

PERIÓDICOS CIENTÍFICOS BRASILEIROS INDEXADOS NO GOOGLE SCHOLAR METRICS

BRAZILIAN SCIENTIFIC JOURNALS INDEXED IN GOOGLE SCHOLAR METRICS

Adilson Luiz Pinto¹
Fabio Lorensi Do Canto²
Edson Mario Gavron³
Marcos Talau⁴

RESUMO

Este trabalho tem o objetivo de identificar quantos periódicos científicos brasileiros são indexados pelo Google Scholar Metrics e o de descrever as características das publicações. As variáveis analisadas são o escopo temático e o impacto (índice e mediana h5), a classificação Qualis, os idiomas do conteúdo, a periodicidade e a vinculação institucional. Foram recuperados dados de 1.906 periódicos científicos brasileiros contidos no UlrichsWeb. Foi desenvolvido um *script* para busca e extração automatizada dos dados no sistema. Para os títulos não encontrados foi realizada uma busca manual. Foram identificados 1.063 periódicos brasileiros no Google Scholar Metrics, que correspondem a 55% do total pesquisado. Constatou-se que 33% das publicações são de Ciências Humanas, 17% de Ciências da Saúde e Medicina e 17% de Ciências Sociais; os periódicos com maior média do índice h5 são Ciências da Saúde (11,34) e Ciências Agrárias e Biológicas (9,11) e aproximadamente 90% do total de publicações possui índice h5 inferior a 10. Verificou-se ainda que aproximadamente 50% dos periódicos é mantido por universidades públicas federais ou estaduais, 54% publica conteúdo somente em português e 65% é semestral ou trimestral. Pode-se afirmar que o Google Scholar Metrics é uma fonte relevante de análise de impacto de periódicos brasileiros, principalmente de publicações acadêmicas, das áreas de Ciências Humanas e Sociais e sem outro indicador de impacto. No entanto, o uso desse sistema para fins avaliativos oficiais deve ser realizado com cautela, considerando que o sistema apresenta inconsistências e limitações.

Palavras-chave: Google Scholar Metrics. Periódicos científicos brasileiros. Índice h5.

ABSTRACT

This work aims to identify how many Brazilian scientific journals are indexed in Google Scholar Metrics and to describe the characteristics of the publications. The variables analyzed are the subject and impact (h5-index and h5-median), the Qualis Brazilian journal classification, the languages, the periodicity and the type of publisher. Data from 1,906 Brazilian scientific journals indexed in UlrichsWeb were retrieved. A script for automated search and data extraction in Google Scholar Metrics was used. The journals not found were searched manually. As a result, 1,063 Brazilian journals were identified in Google Scholar Metrics (55% of the total). It was found that 33% of publications are in Human Sciences, 17% in Health Sciences and Medicine and 17% in Social Sciences; the journals with the highest average of the h5-index are Health Sciences (11.34) and Agrarian and Biological Sciences (9.11) and approximately 90% of the total publications have an h5-index below 10. It was also found that approximately 50% of the journals is maintained by public federal or state universities, 54% publishes works

¹ Doutor em Documentação pela Universidad Carlos III de Madrid, Espanha. Coordenador do Programa de Pós-Graduação em Ciência da Informação. Universidade Federal de Santa Catarina, Brasil. Bolsista de Produtividade do CNPq. ORCID <https://orcid.org/0000-0002-4142-2061>. E-mail: adilson.pinto@ufsc.br

² Doutorando pelo Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Santa Catarina, Brasil. ORCID: <https://orcid.org/0000-0002-8338-1931>. E-mail: fabio.ic@ufsc.br

³ Doutorando pelo Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Santa Catarina, Brasil. ORCID: <https://orcid.org/0000-0002-8761-2178>. E-mail: e.m.gravon@gmail.com

⁴ Doutorando em Redes de Computadores pelo Programa de Pós-Graduação em Engenharia Elétrica e Industrial da Universidade Tecnológica Federal do Paraná, Brasil. Professor adjunto da Universidade Tecnológica Federal do Paraná, Brasil. ORCID: <https://orcid.org/0000-0001-7027-1023>. E-mail: marcostalau@gmail.com

only in Portuguese and 65% is biannual or quarterly. These results suggested that Google Scholar Metrics is a relevant source of impact analysis for Brazilian journals, mainly academic publications, in the areas Humanities and Social Sciences and without any other impact indicator. However, the use of this system for official evaluation purposes should be carried out with caution, owing to the system has inconsistencies and limitations.

Keywords: Google Scholar Metrics. Brazilian scientific journals. H5-index.

Artigo submetido em 16/06/2020 e aceito para publicação em 03/07/2020

1 INTRODUÇÃO

O Google Scholar Metrics - GSM foi lançado em 2012 com a finalidade mensurar o impacto de publicações por meio das citações coletadas no Google Scholar - GS. O sistema utiliza como indicador bibliométrico o índice h (HIRSCH, 2005), com período de cinco anos, e calcula a mediana do número de citações recebidas pelos trabalhos que compõe o índice h5. Além do recurso de busca por título, o GSM elabora *rankings* de publicações de maior impacto por categoriais e subcategorias de assunto e também por doze idiomas (LÓPEZ-CÓZAR; CABEZAS-CLAVIJO, 2012; 2013; JACSÓ, 2012).

Ao contrário das bases de dados tradicionais, que funcionam a partir de sistemas fechados e de padrão elitista, ou seja, com enfoque em publicações de maior prestígio, o GSM indexa fontes acadêmicas disponíveis na web, o que resulta em uma cobertura mais ampla (ORDUÑA-MALEA; LÓPEZ-CÓZAR, 2014). Essa política de indexação inclui publicações regionais, de idiomas que não o inglês e de áreas do conhecimento normalmente não favorecidas pelas demais bases de dados científicas, tais como as Ciências Humanas e Sociais (BORNMANN et al., 2016; MINGERS; MEYER, 2017).

