

Panorama da produção científica do Brasil além da indexação: uma análise exploratória da comunicação em periódicos

A panorama of Brazilian scientific production beyond indexing: an exploratory analysis of communication in journals

Rogério MUGNAINI¹  0000-0001-9334-3448

Rafael Jeferson Pezzuto DAMACENO²  0000-0003-4910-1534

Luciano Antonio DIGIAMPIETRI³  0000-0003-4890-1548

Jesús Pascual MENA-CHALCO²  0000-0001-7509-5532

Resumo

Indicadores de produção científica são usualmente utilizados para a avaliação da comunidade acadêmica. Como parâmetro de qualidade, as bases de dados bibliográficas são empregadas para a seleção dos periódicos mais importantes, o que as confere um papel significativo no processo avaliativo. Por outro lado, ao se restringir as análises à produção indexada, abre-se mão do todo, correndo-se o risco de desqualificar a produção em periódicos não indexados. O Qualis busca contornar esse problema, ao possibilitar que as áreas valorizem os periódicos não indexados, mas nem sempre consideram o volume de artigos publicados neles. Este trabalho apresenta uma análise exploratória da dispersão da produção científica dos 260.663 pesquisadores, no nível de doutorado, registrados na Plataforma Lattes. Para a determinação do panorama nacional, foram considerados todos os artigos completos publicados em periódicos por esses pesquisadores entre os anos de 1998 e 2016. Esse panorama contempla análises discriminadas por grande área de atuação dos pesquisadores, o país de publicação e a indexação das publicações nos diferentes periódicos. Observaram-se tendências de internacionalização, assim como a importância de periódicos nacionais como veículos de publicação de parte significativa da produção científica de algumas áreas. Finalmente, ficou evidenciada a potencial limitação de estudos que não considerem a produção científica em periódicos não indexados, ou estudos restritos às bases Scopus e/ou *Web of Science*, ignorando a *Scientific Electronic Library Online*.

Palavras-chave: Bases bibliográficas. Bibliometria. Periódicos científicos. Produção.

¹ Universidade de São Paulo, Escola de Comunicações e Artes, Programa de Pós-Graduação em Ciência da Informação. Av. Professor Lúcio Martins Rodrigues, 443, Butantã, 05508-020, São Paulo, SP, Brasil. Correspondência para/Correspondence to: R. MUGNAINI. E-mail: <mugnaini@usp.br>.

² Universidade Federal do ABC, Centro de Matemática, Computação e Cognição, Programa de Pós-Graduação em Ciência da Computação. Santo André, SP, Brasil.

³ Universidade de São Paulo, Escola de Artes, Ciências e Humanidades, Programa de Pós-Graduação em Sistemas de Informação. São Paulo, SP, Brasil.

Apoio: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (Processo nº 2012/00255-6).

Recebido em 13 de março de 2019, versão final reapresentada em 28 de junho de 2019 e aprovado em 30 de agosto de 2019.

Como citar esse artigo/How to cite this article

Mugnaini, R. et al. Panorama da produção científica do Brasil além da indexação: uma análise exploratória da comunicação em periódicos. *Transinformação*, v.31, e190033, 2019. <http://dx.doi.org/10.1590/2318-0889201931e190033>



Abstract

Indicators of scientific production are generally used in the evaluation of the academic community. As a quality parameter, the bibliographic databases are used for the selection of the most important scientific journals, a relevant role in the evaluation process. On the other hand, when the analysis is restricted to the indexed production, there is a risk of disqualifying production in journals that are not indexed. Qualis aims to address this issue by enabling areas to value non-indexed journals, even as they do not always consider the volume of articles published in these journals. This paper presents an exploratory analysis of the dispersion of the scientific production of 260,663 researchers with doctorate level registered in the Lattes Platform. For the determination of the national panorama, we considered all the full papers they published in scientific journals between 1998 and 2016. This panorama includes analyzes divided by the researchers' knowledge areas, the publications' countries, as well as the indexation of the publications in the different journals. Internationalization trends were observed, as well as the importance of national journals as vehicles for the publication of a significant part of the scientific production in some areas. Additionally, we observed the limitation of studies that do not consider scientific production in non-indexed journals, or studies restricted to Scopus and/or Web of Science databases that do not consider Scientific Electronic Library Online.

Keywords: Bibliographic databases. Bibliometrics. Scientific periodicals. Production.

Introdução

A utilização de periódicos científicos e sua centralidade no processo de comunicação científica é um tema de muita discussão na literatura científica, seja pela questão da circulação e acesso ao conhecimento, seja pela necessidade de aferição dos logros da pesquisa, conforme enfocados em processos avaliativos de diversos níveis (Davyt; Velho, 2000; Guédon, 2010). A consolidação do periódico é reforçada também por fontes secundárias (como serviços de indexação e resumo), que se originaram há décadas (Cendón, 2000), se desenvolvendo sobremaneira, resultando em bases de dados bibliográficas e oferecendo insumo não apenas para processos avaliativos (Neuhaus; Daniel, 2008; Mckiernan *et al.*, 2019). Por outro lado, os livros não tiveram o mesmo tratamento. Apresentaram, bem tardiamente, iniciativas, muitas delas buscando replicar a trajetória dos periódicos, porém sem repercutir de maneira semelhante (Gorraiz; Purnell; Glänzel, 2013) – mesmo com os esforços empreendidos no contexto brasileiro (Menandro *et al.*, 2011; Lima, 2013) –, e enfrentando diversos entraves para trilhar a rota do acesso aberto (Gimenéz-Toledo, 2018). Assim, a importância do periódico como veículo de comunicação na ciência se mantém, alterando práticas e culturas de áreas de conhecimento mais afeitas à comunicação no formato de livro ou de capítulo de livro.

Desde a década de 1990, o desenvolvimento das tecnologias de informação viabilizou o acesso aos periódicos, apesar de esse acesso ser possível quase exclusivamente de forma paga. O problema do acesso representou grandes dificuldades para pesquisadores de países não centrais, os quais, embora se beneficiassem do formato eletrônico e da disponibilidade *online*, se depararam com outras barreiras, que implicavam altos custos e direitos autorais (Russell, 2001). Esses impasses deram lugar ao movimento de Acesso Aberto, o qual, mesmo já estando em marcha desde aquela década, inicialmente consistindo de esforços de pesquisadores criando e disponibilizando periódicos individuais, posteriormente teve adesão de editoras profissionais – que alteraram a forma de pagamento, transferindo o custo de publicação para o autor –, e há quase uma década é marcado pela adesão de editoras comerciais – que garantem a disponibilidade de artigos individuais mediante cobrança de taxa (Björk *et al.*, 2010).

Em um âmbito mais abrangente, o movimento da Ciência Aberta atenta não apenas para a disponibilidade da informação publicada e seu processo de comunicação, mas para o modo de produção da pesquisa, que, conforme observam Albagli, Clinio e Raychtock (2014), vê-se submetido a obstáculos legais e econômicos que impedem sua circulação, avanço e difusão. Nesse contexto, os dados da pesquisa passam a ser disponibilizados, viabilizando tanto a reutilização como o escrutínio, podendo-se certificar a integridade das evidências obtidas a partir dos dados de pesquisa.

Assim, constata-se atualmente diversas iniciativas em prol da abertura das fontes de informação (bases de dados bibliográficas, repositórios, bases de currículos, diretórios, catálogos, dentre outras), fazendo com que não somente o artigo, mas todo o aparato tecnológico que o envolve, seja acessível. Dessa maneira, essas fontes – ainda que timidamente –, passam a facilitar a descarga de dados, que permitirão a inserção de maior transparência nos procedimentos metodológicos de estudos bibliométricos e cientométricos. O que se quer esclarecer é que a restrição não diz respeito apenas à natureza comercial das bases de dados comumente usadas – como *Web of Science* (WoS) ou Scopus –, mas às barreiras impostas para descarga e tratamento dos dados. Por essa razão, os estudos bibliométricos permanecem com um longo caminho a percorrer, rumo ao que postula a Ciência Aberta.

