de Dados Conectados, essa atividade é relacionada à fase de Busca, Navegação e Exploração para recuperar e visualizar os recursos digitais compartilhados.

3 TABELA DE ESPECIFICAÇÕES

Área de Conhecimento	IK. Projeto, desenvolvimento, implementação e manutenção
Área de assunto específica	preservação de recursos digitais na Web de Dados como dados abertos conectados para estudos bibliométricos/cientométricos
Idioma	Multilíngue
Tipo de dados	Texto em TSV

Como os dados foram adquiridos	a) a extração de dados brutos ocorreu em sítios especializados: (SUCUPIRA, 2020); (SJR, 2020); e (JOURNAL METRICS, 2020) via navegadores de Internet; b) o pré-processamento foi realizado com o uso de <i>scripts</i> PHP para inserir os dados em uma base de dados mantida em uma instância local do Sistema Gerenciador de Banco de Dados MySQL (MYSQL, 2020); c) a partir do Sistema Gerenciador de Banco de Dados, o processamento envolveu a migração dos dados para arquivos no formato CSV. E do formato CSV, os dados foram mapeados para o formato RDF, mediante o uso da ferramenta Sparqlify (SPARQLIFY, 2020) e um <i>script</i> de mapeamento; e d) os conjuntos de dados RDF são compartilhados como dados abertos conectados na Web de Dados, no <i>endpoint</i> http://lod.unicentro.br/sparql, em uma instância do Servidor OpenLink Virtuoso (VIRTUOSO, 2020).
Estado dos dados	Filtrados
Parâmetros para coleta de dados	conjunto completo de dados
Descrição da coleta de dados	os conjuntos de dados são adquiridos em arquivos nos formatos CSV ou XLS
Localização da fonte de dados	os dados originais são extraídos a partir dos sítios (SUCUPIRA, 2020); (SJR, 2020); e (JOURNAL METRICS, 2020)
Acessibilidade de dados	- junto ao artigo no formato TSV, em arquivos compactados (Tabela 1); - no <i>endpoint</i> (http://lod.unicentro.br/sparql), formalizando uma consulta SPARQL (vide Figura 4 e Listagem 1). Os dados podem ser recuperados nos formatos HTML⁷, JSON⁸, JavaScript. Turtle⁹, RDF/XML¹⁰, NT¹¹ e CSV em consonância aos requisitos do consumidor de dados; ou - via <i>link</i> encurtado (vide Quadro 1) em formato HTML.
Artigo de pesquisa relacionado	RAUTENBERG, Sandro; HILD, Tony Alexander; SOUZA, Lucélia de. Curadoria Digital de dados e Web de Dados: mantendo Dados Abertos Conectados para estudos bibliométricos e cientométricos. Em Questão, v. 24, p. 29-47, Edição Especial 6 EBBC, 2018, disponível em: https://seer.ufrgs.br/EmQuestao/article/view/86519, DOI: http://dx.doi.org/10.19132/1808-5245240.29-47

3.1 Descrição do conjunto de dados

Os índices cientométricos são relacionados e disponibilizados em 15 subconjuntos de dados em consonância aos anos de 2005 a 2019. Vale destacar que 2019 é último ano

⁷ Hyper Text Markup Language.

⁸ JavaScript Object Notation.

⁹ Terse RDF Triple Language.

¹⁰ Resource Description Framework/Extensible Markup Language.

¹¹ N-Triple.

de referência para disponibilização dos dados brutos nas fontes originais, dado o período de preparação e escrita deste artigo. Cada conjunto contempla 10 colunas (metadados) em um arquivo TSV, como segue:

1. **ISSN** – identificador inequívoco de um periódico;
2. **nameJournal** – nome do periódico atrelado ao ISSN;
3. **year** – ano referência em que o valor do índice é atrelado;
4. **idAreaCNPq** – identificador da área de conhecimento do CNPq a que a avaliação QUALIS é atrelada;
5. **nameAreaCNPq** – nome da área de conhecimento CNPq;
6. **QUALIS** – valor do estrato de avaliação do periódico em determinada área de conhecimento CNPq, dado um ano específico;
7. **idSubAreaSCOPUS** – identificador da subárea de conhecimento do SCOPUS a que uma avaliação SRJ ou SNIP é vinculada;
8. **nameSubAreaSCOPUS** – nome da área de conhecimento SCOPUS;
9. **SJR** – valor do índice SJR alcançado por um periódico em uma subárea do conhecimento SCOPUS, dado um ano específico; e
10. **SNIP** – valor do índice SNIP alcançado por um periódico em uma subárea do conhecimento SCOPUS, dado um ano específico.