Por esse motivo, o GSM vem sendo considerado uma ferramenta alternativa para avaliação de periódicos científicos, especialmente daqueles que não possuem outro indicador bibliométrico gerado em uma base de dados (HARZING, 2014). Na recente avaliação de meio termo do Qualis/Capes em meados de 2019, o índice h5 do GSM já foi utilizado para classificação de periódicos que não possuem Fator de Impacto ou Citescore (CAPES, 2019a). Essa novidade confere relevância ao GSM no cenário de avaliação científica brasileira, tendo em vista que o número de periódicos brasileiros indexadas no Journal Citation Reports - JCR ou na Scopus Elsevier é modesto em comparação com o total de publicações nacionais.

No entanto, ainda pairam dúvidas acerca da cobertura e do funcionamento do GSM em relação às publicações brasileiras. O número total de periódicos indexados é desconhecido. A obtenção desses dados é dificultada pela impossibilidade de delimitação geográfica no mecanismo de busca do sistema (LÓPEZ-CÓZAR; ORDUÑA-MALEA; MARTÍN-MARTÍN, 2019), o que exige a criação de métodos

PERIÓDICOS CIENTÍFICOS BRASILEIROS INDEXADOS NO GOOGLE SCHOLAR METRICS
experimentais de raspagem de dados web. As características das publicações nacionais indexadas também são desconhecidas, diante da inexistência de pesquisas com esse escopo.

Nesse contexto, essa pesquisa tem por finalidade responder as seguintes questões: quantos periódicos científicos brasileiros são indexados no GSM? Quais são as características dessas publicações no que se refere ao escopo temático, impacto (índice h5 e mediana h5), classificação Qualis, idiomas do conteúdo, periodicidade e natureza do vínculo institucional?

Com os resultados obtidos, pretende-se aprofundar a discussão sobre o uso do índice h5 na avaliação de periódicos brasileiros a partir da nova metodologia de avaliação do Qualis. Pretende-se ainda apresentar um panorama da cobertura de periódicos brasileiros visando contribuir para divulgação dessa ferramenta bibliométrica perante a comunidade científica brasileira.

2 GOOGLE SCHOLAR METRICS

Em abril 2012 foi lançado o GSM, uma extensão do GS voltada para mensurar o impacto de publicações a partir dos dados de citação coletados por essa plataforma em fontes acadêmicas da web. Essa ferramenta bibliométrica, aliada ao Google Scholar Citations - GSC lançado alguns meses antes (ORTEGA, 2015), consolidou a posição da Google na área de informação científica e acirrou a concorrência com as demais bases de dados dedicadas à avaliação de impacto científico. Desde então, o GSM passou a ser considerada uma fonte alternativa ao JCR e a Scopus, as duas bases de dados mais utilizadas para fins de avaliação da produção científica (JACSÓ, 2012; LÓPEZ-CÓZAR; CABEZAS-CLAVIJO, 2012, 2013; ORDUÑA-MALEA; LÓPEZ-CÓZAR, 2014; GU; BLACKMORE, 2017).

2.1 Funcionamento e características

Mantendo a mesma política dos demais recursos do GS, o GSM é gratuito e independe de identificação do usuário. O reconhecimento da marca Google aliado a desnecessidade de instalação de software ou *plug-in* adicional são fatores que favorecem a utilização. O sistema mantém a tendência minimalista, com interface e recursos limitados em comparação com outras bases de dados científicas (JACSÓ, 2012; LÓPEZ-CÓZAR; CABEZAS-CLAVIJO, 2013).

O indicador bibliométrico utilizado pelo GSM é o índice h5, ou seja, o índice de Hirsch (2005) com uma janela de citação de cinco anos. Essa configuração é considerada acertada, pois o índice h é uma métrica reconhecida e fácil de calcular (LÓPEZ-CÓZAR; CABEZAS-CLAVIJO, 2012). Além disso, o período de cinco anos implica em uma medida mais abrangente que os períodos de dois ou

três anos usualmente utilizados (HARZING, 2014; WALTMAN, 2016). O período de análise atual do GSM engloba trabalhos publicados entre 2014 e 2018 e citações coletadas até julho de 2019 (LÓPEZ-CÓZAR; MARTÍN-MARTÍN, 2019). Também é calculada a mediana h5, isto é, a mediana do número de citações recebidas pelos artigos que integram o índice h5 de cada publicação. Mais do que uma métrica independente, a mediana h5 serve como critério de desempate para publicações com o mesmo índice h (JACSÓ, 2012).

Há três formas de recuperar as publicações indexadas no GSM: visualizar listas por doze idiomas, visualizar listas por categorias e subcategorias assuntos ou pesquisar por título. Os idiomas disponíveis são o inglês, chinês, português, espanhol, alemão, russo, francês, japonês, coreano, polonês, ucraniano e indonésio (ORDUÑA-MALEA; LÓPEZ-CÓZAR, 2014). São apresentadas apenas as 100 publicações de maior impacto em cada idioma. Já a visualização por assunto é limitada a 20 publicações de maior impacto divididas em oito categorias de assuntos e 290 subcategorias de assuntos, mas nestes casos somente no idioma inglês (GOOGLE SCHOLAR, 2019).

A busca por publicações pode ser feita somente por título ou palavra do título, digitado em campo único, sem opção de busca avançada. O número de resultados da busca também é limitado a 20 registros. O índice h5 é baseado em dados de citação coletados no GS. É possível acessar os artigos que compõe o índice h5, bem como os trabalhos citantes (GOOGLE SCHOLAR, 2019).

A indexação dos títulos é automatizada a partir das fontes já mapeadas pelo GS, o que inclui repositórios, bases de dados, sítios de universidades, portais de periódicos além de outras fontes acadêmicas. Os critérios mínimos de inclusão no GSM são que a publicação tenha pelo menos 100 trabalhos publicados no período de avaliação e que tenha recebido pelo menos uma citação (índice $h5 > 0$). Aspectos técnicos, contudo, podem favorecer o trabalho do Googlebot, como a edição dos metadados e o uso das versões atualizadas dos softwares repositórios (Eprints, Digital Commons e DSpace) e de editoração recomendados (Atypon, Highwire, Ingenta, Silverchair e OJS). Informações técnicas estão disponíveis em um guia para desenvolvedores web (GOOGLE SCHOLAR, 2019).