Mas é importante considerar que o contexto internacional tem evidenciado um olhar para além das publicações e mirado nas citações, que consiste em um outro aspecto relacionado à pesquisa científica e que há muito tem sido alvo de mensuração. Merece destaque a Iniciativa para Citações Abertas (Projeto I4OC), a qual convocou editoras a disponibilizar as listas de referências de seus artigos. Por meio das informações armazenadas pelo Crossref, viabilizou-se, logo de início, cerca de 500 milhões de referências, graças ao papel dessa empresa sem fins lucrativos na atribuição dos Indicadores de Objetos Digitais (Shotton, 2018).

Na esteira dessa iniciativa, outras fontes já vêm surgindo há algum tempo, combinando, da mesma forma, metadados da Crossref e do PubMed: Lens.org, que também utiliza o *Microsoft Academic* e ainda cruza dados de citações em patentes provenientes de várias fontes; e Scilit, cujo nome mescla as palavras “literatura” e “científica”. Algumas outras menções se fazem igualmente necessárias: o *Digital Science Dimensions*, o qual, além de se destacar pela cobertura, que abrange não só a Scopus, agrega dados de patentes, métricas alternativas, entre outras; o 1Findr, cujo potencial tem se evidenciado em mapear a produção científica em Acesso Aberto, bem como prever a possibilidade de realização de buscas nas referências dos artigos; e o antes mencionado *Microsoft Academic*, o qual, ademais de oferecer a análise de citações, permite a busca semântica.

No Brasil, os estudos bibliométricos contaram com fontes de informação diversas, como (sem a pretensão de ser exaustivo) PubMed e Lilacs, no caso de áreas da saúde (Packer; Tardelli; Castro, 2007); *Scientific Electronic Library Online* (SciELO), para a América Latina (Mugnaini; Digiampetri; Mena-Chalco, 2014); Plataforma Lattes, cujas informações curriculares abarcam a trajetória acadêmica no Brasil de forma mais completa (Damaceno *et al.*, 2019); bancos de teses e dissertações diversos (Araújo; Alvarenga, 2011); e o próprio Google Acadêmico, que oferece uma variedade de tipos de documentos de todo o mundo em múltiplas áreas e idiomas (Caregnato, 2011). Sem discorrer sobre a dificuldade para descarga e/ou tratamento dos dados, essas fontes têm sido utilizadas em estudos publicados tanto em teses e dissertações como em eventos e periódicos.

Se, no contexto brasileiro, a restrição imposta por bases de dados comerciais continua tão presente e a abertura da relevante informação que detêm tão impensável, o que dizer sobre a disponibilidade de citações, haja vista o custo de acessá-las e processá-las para desenhos específicos de metodologia de análise? A complexidade computacional envolvida em um plano de refinamento de um indicador simples como o Fator de Impacto (Vanclay, 2012) é um exemplo claro da impossibilidade de se propor alternativas, o que se deve principalmente à restrição do acesso aos dados.

A realidade das recentes transformações deve ser considerada para avanços dos estudos quantitativos da ciência, significando que estudos que utilizem fontes brasileiras devem primar por conhecer especificidades da produção científica nacional, não se limitando ao que oferecem as bases tradicionais, ainda que isso implique enfrentar problemas e limitações típicas de fontes de dados de natureza diversa. A vantagem consiste em acessar os dados por completo, o que evita a insegurança de se lidar com um universo obscuro e muitas vezes incompleto – diferentemente do acesso que diversos grupos internacionais têm às bases comerciais, mediante alto custo (Costas, 2017).

É preocupante o fato de que, na avaliação de programas de pós-graduação no Brasil – mais especificamente no Qualis –, as bases de dados bibliográficas comerciais representam quase a totalidade de critérios de avaliação de produção científica, servindo como: (1) fontes exclusivas de indicadores de impacto – principalmente os índices de citação Scopus e WoS –, quando se trata das áreas associadas às ditas ciências duras, crescentemente assolando as humanas e sociais; e (2) parâmetro de qualidade de periódicos, levando em conta seu processo seletivo – aspecto que se observa de maneira generalizada entre as áreas (Mugnaini, 2015).

Por essa razão, estudos bibliométricos precisam delinear um contexto mais amplo da produção científica nacional, iluminando espaços de discussão excluídos da chamada corrente principal (Beigel, 2014; Rafols, 2017; Aksnes; Sivertsen, 2019). Esses esforços visam informar a política científica nacional ou quaisquer exercícios avaliativos que insistam em ignorar a ciência que não está presente nessas bases, se furtando de avaliá-la. Cabe destacar, contudo, que quaisquer esforços deverão necessariamente apontar para abordagens bibliométricas avaliativas, as quais deverão, por sua vez, almejar a análise de impacto, o que implicará o desenvolvimento de bases de dados e ferramentas que se beneficiem do movimento de abertura da ciência e, conforme já se afirmou, que abrangem as citações.

Neste momento, este estudo, tirando proveito da abrangência oferecida pela Plataforma Lattes em relação ao sistema de ciência e tecnologia brasileiro, propõe o mapeamento da produção científica brasileira, baseada nos periódicos utilizados para publicação. Trata-se de uma análise exploratória de dispersão da produção, considerando a área dos pesquisadores, os diferentes períodos de avaliação da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e o país de publicação e indexação dos periódicos. Para tanto, recorre a informações de diversas bases de periódicos, visando identificar a indexação e corrigir informações constantes dos currículos, como forma de grafia do título, *International Serial Standard Number* (ISSN) e país de publicação. Vê-se, nesta abordagem, um passo necessário para se distinguir as características dos veículos utilizados pela comunidade científica brasileira, assim como sua importância, como base no seu uso para publicação.

Procedimentos Metodológicos

O procedimento adotado para o levantamento deste panorama está composto de quatro etapas: (1) coleta de dados, (2) deduplicação de artigos, (3) associação periódico-ISSN, e (4) agregação e classificação dos artigos. A seguir, descrevemos em detalhes todas as etapas consideradas neste trabalho.

O estudo reúne informações do currículo acadêmico de 260.663 doutores registrados na Plataforma Lattes. Os dados foram coletados em agosto de 2017 por meio do serviço *online Lattes Extrator (web service)*⁴. Esse serviço permite coletar um conjunto de currículos em formato *eXtended Markup Language* (XML), uma linguagem que organiza os dados de forma hierárquica. Através de um *software* desenvolvido sob medida na linguagem de programação Python, cada currículo foi pré-processado de forma a extrair os seguintes dados: grande área e área, ambas do campo “Áreas de Atuação”, e artigos completos publicados em periódicos, do campo “Produção Bibliográfica”. Cada publicação é composta por um ISSN, título do periódico, título do artigo e ano da publicação. Todas as informações textuais foram normalizadas, isto é, foram transformadas para caixa baixa, e os acentos e caracteres não ocidentais foram removidos.

⁴ O CNPq disponibiliza às instituições de ensino e pesquisa (entre outras) as possibilidades de extração dos dados da Plataforma Lattes usando a ferramenta Lattes Extrator. A informação sobre acordos institucionais está disponível em <http://memoria.cnpq.br/web/portal-lattes/acordos-institucionais>. Adicionalmente, todos os números identificadores Lattes contemplados neste trabalho foram disponibilizados livremente em 08 de julho de 2017 no portal da Plataforma Lattes: <http://memoria.cnpq.br/web/portal-lattes/extracoes-de-dados>. Acesso em: 6 ago, 2019.