Na Figura 3 são ilustrados alguns registros de dados recuperados, os quais representam exemplos de como as tuplas relacionam os índices Qualis, SJR e SNIP.

Figura 3: Exemplos de registros em um conjunto de dados

issn	nameJournal	year	idAreaCNPq	nameAreaCNPq	QUALIS	idSubAreaSCOPUS	nameSubAreaSCOPUS	SJR	SNIP
[...]	[...]	[...]	[...]	[...]	[...]	[...]	[...]	[...]	[...]
"0943-7444"	"Knowledge Organization"	2012	"38"	"EDUCAÇÃO"	"B1"	"3309"	"Library and Information Sciences"	"0.3660"	"0.8570"
"0943-7444"	"Knowledge Organization"	2012	"45"	"INTERDISCIPLINAR"	"B1"	"3309"	"Library and Information Sciences"	"0.3660"	"0.8570"
"0943-7444"	"Knowledge Organization"	2012	"1"	"MATEMÁTICA / PROBABILIDADE E ESTATÍSTICA"	"B3"	"3309"	"Library and Information Sciences"	"0.3660"	"0.8570"
"0943-7444"	"Knowledge Organization"	2012	"31"	"COMUNICAÇÃO E INFORMAÇÃO"	"A1"	"3309"	"Library and Information Sciences"	"0.3660"	"0.8570"
[...]	[...]	[...]	[...]	[...]	[...]	[...]	[...]	[...]	[...]

Fonte: dados da pesquisa.

Os subconjuntos de dados são disponibilizados de diversas formas. Estaticamente, relacionado a esse artigo, um arquivo compactado denominado subconjuntosDados.zip é

fornecido. Conforme a Tabela 1, são compactados 15 subconjuntos anuais, os quais foram coletados no *endpoint* <http://lod.unicentro.br/sparql> em 20 de março de 2021.

Tabela 1: subconjunto de dados e tuplas compartilhadas

ARQUIVO	TUPLAS DISPONIBILIZADAS
2005.tsv	60.548
2006.tsv	62.504
2007.tsv	65.186
2008.tsv	76.458
2009.tsv	79.494
2010.tsv	82.655
2011.tsv	180.538
2012.tsv	185.388
2013.tsv	188.916
2014.tsv	192.296
2015.tsv	125.582
2016.tsv	214.382
2017.tsv	219.031
2018.tsv	222.019
2019.tsv	222.607
TOTAL	2.177.604

Fonte: dados da pesquisa.

Os subconjuntos anualizados também podem ser recuperados dinamicamente de duas maneiras. Primeiramente, pode-se acessar os subconjuntos de dados via consultas SPARQL customizadas e submetidas ao *endpoint* <http://lod.unicentro.br/sparql>, conforme ilustrado na Figura 4, onde:

Figura 4: Interface do *endpoint* <<http://lod.unicentro.br>>



Fonte: adaptado de (RAUTENBERG; HILD; SOUZA, 2018).

- A) o endereço de Internet que implementa os serviços para acessar os recursos digitais;
- B) a área para desenvolver ou colar uma consulta SPARQL aos recursos digitais de um determinado conjunto; e

C) as opções disponibilizadas de formatação dos recursos digitais a serem recuperados.