É recomendado, portanto, que editores científicos atentem para os requisitos técnicos do sistema, visando facilitar a indexação automatizada e, por consequência, a apresentação adequada dos metadados e a coleta de dados de citação. Essas medidas podem minimizar a ocorrência de erros, que tendem a ser em maior número no GSM em comparação a bases de dados com controle bibliográfico.

2.2 Críticas e limitações

Em que pese o GSM seja visto com bons olhos por pesquisadores de estudos métricos, determinados aspectos relacionados ao funcionamento e à confiabilidade desse sistema fazem com a sua utilização para fins avaliativos oficiais seja realizado com cautela e, em alguns casos, seja desaconselhado (LÓPEZ-CÓZAR; CABEZAS-CLAVIJO, 2012, 2013; JACSÓ, 2012).

Um dos principais problemas é a ausência de padronização de títulos, o que dificulta a recuperação, considerando a inexistência de um recurso de busca por ISSN. A falta de controle bibliográfico acarreta também duplicações de registros e erros em contagens de citações (LÓPEZ-CÓZAR; CABEZAS-CLAVIJO, 2012; JACSÓ, 2012). A impossibilidade de exclusão das autocitações no cálculo do índice h5 é apontado como limitação importante. O receio de que autocitações propositais elevem o impacto de publicações foi inclusive manifestado pela Capes, ao definir o índice h5 como critério da próxima avaliação Qualis (CAPES, 2019b).

Erros em contagens de citações, não apenas no GSM, mas também das demais ferramentas que se utilizam das citações coletadas pelo GS são frequentemente mencionadas em denúncias de falta de credibilidade dos dados apresentados pelas ferramentas Google (JACSÓ, 2010, 2011; TEIXEIRA DA SILVA, 2018). Muitas dessas questões têm origem no próprio sistema de coleta de citações baseados em documentos web, que apresentam maior instabilidade e podem desaparecer e reaparecer repentinamente na Internet.

Outra problemática diz respeito a possibilidade de manipulação do índice h5 por meio de citações falsas (LÓPEZ-CÓZAR; ROBINSON-GARCÍA; TORRES-SALINAS, 2014). A existência de más práticas em citações baseadas em documentos web é mencionada também por López-Cózar e Martín-Martín (2019) que sugerem que o uso do GSM para avaliação de periódicos não pode ser limitado a identificação do índice h5, mas deve ser acompanhada de uma verificação das citações que compõe o indicador visando a detectar citações irregulares.

Além disso, os recursos limitados do GSM não permitem distinguir quais periódicos são revisados por pares e quais não são, disfunção que pode definir essa fonte como de natureza acadêmica, mas não estritamente científica (LÓPEZ-CÓZAR; CABEZAS-CLAVIJO, 2012). A inexistência de um recurso de busca avançada e a não divulgação de uma lista geral de periódicos indexados dificultam análises amplas. Muitos dos estudos com enfoque no tamanho e na cobertura do GSM apresentam tão somente resultados estimativos (GU; BLACKMORE, 2017; LÓPEZ-CÓZAR; ORDUÑA-MALEA; MARTÍN-MARTÍN, 2019).

2.3 Tamanho e cobertura

O sistema de indexação automatizada utilizado no GSM resulta em uma ampla cobertura, mas o número total de publicações ainda é desconhecido. Extraindo as listas de publicações por categorias e subcategorias de assuntos e por idiomas, foi possível identificar aproximadamente 5.565 publicações (LÓPEZ-CÓZAR; CABEZAS-CLAVIJO, 2013). Esse conjunto, todavia, contém somente as publicações de maior impacto, sendo que os rankings por assunto abrangem apenas publicações de língua inglesa (GOOGLE SCHOLAR, 2019).

As pesquisas envolvendo o tamanho e cobertura do GSM restaram inconclusivas, ou apresentaram apenas estimativas. Gu e Blackmore (2017) constataram que de 41.787 periódicos extraídos do UlrichsWeb, somente 10.354 (20,8%) estavam indexados no GSM. O método utilizado, contudo, exigia a correspondência exata entre a grafia do título nas duas fontes. Logo, a ausência de padronização de entradas pode ter contribuído para um resultado significativamente inferior ao número real, o que evidencia a necessidade de realizar pesquisas por variações do título, visando diminuir a ocorrência de resultados falsos negativos.

Diante dessa imprecisão no tamanho real da base, estima-se que seja algo em torno de 40 mil total de periódicos indexados no GSM (LÓPEZ-CÓZAR; CABEZAS-CLAVIJO, 2013). Outra em estimativa aponta que de 9.000 periódicos das áreas de artes, humanidades e ciências sociais, aproximadamente 27% são do Estados Unidos, 21% do Reino Unido, 9% da Espanha, 5% do Brasil, 4% da Holanda, 3% França, e 2% Alemanha e 29% de outros países (LÓPEZ-CÓZAR; ORDUÑA-MALEA; MARTÍN-MARTÍN, 2019).

Pesquisadores vêm monitorando o índice h5 dos periódicos espanhóis. De 2.668 publicações com origem nesse país contidas no UlrichsWeb, foram encontradas 1.299 (48,7%) na edição de 2016 (LÓPEZ-CÓZAR; ORDUÑA-MALEA; MARTÍN-MARTÍN, 2019) e 1.448 (54,3%), na edição de 2019 do GSM (LÓPEZ-CÓZAR; MARTÍN-MARTÍN, 2019). Esse número é quase o triplo do número de periódicos espanhóis indexados na Scimago Journal Rank (571) mais de dez vezes o de indexados no JCR (126). As áreas com maior cobertura de publicações espanholas são as Ciências Sociais (48,7%), as Artes e Humanidades (29%), as Ciências da Saúde (11,2%) as Ciências Naturais e as Engenharias (11,1%). Foram detectados ainda 95 periódicos com entradas duplicadas, o que corresponde a 6,5% de duplicação nessa última versão do sistema (LÓPEZ-CÓZAR; MARTÍN-MARTÍN, 2019).