O período de análise corresponde a seis ciclos avaliativos completos, determinados pela CAPES na avaliação dos programas de pós-graduação, cobrindo o período de 1998 a 2016. Esse período corresponde a cinco triênios e um quadriênio: 1998-2000, 2001-2003, 2004-2006, 2007-2009, 2010-2012, 2013-2016.

A produção bibliográfica foi organizada segundo oito grandes áreas do conhecimento, conforme definição do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), a saber: Ciências Agrárias (AGR), Ciências Biológicas (BIO), Ciências da Saúde (SAU), Ciências Exatas e da Terra (EXA), Ciências Humanas (HUM), Ciências Sociais Aplicadas (SOC), Engenharias (ENG), e Linguística, Letras e Artes (LIN). Essas informações correspondem à primeira grande área de atuação declarada pelos acadêmicos nos seus currículos. A Tabela 1 apresenta os dados gerais, de modo a permitir a análise do volume de informações consolidadas, de acordo com cada grande área de conhecimento.

Deduplicação de artigos

A fim de evitar duplicação na contagem dos artigos, as coautorias foram identificadas por meio de casamento exato dos títulos registrados nos artigos, caso o periódico e ano também coincidisse (isto é, um artigo em coautoria no conjunto de dados seria aquele publicado com o mesmo título, no mesmo periódico e no mesmo ano). Assim, a única duplicidade permitida decorreu do fato de dois ou mais autores pertencerem a grandes áreas diferentes, resultando na contagem de uma publicação para cada área, independentemente do número de coautores de uma mesma grande área. É importante destacar que essa etapa é fundamental para se obter a quantidade real de artigos publicados pelos pesquisadores componentes no conjunto de dados.

Associação periódico-ISSN

Algumas informações sobre os periódicos foram obtidas de outras fontes de dados, de forma a eliminar duplicidades provenientes de diferentes formas de grafia do nome do periódico, ISSNs advindos de diferentes formatos (impresso, CD-ROM, *online*) e diferentes países de publicação. A base de ISSNs⁵ foi utilizada, assim como uma lista disponível pela instituição ISSN.org, a qual permite relacionar ISSNs de um mesmo periódico por meio do ISSN-L.

Tabela 1. Universo de pesquisadores e artigos, discriminados por grande área de conhecimento – período de 1998 a 2016.

Grande área	Pesquisadores (n)	Artigos (n)	Artigos/Pesquisador(%)
Ciências Agrárias	26.479	307.704	11,62
Ciências Biológicas	35.417	381.941	10,78
Engenharias	24.890	182.516	7,33
Ciências Exatas e da Terra	38.682	348.873	9,02
Ciências Humanas	44.634	338.494	7,58
Linguística, Letras e Artes	16.241	105.592	6,50
Ciências da Saúde	44.952	584.589	13,00
Ciências Sociais Aplicadas	29.368	238.118	8,11
Total	260.663	2.487.827	9,54

Fonte: Elaborada pelos autores com dados da Plataforma Lattes em dezembro de 2018.

⁵ O país associado a cada ISSN foi extraído por meio de consultas ao *website* do *International Standard Serial Number* (2019). Uma lista com a relação entre ISSN e ISSN-L foi solicitada ao ISSN para unificar os diferentes formatos em um único número serial por periódico, independentemente do formato (denominado ISSN-L).

Adicionalmente, as fontes de informação Latindex e Ulrich foram úteis para a correta associação do país de publicação dos periódicos. Os índices de citação – SciELO, WoS/*Journal Citation Reports* e Scopus –, além de servirem para o mesmo fim (limpeza e unificação de dados), permitiram identificar se os periódicos estavam indexados neles ou não.

Não foi considerado o ano de indexação dos periódicos nos índices de citação. Independentemente de quando determinado periódico entrou em alguma das bases, o fato de ser indexado no presente faz com que todos os artigos sejam atribuídos à base, como se o periódico estivesse indexado desde a sua aparição no currículo de um pesquisador. Essa simplificação se justifica por o estudo não pretender analisar o efeito da crescente indexação ao longo do período, mas sim o volume de artigos publicados em periódicos que nunca foram indexados nessas bases.

Agregação e classificação dos periódicos e artigos

Para a análise da dispersão das publicações dos pesquisadores de cada área – e nos vários ciclos avaliativos –, foi utilizada a Lei de Dispersão de Bradford (Bradford, 1961), a fim de organizá-las em Zonas de Bradford. As zonas dividem os artigos publicados em três conjuntos de mesmo tamanho, abrangendo, contudo, diferentes quantidades de periódicos. A Zona 1 contemplou o núcleo de periódicos mais utilizados, com o menor número de periódicos; a Zona 2 teve uma quantidade intermediária de periódicos; e a Zona 3, cuja dispersão é maior, por apresentar o maior número de periódicos, teve mais baixa frequência de artigos.

Além de serem calculadas para os diferentes períodos, as Zonas de Bradford da produção científica foram determinadas para as diferentes grandes áreas do conhecimento. Não foram consideradas as categorias temáticas, como fazem tradicionalmente os estudos que utilizam bases de dados, e sim a grande área de atuação do pesquisador registrada no próprio currículo. Por essa razão, um periódico foi classificado em tantas áreas quanto houvesse sido utilizado para publicação. Isso porque o estudo pretende analisar a dispersão da produção científica da área em todos os periódicos utilizados por ela.

De posse da identificação das Zonas de Bradford, procedeu-se à análise da dispersão da produção de cada área tendo em vista duas variáveis: nacionalidade e indexação dos periódicos. Para a nacionalidade do periódico, considerou-se o país de publicação segundo as seguintes categorias: Brasil (periódicos editados no Brasil), América Latina e Caribe (exceto Brasil), ou Outros (demais países não considerados nas categorias anteriores). Para a indexação, foram consideradas três categorias: SciELO (periódicos indexados na SciELO, estando ou não indexados em outras bases), Scopus/WoS (periódicos não indexados na SciELO e indexados na Scopus e/ou na WoS), e não indexados (periódicos não indexados em nenhuma das três bases anteriores).

Resultados

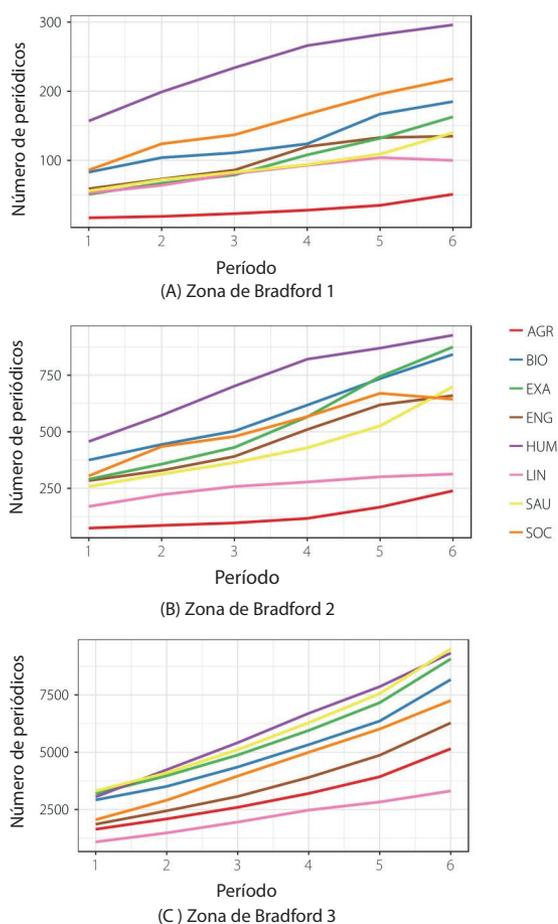
A dispersão da produção científica em periódicos, segundo as diferentes Zonas de Bradford, revela particularidades de cada grande área, à medida que se pode observar o tamanho do conjunto de periódicos que compõem o núcleo. Como observado na Tabela 1, a área SAU apresenta a maior quantidade de pesquisadores e artigos, sendo seguida por HUM, no que diz respeito a pesquisadores, mas por BIO quando se consideram artigos. Apesar do volume dessas áreas, nota-se, na Figura 1, que HUM ostenta produção em um maior número de periódicos, em cada uma das zonas, sendo superada por SAU na Zona 3 apenas no último período. É importante destacar ainda a área de SOC, com o segundo maior volume de periódicos na zona núcleo, e comportamento parecido a BIO nas demais zonas.