Destaca-se que o endpoint <http://lod.unicentro.br/sparql> faz uso de uma instância do Servidor OpenLink Virtuoso (VIRTUOSO, 2020), o qual armazena os dados no formato RDF e acessa os recursos via consultas SPARQL. A escolha dessa base tecnológica encontra consonância aos preceitos dos Dados Abertos Conectados (W3C, 2020a) para privilegiar: a) o grau máximo de abertura e consumo de recursos digitais na Web de Dados (5-STAR, 2021); e b) a diversidade de formatos de apresentação dos recursos recuperados (como destacado na Figura 4c: HTML, JSON, JavaScript, Turtle, RDF/XML, N-Triples ou CSV). Em face disso, a referida escolha tecnológica permite o consumo automatizado dos dados compartilhados por parte de aplicações computacionais que usam a Web de Dados como um repositório universal de recursos digitais. Ou seja, o referido *endpoint* torna-se a interface que possibilita o reúso, a compreensão, a interligação, a descoberta, a confiança, o acesso, a interoperabilidade e a processabilidade de recursos digitais na Web de Dados para com agentes computacionais (robôs) ou aplicações cientométricas baseadas na web.

Em relação aos dados disponibilizados neste artigo, a partir da interface do *endpoint* <http://lod.unicentro.br/sparql>, consultas SPARQL podem ser customizadas e permitir a exploração dos recursos RDF dos índices Qualis, SJR e SNIP, como exemplificado na consulta implementada na Listagem 1.

Listagem 1: Consulta SPARQL para recuperar subconjuntos de dados a partir do *endpoint* <http://lod.unicentro.br/sparql> - ano 2005

```
01 PREFIX qualis: <http://lod.unicentro.br/QualisBrasil/>
02 PREFIX sjr: <http://lod.unicentro.br/SJR/>
03 PREFIX snip: <http://lod.unicentro.br/SNIP/>
04 PREFIX dc: <http://purl.org/dc/elements/1.1/>
05 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
06 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
07 PREFIX foaf: <http://xmlns.com/foaf/0.1/>
08 PREFIX bibo: <http://purl.org/ontology/bibo/>
09
10 SELECT ?issn ?journal ?year ?idAreaCNPq ?areaCNPq ?QUALIS ?idSubAreaSCOPUS
11 ?subAreaSCOPUS ?SJR ?SNIP WHERE {
12   {
13     SELECT DISTINCT ?issn ?journal ?year ?idSubAreaSCOPUS ?subAreaSCOPUS ?SJR WHERE
14     {
15       ?eSJR rdf:type sjr:Evaluation .
16       ?eSJR sjr:hasJournal ?j .
17       ?eSJR sjr:hasSubAreaScopus ?s .
18       ?eSJR sjr:hasYearEvaluation ?y .
19       ?eSJR sjr:hasScore ?sc .
20
21       ?j bibo:issn ?issn .
22       ?j foaf:name ?journal .
23       ?s dc:identifier ?idSubAreaSCOPUS .
24       ?s dc:title ?subAreaSCOPUS .
25       ?y rdf:value ?year .
```

```

26      ?sc rdf:value      ?SJR .
27      FILTER (?year = "2005") .
28  }
29  }
30  OPTIONAL
31  {
32      SELECT DISTINCT ?issn ?journal ?year ?idAreaCNPq ?areaCNPq ?QUALIS WHERE
33      {
34          ?eQualis rdf:type qualis:Evaluation .
35          ?eQualis qualis:hasJournal ?j .
36          ?eQualis qualis:hasKnowledgeField ?s .
37          ?eQualis qualis:hasYearEvaluation ?y .
38          ?eQualis qualis:hasScore ?sc .
39
40          ?j bibo:issn ?issn .
41          ?j foaf:name ?journal .
42          OPTIONAL { ?s dc:identifier ?idAreaCNPq } .
43          OPTIONAL { ?s dc:title ?areaCNPq } .
44          ?y rdf:value ?year .
45          ?sc rdf:value ?QUALIS .
46      FILTER (?year = "2005") .
47      }
48  }
49  OPTIONAL
50  {
51      SELECT DISTINCT ?issn ?journal ?year ?idSubAreaSCOPUS ?subAreaSCOPUS ?SNIP WHERE
52      {
53          ?eSNIP rdf:type snip:Evaluation .
54          ?eSNIP snip:hasJournal ?j .
55          ?eSNIP snip:hasSubAreaScopus ?s .
56          ?eSNIP snip:hasYearEvaluation ?y .
57          ?eSNIP snip:hasScore ?sc .
58
59          ?j bibo:issn ?issn .
60          ?j foaf:name ?journal .
61          ?s dc:identifier ?idSubAreaSCOPUS .
62          ?s dc:title ?subAreaSCOPUS .
63          ?y rdf:value ?year .
64          ?sc rdf:value ?SNIP .
65      FILTER (?year = "2005") .
66      }
67  }
68  }
69  ORDER BY ?journal

```

Fonte: dados da pesquisa.