Estudos com delimitação temática também foram conduzidos. López-Cózar e Caballero (2013) constataram que de 663 periódicos da área de Comunicação de diversos países, 277 foram localizadas no GSM, número que representa 41,7% do total. Já Costa, Canto e Pinto (2020), identificaram que de 46 periódicos brasileiros de Ciência da Informação listados na BRAPCI, 21 estava indexado no GSM, número que representa 45% do total.

A cobertura mais ampla, inclusive nas áreas de Ciências Humanas e Sociais, de publicações de outros idiomas que não o inglês e de países emergentes, faz com que o GSM adquira relevância no processo de avaliativo da ciência, suprimindo uma lacuna deixada pela JCR e pela Scopus na geração de indicadores de impacto (BORNMANN et al., 2016). Além disso, o uso do índice h com período de cinco anos favorece a avaliação desse escopo da ciência, pois é um período de análise mais amplo em contraste ao período de dois e de três anos, respectivamente, adotados pelo Fator de Impacto e pelo CiteScore (WALTMANN, 2016).

3 PROCEDIMENTOS METODOLÓGICOS

A coleta e o tratamento dos dados foram realizados entre os meses de março e maio de 2020. Por meio do diretório UlrichsWeb, foram recuperados os registros de 1.906 periódicos científicos brasileiros ativos, de conteúdo acadêmico/científico e disponíveis online. Foram extraídos os dados dos campos ID (nº identificador do UlrichsWeb), título e subtítulo, ISSN, escopo temático, vínculo institucional, link de acesso, variações do título, idiomas do conteúdo e periodicidade. Os campos título e subtítulo foram convertidos para arquivo de texto, visando realizar a busca no GSM. Foi realizada a correção ortográfica e a exclusão do termo “(online)” presente após o título em um número significativo de registros.

Foram definidas duas técnicas de busca, uma automatizada e outra manual. A busca automatizada foi realizada por meio de um *script* em linguagem *Python* para pesquisa dos títulos e raspagem dos dados html apresentados no GSM (título, índice h5 e mediana h5). O sistema foi configurado recuperar resultados idênticos ou com pelo menos 95% de similaridade com o termo pesquisado. Além disso, foram recuperados resultados nos quais o termo pesquisado estava contido em uma expressão maior, para posterior análise de correspondência entre o periódico pesquisado e o recuperado.

Os títulos não encontrados retornaram resultado zero. Nestes casos, procedeu-se com a técnica de busca manual. Além do título expresso no UlrichsWeb, foram realizadas buscas por variações do

título e abreviaturas. A base de dados do ISSN e o site do periódico pesquisado foram utilizados para saneamento de dúvidas acerca de eventuais homônimas, duplicações e variações por idioma. Em alguns casos, foi necessário acessar o texto completo de artigos do índice h5 para análise de correspondência.

Os dados obtidos foram inseridos na lista geral com os demais dados do UlrichsWeb. A partir disso, foram realizadas as análises de impacto (índice h5 e mediana h5), idiomas do conteúdo, periodicidade e vínculo institucional. Para análise do escopo temático, foi realizada a correspondência dos assuntos apresentados no UlrichsWeb para as grandes áreas do conhecimento da Capes. As Ciências Agrárias e as Ciências Biológicas foram agrupadas e a área Multidisciplinar da Capes não foi considerada. Para a análise do Qualis, foram utilizados os dados extraídos da Plataforma Sucupira (2013-2016), bem como o resultado da avaliação de meio termo do novo Qualis (2019) divulgado pela Capes.

4 RESULTADOS E DISCUSSÃO

Foram encontrados 1.063 periódicos científicos brasileiros indexados no GSM, número que corresponde a 55,77% do total de 1.906 publicações utilizadas de base para a pesquisa. Esse resultado percentual é quase idêntico ao obtido por (LÓPEZ-CÓZAR; ORDUÑA-MALEA; MARTÍN-MARTÍN, 2019) em pesquisa com periódicos da Espanha, indicando que esse padrão pode ser utilizado como hipótese em pesquisas futuras.

4.1 Escopo temático e impacto

A Tabela 1 mostra a distribuição percentual dos periódicos de acordo com o escopo temático, bem como os valores médios do índice h5 e da mediana h5. As áreas do conhecimento com maior cobertura são Ciências Humanas - CHS, Ciências da Saúde e Medicina - CSM e Ciências Sociais Aplicadas - CSA, que juntas abrangem quase 70% do total de publicações. As áreas com menor cobertura são Linguística, Literatura e Artes - LLA, Ciências Exatas e da Terra - CET e Engenharias e Tecnologia - EGT, que possuem um percentual aproximado de 7% cada uma. Ciências Agrárias e Biológicas – CAB engloba quase 10% de publicações.

Tabela 1 – Escopo temático e impacto dos periódicos, por área do conhecimento

Áreas do conhecimento	%	h5	h5-m
Ciências Humanas - CHS	33,11	5,98	8,68
Ciências da Saúde e Medicina - CSM	17,80	11,94	16,15
Ciências Sociais Aplicadas – CSA	17,72	8,01	11,24
Ciências Agrárias e Biológicas - CAB	9,69	9,65	13,17
Linguística, Literatura e Artes - LLA	7,37	3,69	5,33
Ciências Exatas e da Terra - CET	7,37	7,13	9,75
Engenharias e Tecnologia - EGT	6,95	7,38	10,38
Total	100	7,68	10,67

Fonte: Dados da pesquisa (UlrichsWeb).

Quanto ao impacto das publicações, a média geral do índice h5 é de 7,68 enquanto o da mediana h5 é de 10,67. As áreas com maiores médias são a CSM, CAB e CSA.