No outro extremo, tem-se AGR e LIN, que evidenciam mais concentração de produção em cada uma das zonas (Figura 1). Esses aspectos são compreensíveis à medida que, conforme se observou na Tabela 1, são as áreas

com menor quantidade de periódicos e, no caso de LIN, também de pesquisadores e artigos. Deve-se considerar que boa parte da produção da área de LIN é publicada em livros (Fukahori, 2017), assim como se nota nas Ciências Sociais e Humanas em geral (Sivertsen, 2019). Essa realidade pode advir de uma especificidade das áreas, a talvez um menor percentual de interdisciplinaridade, dado que, como apontam Mena-Chalco *et al.* (2014), apenas 10% dos pesquisadores de LIN estão conectados a outros pesquisadores (de LIN ou demais grandes áreas) em termos de coautoria científica.

As áreas de ENG e EXA apresentam perfis bastante similares na Zona 1, havendo um distanciamento de EXA na Zona 2, a partir do período 5, e na Zona 3, ao longo de todo o período. Depreende-se que a produção em EXA se distribui para um número maior de periódicos, haja vista que seu número de pesquisadores e artigos é significativamente maior (Tabela 1). Cabe pesar ainda a tendência de ENG de publicar em anais de eventos (Mueller, 2005), o que também explica a menor quantidade de periódicos.

Por fim, é importante observar os números absolutos no último período da Zona 1, com 296 periódicos que publicam um terço dos artigos de HUM, seguido por 218 periódicos de SOC. Já na Zona 2, HUM apresenta 927 periódicos, seguida por EXA, com 875. E, na Zona 3, lideram SAU e HUM, havendo publicado em, respectivamente,



9.502 e 9.322 periódicos no último período. Como se pode notar, a tendência é o número de periódicos aumentar ao longo dos períodos. Contudo, ao fazer uma comparação com os resultados de Mugnaini, Digiampietri e Mena-Chalco (2014), cujos dados se restringiam às bases SciELO e WoS, as Ciências Humanas não apresentam essa magnitude nas Zonas 1 e 2, indicando que periódicos não indexados são muito utilizados nessas áreas, como se verificará adiante.

Por outro lado, a proporcionalidade entre o tamanho das zonas também permite diferenciar as áreas (dados não apresentados). Para a área de AGR, ao longo do período, em média, a Zona 3 é cerca de 24 vezes maior que a Zona 2, denotando que o número de periódicos explode na zona de maior dispersão (o que pode advir de um esforço para publicação em novos periódicos, o que será constatado adiante). A área a seguir é SAU, com a Zona 3 em média aproximadamente 14 vezes maior que a Zona 2. A área com menor relação entre essas zonas é ENG (7,8). Da mesma forma, ao se comparar o tamanho da Zona 2 com a Zona 1 (dados não apresentados), a área de EXA tem a maior média (5,5), seguida de ENG e SAU (empatadas com 4,6). A área de HUM tem a menor média (3,0). A Zona 2 mais volumosa pode retratar um volume mais significativo de periódicos que estão se consolidando como veículo de destacada importância na área, obviamente candidatos a integrar a Zona 1 (núcleo).

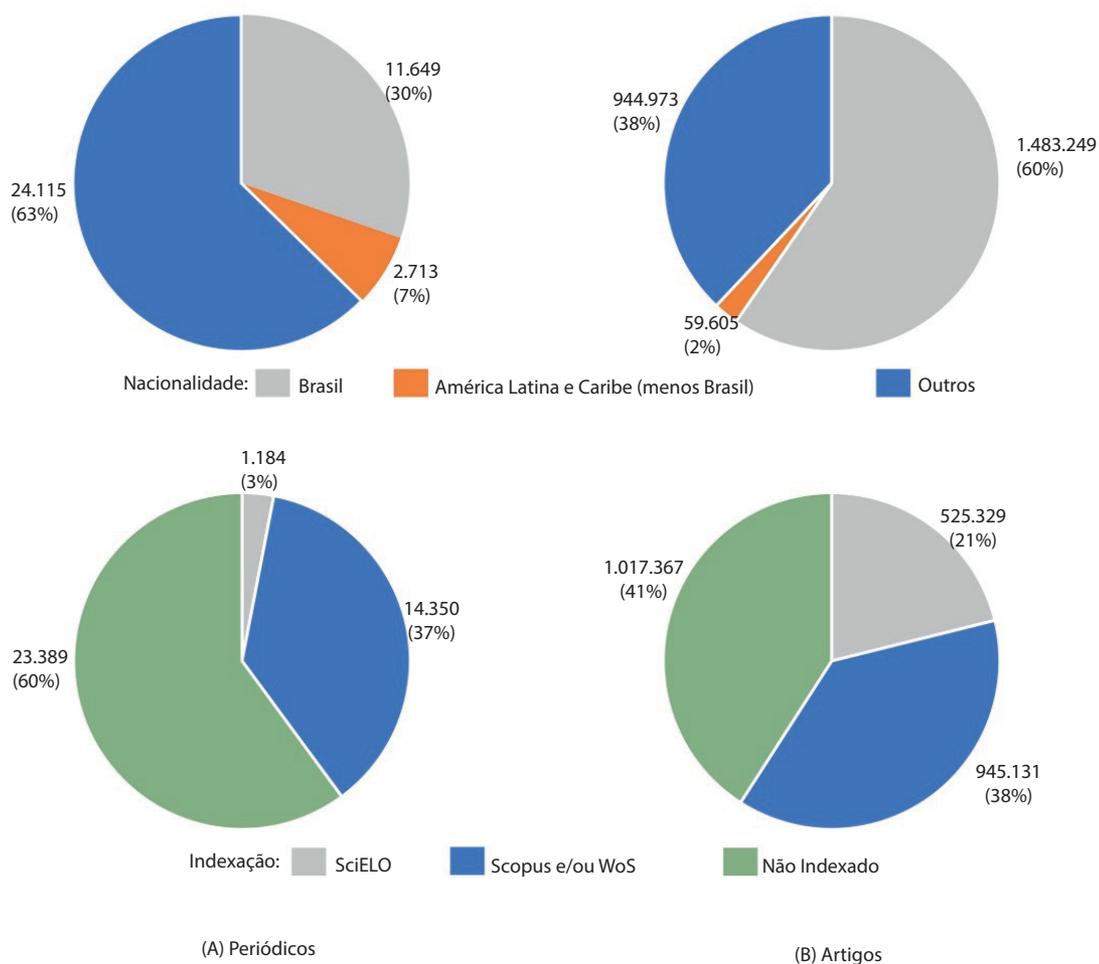


Figura 2. Percentual de (A) periódicos e (B) artigos – período de 1998 a 2016.

Nota: A linha superior está relacionada com a nacionalidade e a linha inferior está relacionada com a indexação.

Fonte: Elaborada pelos autores com dados da Plataforma Lattes em dezembro de 2018.

De acordo com as características observadas até aqui, a análise do percentual de periódicos de diferentes países e indexados em diferentes bases permitirá o delineamento de um panorama mais rico da produção científica brasileira. No entanto, antes de aprofundar a análise entre áreas e ao longo dos períodos e zonas, é fundamental conhecer o cenário das duas variáveis de interesse neste estudo (nacionalidade e indexação dos periódicos) e qual seu comportamento quando se diferenciam duas unidades de análise: os periódicos, conforme se tratou até aqui; e o volume de artigos publicados, que é o que se passa a diferenciar a partir de agora.