Na consulta SPARQL exemplificada, é importante observar que esta customiza a recuperação do subconjunto de registros no ano 2005 (observe as linhas 27, 46 e 65). Neste sentido, a consulta também pode ser reutilizada, ao adaptar suas linhas de filtragem para adquirir outros subconjuntos. Por exemplo, à medida que as fontes originais sejam incrementadas, outros subconjuntos podem constituídos, compreendendo os registros de dados para os anos vindouros. Isso também permite a obtenção de dados atualizados dos índices SJR e SNIP, ao considerar que para estes índices, os fatores de impacto dos periódicos são revistos em relação aos quantitativos de citações realizadas no triênio vigente.

Por fim, a outra forma dinâmica de obtenção dos subconjuntos de dados é mediada por *links* encurtados que implementam consultas SPARQL (Quadro 1), considerando os anos no intervalo [2005; 2019].

Quadro 2: Arquivos de subconjunto de dados e seus *links* encurtados

Ano referência	Link encurtado
2005	http://tinyurl.com/8689y8c7
2006	http://tinyurl.com/yvbsmxhf
2007	http://tinyurl.com/ivucy6rb
2008	http://tinyurl.com/h37hsn85
2009	http://tinyurl.com/hvun3e5y
2010	http://tinyurl.com/647fju5
2011	http://tinyurl.com/2kezy73r
2012	http://tinyurl.com/4sm7adbh
2013	http://tinyurl.com/3hfjcsxw
2014	http://tinyurl.com/3z9bf6sb
2015	http://tinyurl.com/3kta888s
2016	http://tinyurl.com/yrf4enk9
2017	http://tinyurl.com/5h6uvdn5
2018	http://tinyurl.com/49j9pfc3
2019	http://tinyurl.com/2vz8ceve

Fonte: dados da pesquisa.

Cabe ressaltar que, no sentido da Curadoria Digital de Dados, o Quadro 1 expressa o incremento dos conjuntos de dados disponibilizados, uma vez que o histórico disponibilizado anteriormente compreendia o período [2005; 2016] (RAUTENBERG; HILD; SOUZA, 2018).

REFERÊNCIAS

5-STAR. **5-Star Open Data**. Disponível em: <http://5stardata.info/en>. Acesso em: 17 set. 2021.

AUER, S. Introduction to lod2. In: **Linked Open Data – Creating Knowledge Out of Interlinked Data**. AUER, S.; BRYL, V.; TRAMP, C (Ed.). Lecture Notes in Computer Science. Springer-Verlag, 2014.

DIGITAL CURATION CENTER. **DCC Curation Lifecycle Model**. Disponível em: <http://www.dcc.ac.uk/sites/default/files/documents/publications/DCCLifecycle.pdf>. Acesso em: 16 mar. 2020.

JOURNAL METRICS. **Journal Metrics - Scopus.com**. Disponível em: <https://www.scopus.com/sources>. Acesso em: 16 mar. 2020.

MySQL. **MySQL**. Disponível em: <https://www.mysql.com/>. Acesso em: 16 mar. 2020.

RAUTENBERG, S.; HILD, T. A.; SOUZA, L. de. Curadoria Digital de dados e Web de Dados: mantendo Dados Abertos Conectados para estudos bibliométricos e cientométricos. **Em Questão**, v. 24, p. 29-47, Edição Especial 6 EBBC, 2018, disponível em:

<https://seer.ufrgs.br/EmQuestao/article/view/86519>, DOI: <http://dx.doi.org/10.19132/1808-5245240.29-47>.

RAUTENBERG, S.; MOTYL, S.; BURDA, A.; SILVERIO, A.; MOURA, F. M. Dados Abertos Conectados e Gestão do Conhecimento: estudos de caso cientométricos em uma universidade brasileira. **Perspectivas em Ciência da Informação**, v. 22, p. 116-142, 2017a, disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2885>. Acesso em: 19 mai. 2021.