A Tabela 3 mostra a distribuição de periódicos de acordo com faixas de valores do índice h5 e da mediana h5. Quase a metade dos periódicos está distribuída na primeira faixa de impacto, com índice h5 de valores entre um e cinco. Um segundo grupo de aproximadamente 30% dos periódicos, se posiciona em uma segunda faixa de valores, com índices h5 entre seis e dez. Menos de 10% possui índices h5 superiores a 15, valor a partir do qual as publicações podem ser consideradas de alto impacto no GSM.

Tabela 2 – Distribuição (nº e %) dos periódicos conforme índice h5 e mediana h5

Valor	h5		h5-m	
	nº	%	nº	%
Entre 1 e 5	521	49,01	307	28,88
Entre 6 e 10	321	30,20	386	36,31
Entre 11 e 15	119	11,19	164	15,43
Entre 16 e 20	60	5,64	87	8,18
Entre 21 e 25	19	1,79	57	5,36
Entre 26 e 30	15	1,41	23	2,16
Entre 31 e 40	6	0,56	29	2,73
Mais de 40	2	0,19	10	0,94
Total	1.063	100	1.063	100

Fonte: Dados da pesquisa (GSM).

Com padrão de distribuição semelhante, aproximadamente 29% dos periódicos possui mediana h5 com valores entre um e cinco, seguida de um conjunto de 36% publicações com valores de mediana

entre cinco e dez. Esses dados confirmam a política não elitista do GSM, que possibilita a avaliação de um grande número de periódicos brasileiros e de diferentes níveis de impacto de citação.

4.2 Classificação Qualis

A Tabela 2 mostra a distribuição dos periódicos de acordo com os resultados do Qualis Periódicos das avaliações do quadriênio 2013-2016 e de meio termo de 2019. Referente ao período 2013-2016, as publicações foram classificadas nas áreas as quais possuíam aderência, sendo identificadas 9.595 avaliações distintas para 963 periódicos dentre os analisados (100 não foram avaliados).

Tabela 3 – Distribuição (nº e %) dos periódicos conforme a classificação Qualis

Qualis (2013-2016)			Qualis (2019)		
Estrato	n.	%	Estrato	n.	%
A1	186	1,94	A1	115	11,10
A2	449	4,68	A2	136	13,13
B1	1294	13,49	A3	138	13,32
B2	1442	15,03	A4	155	14,96
B3	1588	16,55	B1	148	14,29
B4	2016	21,01	B2	121	11,68
B5	1800	18,76	B3	86	8,30
C	820	8,55	B4	77	7,43
--	--	--	C	60	5,79
Total	9.595	100%	Total	1.036	100%

Fonte: Dados da pesquisa (Qualis Periódicos).

Apenas 7% das avaliações, aproximadamente, resultaram em classificação nos estratos A1 e A2, considerados de excelência internacional. Um segundo conjunto, com aproximadamente 35% de avaliações, obteve classificação nos estratos B1 e B2, considerados de excelência nacional. A grande maioria das avaliações, em percentual superior a 60%, obteve classificação nos estratos inferiores (B3, B4 e B5).

Já na avaliação de meio termo (2019), cada periódico foi avaliado em uma única área (área mãe), sendo identificadas 1.036 avaliações (21 não foram avaliados e 5 foram avaliados como não periódicos). Percebe-se uma melhora nos quantitativos classificados nos estratos mais elevados, sendo que aproximadamente a metade das publicações foi avaliada nos estratos A, sendo quase 25% nos estratos A1 e A2.

Considerando a mudança de metodologia do novo Qualis, a comparação entre o desempenho geral dos periódicos resta, em parte, prejudicada. Pode-se afirmar, contudo, que na antiga metodologia de avaliação houve uma concentração maior de classificações nos estratos B3, B4 e B5, provavelmente decorrentes de avaliações de periódicos em áreas de pouca aderência. Por outro lado, na recente avaliação de meio termo, há uma distribuição uniforme no percentual de periódicos classificados em cada estrato, especialmente nos estratos A e nos estratos B1 e B2.

Diante disso, infere-se que a metodologia de classificação de periódicos em uma única área (área mãe) é mais coerente que a anterior. Demanda um número inferior de análises (uma por periódico). Outra melhoria detectada é a nova estratificação, que divide as categorias A e B em um mesmo número de estratos, nos quais os periódicos são classificados com base em sua posição acima ou abaixo da mediana.

4.3 Vínculo institucional

A Tabela 4 mostra a distribuição dos periódicos brasileiros indexados no GSM conforme a natureza do vínculo institucional. Aproximadamente a metade de todos os periódicos analisados é mantido por universidades públicas federais ou estaduais, instituições tradicionalmente responsáveis pela maior parte da produção científica nacional.

Tabela 4 – Distribuição (nº e %) dos periódicos conforme o vínculo institucional

Vínculo institucional	nº	%
Instituição federal de ensino	316	29,73
Instituição privada de ensino	273	25,68
Instituição estadual de ensino	194	18,25
Sociedade ou associação	183	17,22
Fundação de pesquisa	32	3,01
Editora comercial	28	2,63
Órgão público	23	2,16
N/A	14	1,32
Total	1.063	100

Fonte: Dados da pesquisa (UlrichsWeb).

Além disso, observa-se a ocorrência de 25% de publicações mantida por instituições privadas de ensino, o que indica a origem acadêmica de três quartos dos periódicos analisados. Sociedades científicas e associações de pesquisa e profissionais também mantêm um percentual significativo de publicações,

mantendo a tradição histórica dessas instituições na editoração científica. Já as fundações de pesquisa, as editoras comerciais e os órgãos públicos são os vínculos institucionais de menor ocorrência.

4.4 Idiomas e periodicidade

A Tabela 5 mostra que mais da metade dos periódicos publica trabalhos somente no idioma português, o que pode estar relacionado à origem acadêmica das publicações e à abrangência local e nacional dos temas de pesquisa. As áreas de CHS, CSA e LLA são as que possuem maior percentual de publicações somente em português, resultado que acompanha uma tendência de publicações de trabalhos dessas áreas no idioma nativo do país de origem.