A Figura 2 permite constatar que se utilizam mais periódicos de fora da América Latina (azul, 63%), mas, por outro lado, eles publicam 38% dos artigos. Analogamente, periódicos latino-americanos e caribenhos (exceto Brasil) têm seu percentual diminuído (7% contra 2%) quando se comparam periódicos e artigos. Em contrapartida, os 30% de periódicos do Brasil publicam 60% dos artigos, revelando a utilidade dos periódicos domésticos como veículos de publicação dos autores brasileiros.

Ao se analisar a indexação (Figura 2), percebe-se que 60% dos periódicos não estão indexados na SciELO, no Scopus ou no WoS. Este é um dos mais importantes resultados de um estudo dessa magnitude em uma base de currículos, observando o que não é mostrado pelas bases de dados bibliográficas com critérios seletivos. No entanto, essa grande quantidade de periódicos (mais de 23 mil) é utilizada para a publicação de 41% dos artigos. Esse menor percentual de artigos dá lugar aos periódicos SciELO, os quais publicam 21% da produção nacional, haja vista corresponderem a apenas 3% dos periódicos utilizados. Finalmente, os periódicos indexados na Scopus e/ou WoS correspondem a 37% dos periódicos e 38% dos artigos.

A partir das variáveis de nacionalidade dos periódicos, os gráficos da coluna A da Figura 3 permitem observar alguns padrões entre as áreas. Primeiramente, nota-se um conjunto de áreas nas quais predomina a cor cinza (periódicos do Brasil): AGR, HUM, LIN, SAU e SOC. Dentre elas, pode-se verificar que a maior parte apresenta porção significativa da Zona 3 em azul, a qual se mantém ou cresce levemente entre os períodos, denotando que a internacionalização fora da região latino-americana e caribenha é um fenômeno presente; contudo, é percebido na zona de menor frequência. Merecem destaque as áreas AGR e SAU, cuja porção de periódicos internacionais é também nítida e tem crescimento significativo. Ainda, já ocupam essa zona na qual os periódicos estão mais consolidados na prática de comunicação da área.

Dentre as demais áreas, salienta-se a EXA como a que mais publica em periódicos fora da região latino-americana e caribenha, em todas as zonas, mas apresenta leve queda percentual ao longo dos triênios, dando lugar a periódicos do Brasil. Já ENG e BIO têm um comportamento similar, com percentual crescente de periódicos de fora da região nas Zonas 1, 2 e 3, com diminuição da diferença entre as zonas ao longo dos períodos. Esses resultados mostram que essas duas áreas têm buscado mais periódicos estrangeiros para publicação, evidenciando um claro processo de internacionalização.

Quando se parte para a análise da indexação dos periódicos, três grupos de áreas podem ser identificados, de acordo com as barras verdes (periódicos não indexados). O maior percentual é observado em HUM, LIN e SOC. Dentre estas, HUM apresenta em torno de 35% de periódicos SciELO na Zona 1, percentual este que decresce nas Zonas 2 e 3. Essa importância dos periódicos SciELO na área é refletida na própria coleção brasileira da base, que tem um percentual significativo de periódicos da área de humanas. Nesse sentido, é relevante atentar para o alerta de Hicks (2004) de que exercícios avaliativos nessas áreas deveriam contemplar os periódicos locais, assim como a literatura desenvolvida no contexto local – a baixa representatividade dos periódicos indexados nos índices de citação reforça esse argumento.

Outro grupo de áreas é formado por AGR e SAU, cujo percentual de periódicos não indexados é maior na Zona 2, seguido da Zona 3 e então da Zona 1. No entanto, com o passar dos períodos, as Zonas 2 e 3 ficam próximas e, no último triênio, se ordenam crescentemente. Esse efeito resulta de um percentual maior de periódicos indexados na Scopus e/ou WoS nas zonas. Como consequência, a Zona 1, que começa com claro predomínio de

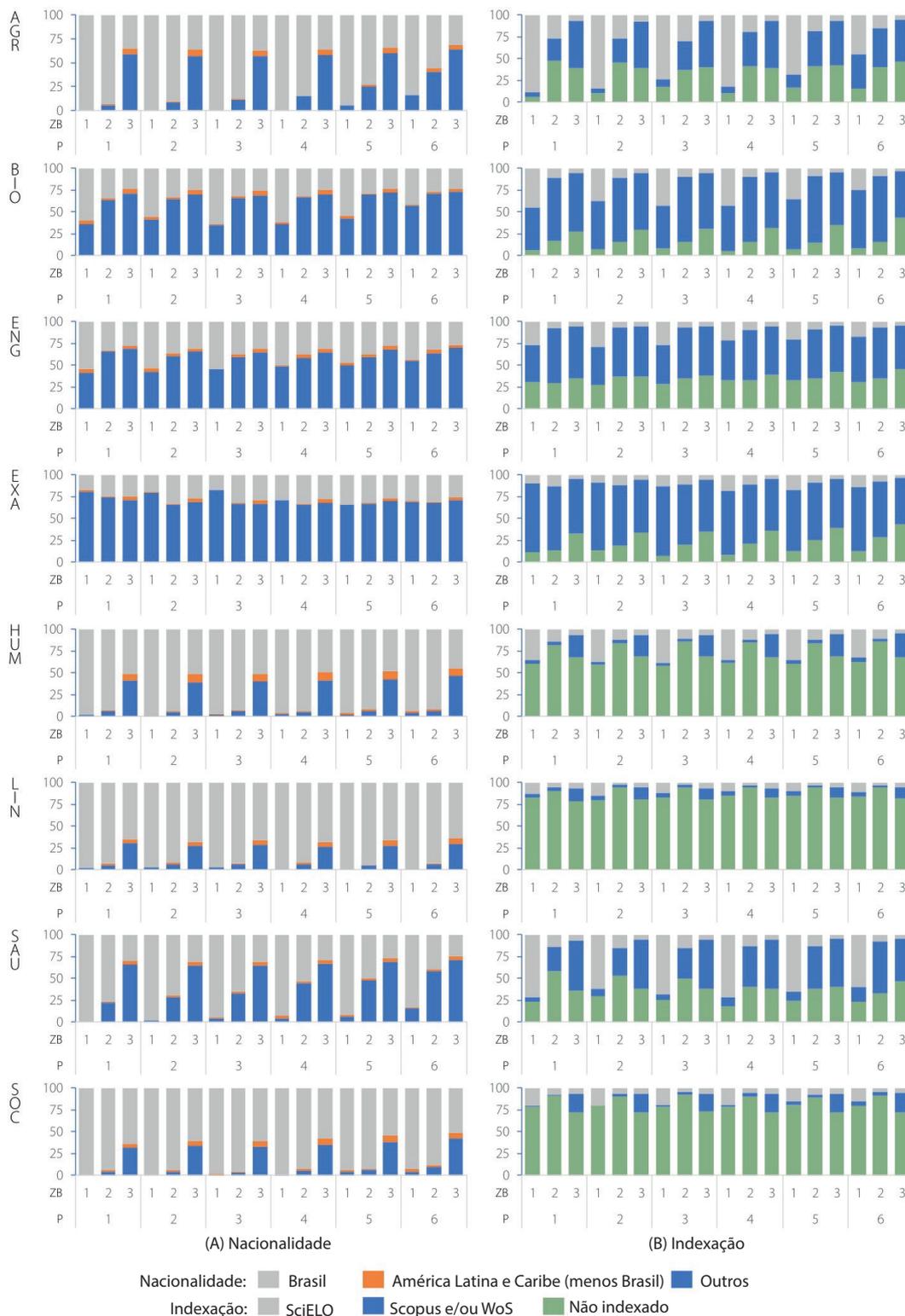


Figura 3. Percentual de periódicos por Zona de Bradford (ZB) e período, discriminado por (A) nacionalidade e (B) indexação.