RAUTENBERG, S.; SOUZA, L.; HAUAGGE, J.; HILD, T.; MICHELON, G.; BURDA, A. representando índices cientométricos como dados abertos conectados. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 28., 2017, Marília. **Anais...** Marília-SP: PPGCI, UNESP, 2017b.

SJR. **Scimago Journal & Country Rank**. Disponível em: <http://www.scimagojr.com/journalrank.php> . Acesso em: 16 mar. 2020.

SPARQLIFY. **Sparqlify - Agile Knowledge Engineering and Semantic Web (AKSW)**. Disponível em: <http://aksw.org/Projects/Sparqlify.html>. Acesso em: 16 mar. 2020.

SUCUPIRA. **Plataforma Sucupira**. Disponível em: <https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/veiculoPublicacaoQualis/listaConsultaGeralPeriodicos.jsf>. Acesso em: 16 mar. 2020.

VIRTUOSO. **OpenLink Virtuoso Home Page**. Disponível em: <https://virtuoso.openlinksw.com/>. Acesso em: 16 mar. 2020.

W3C. **Data on the Web best practices: W3C recommendation 31 January 2017**. Disponível em: <https://www.w3.org/TR/2017/REC-dwbp-20170131/>. Acesso em: 15 mar. 2020a.

W3C. **Web Services Description Requirements**. Disponível em: <https://www.w3.org/TR/2002/WD-ws-desc-reqs-20021028/#normDefs>. Acesso em: 09 set. 2020b.

NOTAS

AGRADECIMENTOS

O primeiro autor agradece à Fundação Araucária pelo suporte financeiro ao projeto de pesquisa intitulado “Curadoria Digital e Dados Abertos Conectados: um estudo da preservação de recursos digitais na Web de Dados para estudos cientométricos” (FA - Convênio 49773/19FA).

CONTRIBUIÇÃO DE AUTORIA

Concepção e elaboração do manuscrito: Rautenberg, S.; Souza, L. de; Hild, T. A.

Coleta de dados: Rautenberg, S.

Análise de dados: Rautenberg, S.

Discussão dos resultados: Rautenberg, S.; Souza, L. de

Revisão e aprovação: Rautenberg, S.; Souza, L. de; Hild, T. A.

CONJUNTO DE DADOS DE PESQUISA

Todo o conjunto de dados que dá suporte aos resultados deste estudo foi publicado no artigo e na seção “Materiais suplementares”.

FINANCIAMENTO

Fundação Araucária (FA - Convênio 49773/19FA).



CONSENTIMENTO DE USO DE IMAGEM

Não se aplica

APROVAÇÃO DE COMITÊ DE ÉTICA EM PESQUISA

Não se aplica.

CONFLITO DE INTERESSES

Não se aplica

LICENÇA DE USO – uso exclusivo da revista

Os autores cedem à **Encontros Bibli** os direitos exclusivos de primeira publicação, com o trabalho simultaneamente licenciado sob a [Licença Creative Commons Attribution](#) (CC BY) 4.0 International. Esta licença permite que **terceiros** remixem, adaptem e criem a partir do trabalho publicado, atribuindo o devido crédito de autoria e publicação inicial neste periódico. Os **autores** têm autorização para assumir contratos adicionais separadamente, para distribuição não exclusiva da versão do trabalho publicada neste periódico (ex.: publicar em repositório institucional, em site pessoal, publicar uma tradução, ou como capítulo de livro), com reconhecimento de autoria e publicação inicial neste periódico.

PUBLISHER

Universidade Federal de Santa Catarina. Programa de Pós-graduação em Ciência da Informação. Publicação no [Portal de Periódicos UFSC](#). As ideias expressadas neste artigo são de responsabilidade de seus autores, não representando, necessariamente, a opinião dos editores ou da universidade.

EDITORES

Enrique Muriel-Torrado, Edgar Bisset Alvarez, Camila Barros, Genilson Geraldo.

HISTÓRICO

Recebido em: 20-03-2021 – Aprovado em: 15-06-2021 – Publicado em: 09-08-2021