Por outro lado, CAB, CSM e EGT são as áreas que registram maior percentual de periódicos que adotam somente o idioma inglês ou o português e o inglês, o que sugere a busca pela maior visibilidade e, conseqüentemente, maior impacto do conteúdo publicado neste idioma. Destaca-se, nesse sentido, que CAB e CSM são as áreas com os periódicos de maiores médias de índice h5 e mediana h5 (Tabela 1), o que sugere uma correlação entre publicação em inglês e índices de impacto de citação mais elevado.

Tabela 5 – Distribuição dos periódicos por idioma, conforme a área (%)

Idioma	CHS	CSM	CSA	CAB	LLA	CET	EGT	Geral
Português	63,00	52,31	64,79	43,36	61,90	44,94	52,38	54,67
Português, inglês e espanhol	9,00	5,09	14,08	10,62	4,76	19,10	11,90	10,65
Português e inglês	3,25	11,57	5,16	17,70	7,14	5,62	11,90	8,91
Inglês	1,50	8,33	3,76	12,39	2,38	7,87	8,33	6,37
Português e espanhol	6,00	5,09	5,16	1,77	2,38	4,49	4,76	4,24
Vários idiomas	17,25	17,59	7,04	14,16	21,43	17,98	10,71	15,17
Total	100%	100%	100%	100%	100%	100%	100%	100%

Fonte: Dados da pesquisa (UlrichsWeb)

Pode-se afirmar que o conjunto de periódicos analisados inclui publicações de variados níveis e em diferentes estágios de desenvolvimento editorial. Logo, os dados apresentados refletem o grau de internacionalização dos periódicos brasileiros, um processo ainda em curso e que não atingiu a totalidade de publicações. A falta de familiaridade de pesquisadores com outros idiomas, sobretudo o inglês, é notório no país, fato que certamente reflete nos resultados apresentados.

A Tabela 6 mostra a distribuição percentual de periódicos de acordo com a periodicidade e a área do conhecimento.

Tabela 6 – Distribuição (%) dos periódicos por periodicidade, conforme a área

	CHS	CSM	CSA	CAB	LLA	CET	EGT	Geral
Semestral	65,00	23,72	47,66	30,77	76,67	42,53	49,40	47,96
Trimestral	10,75	35,81	21,03	33,33	5,56	22,99	22,89	21,77
Quadrimestral	14,5	17,67	23,36	11,11	10,00	19,54	19,28	16,50
Bimestral	0,75	13,02	3,27	6,84	5,56	2,30	1,20	4,71
Anual	6,00	3,72	2,34	6,84	0,00	10,34	1,20	4,35
Irregular	1,75	1,40	0,93	3,42	1,11	0,00	3,61	1,75
Contínua	1,00	2,79	0,47	4,27	0,00	1,15	1,20	1,56
Mensal	0,25	1,86	0,93	3,42	0,00	1,15	1,20	1,26
Total	100%	100%	100%	100%	100%	100%	100%	100%

Fonte: Dados da pesquisa (UlrichsWeb)

Na análise da média geral, a periodicidade mais frequente é a semestral, com quase a metade dos periódicos, seguida das periodicidades trimestral e quadrimestral. Na análise por área do conhecimento, nota-se que as áreas com periodicidade mais frequente são CSM e CAB, que possuem um maior percentual de periódicos trimestrais em relação ao periodicidades mais longas. Por outro lado, LLE, CHS e CSA são as áreas com periodicidade mais longa, reflexo do maior número de publicações semestrais.

Essa periodicidade mais curta reflete em maior produtividade, favorecendo a publicação de 100 ou mais trabalhos no período de cinco anos, critério exigido para indexação do GSM. Nesse sentido, é possível que a produtividade inferior ao exigido impediu a indexação de um número maior de periódicos brasileiros, mesma constatação feita por López-Cózar, Orduña-Malea e Martín-Martín (2019) em relação à periódicos espanhóis.

4.4 Registros duplicados

Embora a identificação de duplicações não seja um dos objetivos desse trabalho, tarefa que exigiria metodologia específica, reputou-se necessário apresentar os registros duplicados detectados durante a coleta de dados, conforme a Tabela 7.

Observa-se que quase a totalidade dos registros apresentam alguma distinção entre os títulos duplicados, resultado do processo de coleta automatizada de metadados e ausência de um sistema eficiente de limpeza dos dados.

Os 19 registros duplicados representam apenas 1,78% do total analisado, quantidade inferior ao percentual de 6,5% de duplicações em periódicos espanhóis (LÓPEZ-CÓZAR; MARTÍN-MARTÍN, 2019). Assim, considerando a similaridade estatística entre aquela e a presente pesquisa, acredita-se que o número de duplicações de periódicos brasileiros no GSM seja significativamente superior.