Nota: AGR: Ciências Agrárias; BIO: Ciências Biológicas; ENG: Engenharias; EXA: Ciências Exatas e da Terra; HUM: Ciências Humanas; LIN: Linguística, Letras e Artes; SAU: Ciências da Saúde; SOC: Ciências Sociais Aplicadas.

Fonte: Elaborada pelos autores com dados da Plataforma Lattes em dezembro de 2018.

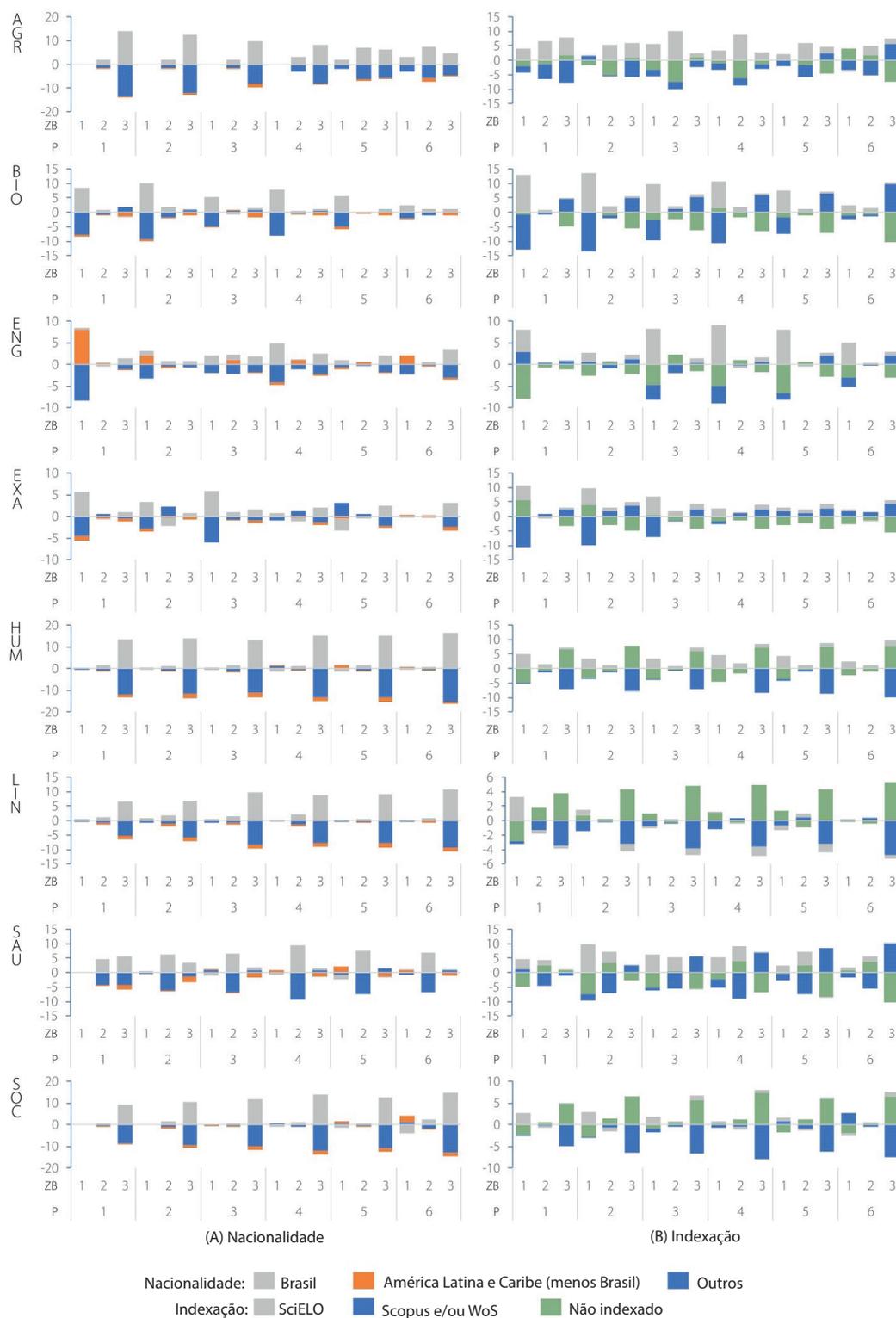


Figura 4. Diferença entre os percentuais de artigos e periódicos por Zona de Bradford (ZB) e período (P), discriminada por (A) nacionalidade e (B) indexação.

Nota: AGR: Ciências Agrárias; BIO: Ciências Biológicas; ENG: Engenharias; EXA: Ciências Exatas e da Terra; HUM: Ciências Humanas; LIN: Linguística, Letras e Artes; SAU: Ciências da Saúde; SOC: Ciências Sociais Aplicadas.

Fonte: Elaborada pelos autores com dados da Plataforma Lattes em dezembro de 2018.

periódicos SciELO, tem esse percentual significativamente diminuído ao longo dos períodos, sobretudo em AGR (e também BIO). Já em SAU, o percentual de SciELO se mantém significativo até o final do período, sendo outra área muito bem representada na coleção SciELO Brasil (Mugnaini; Digiampetri; Mena-Chalco, 2014).

O terceiro grupo é formado por EXA e ENG, com predomínio de periódicos indexados na Scopus e/ou WoS, sendo que ENG apresenta cerca de 30% de periódicos não indexados ao longo dos períodos e uma porção de periódicos SciELO na Zona 1. Em EXA, por sua vez, o percentual de não indexados é crescente à medida que as zonas aumentam. Desses resultados, pode-se depreender que os periódicos do Brasil têm um importante papel, mas os que não são indexados são observados principalmente nas zonas de maior dispersão, significando que estes não se consolidam nas áreas pela frequência de publicação – ao contrário, a consolidação se expressa na Zona 1, o que é mais marcado em EXA.

Agora, para se analisar a distribuição dos artigos, considerando as variáveis nacionalidade e indexação dos periódicos, procede-se à comparação por meio da diferença entre os percentuais dos artigos e dos periódicos (Figura 3). Essa análise – apresentada na Figura 4 –, permite observar o quanto o volume de artigos destoa do volume de periódicos. Para exemplificar, considera-se a análise da Figura 2, na qual se notou que os artigos publicados em periódicos SciELO correspondiam a 21%, ao passo que esses mesmos periódicos representavam 3%, ou seja, uma diferença de 18%.

Observa-se, na Figura 4, que as áreas de AGR, HUM, LIN e SOC publicam um maior percentual de artigos em periódicos do Brasil, principalmente na Zona 3. A exceção é que, para AGR, esse percentual é decrescente ao longo dos períodos e a diferença é positiva, porém menos pronunciada, na Zona 2, aumentando significativamente nos dois últimos períodos. No sentido oposto – com percentual de artigos menor que o de periódicos –, está a indexação na Scopus e/ou WoS, também ocorrendo sobretudo na Zona 3 e decrescendo apenas no caso de AGR. Essas áreas chamam a atenção pelo fato de revelarem diferenças entre os percentuais apenas na Zona 3, que é exatamente a zona na qual as mudanças são mais imediatas, considerando-se a baixa frequência.

Na área de SAU, apresentam-se perfis parecidos com as áreas anteriores – diferenças positivas em relação a artigos em periódicos do Brasil e negativas em relação aos de fora da região latino-americana e caribenha –, mas observadas principalmente na Zona 2. Já as áreas de BIO, ENG e EXA destacam-se pelas barras de diferenças mais significativas na Zona 1, que tendem a se dissipar no último período. Chama a atenção o fato de ENG manifestar diferença positiva de percentual de artigos em periódicos da América Latina e Caribe na Zona 1 no primeiro período, mantendo-se em menor magnitude nos períodos seguintes.