Tabela 7 – Registros de periódicos duplicados

ISSN	Títulos duplicados	h5	m5
1982-5765	Avaliação: Revista da Avaliação da Educação Superior	19	28
	Avaliação: Revista da Avaliação da Educação Superior (Campinas)	10	22
1676-0603	Biota Neotropica	14	18
	Biota Neotropica (Edicao em Ingles)	10	14
2177-6210	Educação Unisinos	9	11
	Educação Unisinos	2	2
1517-7793	GEOgraphia	5	7
	GEOGRAPHIA-UFF	3	4
2358-4114	Gragoatá	5	6
	GRAGOATA-UFF	3	4
2178-1494	Revista Estudos Históricos	8	14
	Estudos Históricos (Rio de Janeiro)	6	10
1809-4783	Informação & Sociedade	8	11
	Informacao & Sociedade-estudos	4	6
2236-0441	Nuances: estudos sobre Educação	8	10
	Nuances-estudos Sobre Educacao	4	6
1807-1929	Revista Brasileira de Engenharia Agrícola e Ambiental	15	18
	Revista Brasileira de Engenharia Agrícola e Ambiental-Agriambi	20	25
2177-3491	Ceres	14	19
	Revista Ceres	14	18
2316-9044	Revista de Direito Sanitário	8	13
	REVISTA DE DIREITO SANITARIO-JOURNAL OF HEALTH LAW	2	2
1980-8178	Revista de Patologia Tropical	9	13
	Revista de Patologia Tropical/Journal of Tropical Pathology	8	12
2236-6040	RITUR-Revista Iberoamericana de Turismo	7	8
	Revista Iberoamericana de Turismo (RITUR)	4	6
2316-7521	Revista Percurso	2	4
	Percurso	3	4
1984-0063	Sleep Science	18	28
	Sleep science (Sao Paulo, Brazil)	18	28
2236-9937	TEOLITERARIA-Revista de Literaturas e Teologias	2	2
	Teoliteraria-Revista Brasileira de Literaturas e Teologias	1	1
1980-265X	Texto & Contexto-enfermagem	20	24
	Texto & Contexto	3	4
2358-6958	Urdimento-Revista de Estudos em Artes Cênicas	8	11
	Urdimento - Revista de Estudos em Artes Cênicas	1	37
2179-8699	Veredas do Direito: Direito Ambiental e Desenvolvimento Sustentável	9	11
	Veredas do direito	6	8

Fonte: Dados da pesquisa (GSM).

A expressão “revista” antes do título, a cidade da publicação entre parenteses e o uso ou não de siglas e subtítulos são alguns dos elementos de distinção entre os títulos duplicados. No caso do periódico *Biota Neotropica* (1676-0603), ambos registros possuem os trabalhos do índice h5 no idioma inglês, o que indica erro na indicação de edição específica deste idioma.

5 CONSIDERAÇÕES FINAIS

O aumento da oferta de serviços de avaliação de científica por parte da Google, materializada por meio do GSC e do GSM, é um fato positivo para a comunidade científica. A gratuidade e a facilidade de uso dessas ferramentas contribuem para popularização das bases de dados e os indicadores bibliométricos para os usuários em geral, parte deles não familiarizada com uso de recursos disponíveis somente mediante assinatura.

Contatou-se que o GSM indexa aproximadamente mais da metade dos periódicos científicos brasileiros, o que representa uma cobertura significativamente superior ao JRC e ao Scopus. A política não elitista de indexação adotado pelo GS possibilita a avaliação de impacto de periódicos usualmente não abrangidos pelas bases de dados comerciais, seja em razão do escopo temático, seja em razão da qualidade científica ou editorial inferior. A partir disso, afirma-se que o GSM é a base de dados que oferece um indicador de impacto de citação com cobertura mais ampla de periódicos científicos brasileiros.

Entretanto, o uso do GSM com fonte oficial de avaliação, conforme proposto pela Capes, deve ser realizada com cautela. A impossibilidade de pesquisa por ISSN dificulta recuperação de publicações com muitas variações do título, bem como a identificação exata de periódicos homônimos. Logo, deve ser tentadas múltiplas formas de pesquisa, visando garantir que os títulos indexados sejam de fato encontrados. Além disso, a possibilidade de erros de cálculo do índice h5, seja por contagem incorreta de citações, seja por más práticas ou por outros motivos, torna necessária uma análise mais detalhada das citações que compõe esse índice em cada publicação.

Destaca-se, contudo, que a não padronização das entradas por título atinge também as publicações nacionais, o que de certo modo afeta a credibilidade dessa base de dados como fonte oficial de avaliação científica, conforme proposto pela Capes. Essa questão, tanto no que tange ao percentual de duplicações em periódicos brasileiros, quanto da diminuição dos índices h5 dos títulos duplicados não puderam ser aprofundadas nesta pesquisa, o que torna recomendável uma investigação específica.

Além disso, a questão das autocitações, aspecto sensível a essa base de dados, também não pôde ser aprofundada. A possibilidade de excesso de autocitações, que podem aumentar artificialmente o índice h5 de periódicos é uma questão relevante. Recomenda-se a fixação de de parâmetros máximos de autocitação por área do conhecimento, bem como mecanismos de detecção automatizada de autocitações na lista de trabalhos que integram o índice h5 de cada publicação.

REFERÊNCIAS

BORNMANN, L. et al. The application of bibliometrics to research evaluation in the humanities and social sciences: An exploratory study using normalized Google Scholar data for the publications of a research institute. **Journal of the Association for Information Science and Technology**, v. 67, n. 11, p. 2778–2789, 1 nov. 2016. Disponível em: <http://doi.wiley.com/10.1002/asi.23627>. Acesso em: 10 maio 2020.

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES). **Aprimoramento do processo de avaliação da pós-graduação**: esclarecimentos a respeito do Qualis Periódico e avaliação da produção intelectual. jul. 2019. Disponível em: https://www.capes.gov.br/images/novo_portal/documentos/DAV/avaliacao/18072019_Esclarecimentos_Qualis2.pdf. Acesso em: 25 abr. 2020.

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES). **Relatório do Qualis Periódicos: Área 31: COMUNICAÇÃO E INFORMAÇÃO**. 2019. Disponível em: https://www.capes.gov.br/images/Relatorio_qualis_periodicos_referencia_2019/Relatorio_qualis_comunicacao_informacao.pdf. Acesso em: 25 abr. 2019.

COSTA, H.; CANTO, F.L. DO; PINTO, A. L. Google Scholar Metrics e a proposta do novo Qualis: impacto dos periódicos brasileiros de Ciência da Informação. **Informação & Sociedade: Estudos**, v. 30, n. 1, 31 mar. 2020. Disponível em: <https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/50676>. Acesso em: 10 abr.2020.

GOOGLE SCHOLAR. **Google Scholar Metrics**. jul., 2019. Disponível em: https://scholar.google.com/citations?view_op=top_venues&hl=en. Acesso em: 20 abr. 2020.