Quanto à variável indexação, evidencia-se o efeito SciELO, conforme observado na Figura 2, pronunciando-se nas seguintes áreas: AGR e SAU, em todas as zonas; BIO e ENG, nas Zonas 1; e EXA, nas Zonas 1 dos primeiros três períodos. A indexação nas bases Scopus e/ou WoS se revela positivamente e na maioria dos períodos nas Zonas 3 de BIO, ENG, EXA e SAU. Finalmente, HUM, LIN e SOC mostram diferença positiva do percentual de artigos em periódicos não indexados, sobretudo na Zona 3.

Esses resultados estão alinhados com o que apontaram Mugnaini, Digiampetri e Mena-Chalco (2014) quanto ao aumento do percentual de artigos publicados em periódicos do Brasil e, conseqüentemente, aqueles indexados na SciELO. O que se pode destacar aqui, agora sobre o volume de artigos, é que os periódicos não indexados seguem sendo muito utilizados nas Ciências Sociais e Humanas.

Contrariamente, na maioria das áreas, o percentual de artigos em periódicos indexados nas bases Scopus e/ou WoS é menor em relação ao percentual de periódicos, sendo que, nas ciências duras, se evidencia na Zona 1 e, nas demais áreas, na Zona 3. Isso deve resultar do aumento do número de periódicos nessas zonas, ou seja, a produção está dispersando para periódicos até então não utilizados naquela área. Mas se observa que, nas ciências duras, esses periódicos estão sendo inseridos na zona núcleo, a qual reúne os periódicos mais consolidados em seu processo de comunicação científica, e, ainda, que essa diferença dos percentuais apresenta pequena diminuição

ao longo dos períodos. Nas demais áreas, esses periódicos surgem na zona de maior dispersão, o que resulta do seu uso inicial ou ocasional. Contudo, a diferença dos percentuais aumenta ao longo dos períodos. Esse fato sugere que, nos últimos períodos, diversos novos periódicos estão sendo utilizados para publicação; entretanto, não se nota qualquer sinal de atingirem a Zona 2, que indicaria os primeiros indícios de consolidação.

Esses resultados podem ser fruto dos estímulos decorrentes do uso de indicadores bibliométricos para a classificação de periódicos em áreas de Ciências Sociais e Humanas, acentuados desde a trienal 2010-2012 (Mugnaini, 2015). Esses estímulos fazem com que os pesquisadores tendam a publicar nesses periódicos, mas o façam ocasionalmente; outra questão seria a integração desses periódicos no núcleo da área. Portanto, uma análise do processo avaliativo se faz necessária, já que não são recentes os alertas sobre suas influências no comportamento dos pesquisadores (Butler, 2007), sendo de extrema importância que as áreas avaliem os periódicos nos quais a dita "internacionalização" está se dando. Entretanto, ainda assim, cabe considerar que este estudo utilizou as grandes áreas do conhecimento, sendo essencial estudar esses comportamentos nas áreas específicas.

Considerações Finais

A utilização de fontes de informação exaustivas da produção científica nacional permite a compreensão de características que podem guiar estudos bibliométricos, propiciar monitoramento do fluxo da comunicação científica e guiar políticas para a definição de critérios de avaliação. Este trabalho apresenta, segundo nosso entendimento, informações inéditas sobre a dispersão da produção científica de mais de 260.663 pesquisadores doutores associados à ciência brasileira, revelando tendências de internacionalização, assim como a importância de periódicos nacionais como veículos de publicação de parte significativa da produção científica de algumas áreas.

Fica evidenciada a limitação de estudos que não consideram a produção científica não indexada nas bases de dados examinadas, já que correspondem a 60% dos periódicos utilizados e 41% dos artigos. Mesmo estudos de internacionalização da produção brasileira devem atentar para esses resultados, pois somente 47% dos periódicos não indexados são editados no Brasil. Isso resulta que parte significativa dos periódicos não indexados seja composta de periódicos estrangeiros – 8% da região latino-americana e caribenha e 45% de fora dela –, ainda que, dos artigos publicados por periódicos não indexados, os do Brasil correspondam a 86%.

Ao se analisar a produção sob a ótica das Zonas de Bradford nas diferentes áreas, observou-se que, enquanto as humanidades (HUM, LIN e SOC) utilizam percentual significativo de periódicos não indexados, as demais áreas o fazem em menor percentual, mas proporcionalmente à zona: quanto maior a zona, ou quanto menor a frequência de artigos por periódicos, maior o percentual. A indexação na SciELO revelou-se importante, principalmente para AGR, seguida de SAU, HUM e BIO.

Em relação ao volume de artigos, observou-se que os periódicos não indexados vêm se destacando na zona de maior dispersão (Zona 3) das humanidades (HUM, LIN e SOC), enquanto a publicação em periódicos SciELO é destaque na Zona 1, o que mostra a importância marginal dos primeiros e a consolidação dos segundos. De maneira mais generalizada, entre as zonas de AGR e SAU e na Zona 1 de BIO e ENG, periódicos SciELO demonstraram sua relevância. E a proeminência de artigos publicados em periódicos indexados na Scopus e/ou WoS se deu na Zona 3 de BIO, ENG, EXA e SAU.

Desses resultados, pode-se concluir que, além dos periódicos não indexados, ignorar os periódicos SciELO pode representar limitação significativa em exercícios avaliativos de praticamente todas as áreas – e mais ainda em AGR e SAU. Por outro lado, é essencial analisar se o efeito da publicação em periódicos estrangeiros é de interesse para áreas nas quais não se observa sua consolidação (como Sociais e Humanas).

Acreditamos que este estudo apresenta um panorama de interesse tanto de pesquisadores das diversas áreas como de editores e componentes de comissões de avaliação de instâncias como CAPES e CNPq. No Brasil,

ainda não foi realizado um estudo como este, e pode-se destacar que, pelo fato de outros países não contarem com uma fonte de informação como a Plataforma Lattes, não há estudos dessa magnitude entre a comunidade internacional de estudos bibliométricos.

Finalmente, diversos são os desdobramentos de pesquisa que podem ser considerados como elementos para futuros trabalhos. Em particular, pretendemos proceder com uma análise mais profunda em diferentes áreas específicas, por as grandes áreas serem demasiadamente abrangentes, escondendo contrastes significativos (padrões e/ou anomalias, por exemplo).

Colaboradores

R. MUGNAINI contribuiu com a conceituação, metodologia, escrita e supervisão – rascunho original, revisão e edição. R.J.P. DAMACENO contribuiu com a curadoria de dados, visualização, escrita – rascunho original, revisão e edição. L.A. DIGIAMPIETRI contribuiu com a conceituação, curadoria de dados, metodologia e validação. J.P. MENA-CHALCO contribuiu com a curadoria de dados, metodologia, validação, visualização e escrita – revisão e edição.

Referências

Aksnes, D.W.; Sivertsen, G.A. Criteria-based assessment of the coverage of Scopus and Web of Science. *Journal of Data and Information Science*, v.4, n.1, p.1-21, 2019. Doi: <http://dx.doi.org/10.2478/jdis-2019-0001>

Albagli, S.; Clinio, A.; Raychtock, S. Ciência Aberta: correntes interpretativas e tipos de ação. *Liinc em Revista*, v.10, n.2, 2014. Doi: <http://dx.doi.org/10.18617/liinc.v10i2.749>

Araújo, R. F.; Alvarenga, L. A bibliometria na pesquisa científica da pós-graduação brasileira de 1987 a 2007. *Encontros Bibli*, v.16, n.31, p.51-70, 2011. Doi: <http://dx.doi.org/10.5007/1518-2924.2011v16n31p51>

Beigel, F. Publishing from the periphery: Structural heterogeneity and segmented circuits: The evaluation of scientific publications for tenure in Argentina's CONICET. *Current Sociology*, v.62, n.5, p.743-765, 2014. Doi: <http://dx.doi.org/10.1177/0011392114533977>

Björk, B.-C. *et al.* Open access to the scientific journal literature: Situation 2009. *Plos One*, v.5, n.6, e11273, 2010. Doi: <http://dx.doi.org/10.1371/journal.pone.0011273>

Bradford, S.C. *Documentação*. Rio de Janeiro: Fundo de Cultura, 1961. p.196-216.