GU, X.; BLACKMORE, K. Characterisation of academic journals in the digital age. **Scientometrics**, v. 110, n. 3, p. 1333–1350, 3 mar. 2017. Disponível em: <http://link.springer.com/10.1007/s11192-016-2219-4>. Acesso em: 10 mar. 2020.

HARZING, A.-W. A longitudinal study of Google Scholar coverage between 2012 and 2013. **Scientometrics**, v. 98, n. 1, p. 565–575, 5 jan. 2014. Disponível em: <http://link.springer.com/10.1007/s11192-013-0975-y>. Acesso em: 14 mar. 2020.

HIRSCH, J. E. An index to quantify an individual's scientific research output. **Proceedings of the National Academy of Sciences**, v. 102, n. 46, p. 16569–16572, 15 nov. 2005. Disponível em: www.pnas.org/cgi/doi/10.1073/pnas.0507655102. Acesso em: 10 mar. 2020.

JACSÓ, P. Google Scholar duped and deduped – the aura of “robometrics”. **Online Information Review**, v. 35, n. 1, p. 154–160, 22 fev. 2011. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/14684521111113632/full/html>. Acesso em: 10 mar. 2020.

JACSÓ, P. Google Scholar Metrics for Publications: the software and content features of a new open access bibliometric service. **Online Information Review**, v. 36, n. 4, p. 604–619, 3 ago. 2012. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/14684521211254121/full/html>. Acesso em: 10 mar. 2020.

JACSÓ, P. Metadata mega mess in Google Scholar. **Online Information Review**, v. 34, n. 1, p. 175–191, 23 fev. 2010. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/14684521011024191/full/html>. Acesso em: 10 mar. 2020.

LÓPEZ-CÓZAR, E.; CABALLERO, R. R. El impacto de las revistas de comunicación: comparando Google Scholar Metrics, Web of Science y Scopus. **Comunicar**, p. 45–52, 2013. Disponível em: <http://rabida.uhu.es/dspace/handle/10272/6971>. Acesso em: 10 mar. 2020.

LÓPEZ-CÓZAR, E.; CABEZAS-CLAVIJO, Á. Google Scholar Metrics: an unreliable tool for assessing scientific journals. **El Profesional de la Información**, v. 21, n. 4, p. 419–427, 1 jul. 2012. Disponível em: <https://recyt.fecyt.es/index.php/EPI/article/view/epi.2012.jul.15>. Acesso em: 10 mar. 2020.

LÓPEZ-CÓZAR, E.; CABEZAS-CLAVIJO, Á. Ranking journals: Could Google Scholar Metrics be an alternative to journal citation reports and Scimago journal rank? **Learned Publishing**, v. 26, n. 2, p. 101–114, 1 abr. 2013. Disponível em: <http://doi.wiley.com/10.1087/20130206>. Acesso em: 16 mar. 2020.

LÓPEZ-CÓZAR, E.; CABEZAS-CLAVIJO, Á. Ranking journals: could Google Scholar Metrics be an alternative to Journal Citation Reports and Scimago Journal Rank? **Learned Publishing**, v. 26, n. 2, p. 101–113, 1 abr. 2013. Disponível em: <http://doi.wiley.com/10.1087/20130206>. Acesso em: 10 mar. 2020.

LÓPEZ-CÓZAR, E.; MARTÍN-MARTÍN, A. Índice H de las revistas científicas españolas según Google Scholar Metrics (2014-2018). 4. ed. Granada, 30 out. 2019. Disponível em: <https://doi.org/10.13140/RG.2.2.36649.13923>. Acesso em: 10 mar. 2020.

LÓPEZ-CÓZAR, E.; ORDUÑA-MALEA, E; MARTÍN-MARTÍN, A. Google Scholar as a data source of research assessment. In: GLÄNZEL, W. et al. (eds.). **Springer Handbook of Science and Technology Indicators**. Cham (Suíça): Springer International Publishing, 2019. p. 95-128. Disponível em: http://link.springer.com/10.1007/978-3-030-02511-3_14 <http://link.springer.com/10.1007/978-3-030-02511-3>. Acesso em: 10 mar. 2020.

LÓPEZ-CÓZAR, E. D; ROBINSON-GARCÍA, N.; TORRES-SALINAS, D. The Google Scholar Experiment: how to index false papers and manipulate bibliometric indicators. **Journal of the Association for Information Science and Technology**, v. 65, n. 3, p. 446-54, 2014. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/asi.23056>. Acesso em: 20 abr. 2020.

MINGERS, J.; MEYER, M. Normalizing Google Scholar data for use in research evaluation. **Scientometrics**, v. 112, n. 2, p. 1111–1121, 22 ago. 2017. Disponível em: <http://link.springer.com/10.1007/s11192-017-2415-x>. Acesso em: 10 mar. 2020.

ORDUÑA-MALEA, E.; LÓPEZ-CÓZAR, E. Google Scholar Metrics evolution: an analysis according to languages. **Scientometrics**, v. 98, n. 3, p. 2353–2367, 23 mar. 2014. Disponível em: <http://link.springer.com/10.1007/s11192-013-1164-8>. Acesso em: 10 mar. 2020.

ORTEGA, J. L. How is an academic social site populated? A demographic study of Google Scholar Citations population. **Scientometrics**, v. 104, n. 1, p. 1–18, 9 jul. 2015. Disponível em: <http://link.springer.com/10.1007/s11192-015-1593-7>. Acesso em: 10 mar. 2020.

TEIXEIRA DA SILVA, J. A. The Google Scholar h-index: useful but burdensome metric. **Scientometrics**, v. 117, n. 1, p. 631–635, 20 out. 2018. Disponível em: <http://link.springer.com/10.1007/s11192-018-2859-7>. Acesso em: 10 mar. 2020.

WALTMAN, L. A review of the literature on citation impact indicators. **Journal of Informetrics**, v. 10, n. 2, p. 365-391, 2016. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1751157715300900>. Acesso em: 02 abr. 2020.