Butler, L. Assessing university research: a plea for a balanced approach. *Science and Public Policy*, v.34, n.8, p.565-574, 2007. Doi: <http://dx.doi.org/10.3152/030234207X254404>

Caregnato, S.E. Google acadêmico como ferramenta para os estudos de citações: avaliação da precisão das buscas por autor. *Ponto de Acesso*, v.5, n.3, p.72-86, 2011. Doi: <http://dx.doi.org/10.9771/1981-6766rpa.v5i3.5682>

Cendón, B.V. Serviços de indexação e resumo. In: Cendón, B.V.; Campello, B.S.; Kremer, J.M. (Org.). *Fontes de informação para pesquisadores e profissionais*. Belo Horizonte: Editora UFMG, 2000. p.217-248.

Costas, R. Discussões gerais sobre as características mais relevantes de infraestruturas de pesquisa para a cientometria.

In: Mugnaini, R.; Fujino, A.; Kobashi, N.Y. (Ed.). *Bibliometria e cientometria no Brasil: infraestrutura para avaliação da pesquisa científica na era do Big Data*. São Paulo: Universidade de São Paulo. Escola de Comunicações e Artes, 2017. p.19-42. Disponível em: <http://www.livrosabertos.sibi.usp.br/portalde livrosUSP/catalog/book/129>. Acesso em: 25 jun. 2019.

Damaceno, R.J.P., *et al.* The Brazilian academic genealogy: Evidence of advisor-advisee relationships through quantitative analysis. *Scientometrics*, v. 119, n.1, p.303-333, 2019. Doi: <http://dx.doi.org/10.1007/s11192-019-03023-0>

Davyt, A.; Velho, L. A avaliação da ciência e a revisão por pares: passado e presente. Como será o futuro?. *História, Ciências, Saúde Manguinhos*, v.7, n.1, p.93-116, 2000. Doi: <http://dx.doi.org/10.1590/S0104-59702000000200005>

Fukahori, M.A.V. *As práticas de publicação científica das áreas de Ciência da Informação, Comunicação, Letras e Artes: as preferências dos bolsistas de produtividade em pesquisa PQ1 e PQ-SR do CNPq*. 2017. Dissertação (Mestrado em Ciência da Informação), Universidade Federal de Pernambuco, Recife, 2007.

Gimenéz-Toledo, E. La evaluación de las Humanidades y de las Ciencias Sociales en revisión. *Revista Española de Documentación Científica*, v.41, n.3, p.208, 2018. Doi: <http://dx.doi.org/10.3989/redc.2018.3.1552>

Gorraiz, J.; Purnell, P.J.; Glänzel, W. Opportunities for and limitations of the Book Citation Index. *Journal of the American Society for Information Science and Technology*, v.64, n.7, p.1388-1398, 2013. Doi: <http://dx.doi.org/10.1002/asi.22875>

Guédon, J.C. Acesso Aberto e divisão entre ciência predominante e ciência periférica. In: Ferreira, S.M.S.P.; Targino, M.G. (Org.). *Acessibilidade e visibilidade de revistas científicas eletrônicas*. São Paulo: Senac, 2010. p.21-78.

Hicks, D. The four literatures of social science. In: Moed, H.F.; Glänzel, W.; Schmoch, U. (Ed.). *Handbook of quantitative*

science and technology research. New York: Kluwer Academic, 2004. p.473-496.

International Standard Serial Number. *International Centre*. [S.l.]: ISSN, 2019. Available from: <https://portal.issn.org>. Cited: 02 Feb. 2019.

Lima, N.T. Avaliação desmedida. *Cadernos de Saúde Pública*, v.29, n.9, p.1723-1725, 2013. Doi: <http://dx.doi.org/10.1590/0102-311XCO080913>

Mckiernan, E.C. *et al.* Use of the Journal Impact Factor in academic review, promotion, and tenure evaluations. *PeerJ Preprints*, 2019. Available from: <https://peerj.com/preprints/27638.pdf>. Cited: 6 May 2019. Doi: <http://dx.doi.org/10.7287/peerj.preprints.27638v2>

Mena-Chalco, J.P. *et al.* Brazilian bibliometric coauthorship networks. *Journal of the Association for Information Science and Technology*, v.65, n.7, p.1424-1445, 2014. Doi: <http://dx.doi.org/10.1002/asi.23010>

Menandro, P.R.M. *et al.* Livros à mão cheia: o livro como veículo de produção acadêmica. *Psicologia USP*, v.22, n.2, p.367-386, 2011. Doi: <http://dx.doi.org/10.1590/S0103-6564201100020005>

Mugnaini, R. Ciclo avaliativo de periódicos no Brasil: caminho virtuoso ou colcha de retalhos? *In: Encontro Nacional de Pesquisa em Ciência da Informação*, 16, 2015, João Pessoa. *Anais [...]*. João Pessoa: UFPB, 2015.

Mugnaini, R.; Digiampetri, L.A.; Mena-Chalco, J.P. Comunicação científica no Brasil (1998-2012): indexação, crescimento, fluxo e dispersão. *Transinformação*, v.26, n.3, p.239-252, 2014. Doi: <http://dx.doi.org/10.1590/0103-3786201400030002>

Mueller, S.P.M. A publicação da ciência: áreas científicas e seus canais preferenciais. *DataGramaZero: Revista de Ciência da Informação*, v.6, n.1, 2005. Disponível em: <http://repositorio.unb.br/handle/10482/980>. Acesso em: 27 jun. 2019.

Neuhaus, C.; Daniel, H.D. Data sources for performing citation analysis: An overview. *Journal of documentation*, v.64, n.2, p.193-210, 2008. Doi: <http://dx.doi.org/10.1108/00220410810858010>

Packer, A.L.; Tardelli, A.O.; Castro, R.C.F. A distribuição do conhecimento científico público em informação, comunicação e informática em saúde indexado nas bases de dados MEDLINE e LILACS. *Ciência & Saúde Coletiva*, v.12, p.587-599, 2007. Doi: <http://dx.doi.org/10.1590/S1413-81232007000300009>

Rafols, I. Rumo a indicadores para 'abertura' de políticas de ciência e tecnologia. *In: Mugnaini, R.; Fujino, A.; Kobashi, N.Y. (Org.). Bibliometria e cientometria no Brasil: infraestrutura para avaliação da pesquisa científica na Era do Big Data*. São Paulo: ECA/USP, 2017. p.67-78. Disponível em: <https://doi.org/10.11606/9788572051705>. Acesso em: 25 jun. 2019.

Russell, J.M. Scientific communication at the beginning of the twenty-first century. *International Social Science Journal*, v.53, n.168, p.271-282, 2001.

Shotton D. Funders should mandate open citations. *Nature*, v.553, n.7687, p.129, 2018. Doi: <http://dx.doi.org/10.1038/d41586-018-00104-7>

Sivertsen, G. Understanding and evaluating research and scholarly publishing in the Social Sciences and Humanities (SSH). *Data and Information Management*, v.2, n.3, p.1-11, 2019. Doi: <http://dx.doi.org/10.2478/dim-2019-0008>

Vanclay, J.K. Impact factor: Outdated artefact or stepping-stone to journal certification? *Scientometrics*, v.92, n.2, p.211-238, 2012. Doi: <http://dx.doi.org/10.1007/s11192-011-0561-0>