

Editor

César Antonio Pereira

Conflicto de interesse

Os autores declaram não haver conflitos de interesse.

Recebido

29 dez. 2022

Aprovado

28 abr. 2023

Análise estatística da literatura publicada sobre Organização do Conhecimento no Brasil

Statistical analysis of published literature on Knowledge Organization in Brazil

Rubén Urbizagástegui-Alvarado¹ , Michely Jabala Mamede Vogel² 

¹ University of California, University Libraries, Metada Department. University Avenue, s/n., 92521, Riverside, California, USA.

² Universidade Federal Fluminense, Instituto de Arte e Comunicação Social, Departamento de Ciência da Informação e Programa de Pós-Graduação em Ciência da Informação. Niterói, RJ, Brasil. Correspondência para/Correspondence to: Michely J. M. Vogel. E-mail: michelyvogel@id.uff.br

Como citar este artigo/How to cite this article: Análise estatística da literatura publicada sobre Organização do Conhecimento no Brasil. *Transinformação*, v. 35, e227245, 2023. <https://doi.org/10.1590/2318-0889202335e227245>

Resumo

Para compreender um campo, é preciso conhecer como cresce a literatura por ele produzida. Objetiva-se realizar análise estatística inferencial do crescimento da literatura publicada por pesquisadores brasileiros sobre Organização e Representação do Conhecimento de 1972 até 2021. Apresenta-se uma revisão sobre os estudos de crescimento da literatura, e exploram-se os modelos propostos para medir esse crescimento. Foi realizada a coleta de artigos publicados em periódicos acadêmicos, capítulos de livros e trabalhos apresentados em congressos que tratassem alguns aspectos da ORC em lista abrangente de bases de dados e sítios da web, sendo criada uma base de dados em EndNote X8. A literatura coletada foi analisada à luz do Modelo de Crescimento Exponencial e do Modelo Epidêmico. O estudo resultou em 3.630 artigos, produzidos por 2.535 pesquisadores diferentes, sendo a maior parte em trabalhos de evento (53,3%), seguido por artigos em periódicos (35,2%) e por capítulos de livros (5%). O português foi o idioma de 89,5% dos textos, encontrando-se publicações em inglês, espanhol e francês. Quanto aos autores, 20 são os mais produtivos e representam 5% do total de autores pesquisando e publicando no campo. Conclui-se que é partir do ano 2000 que a publicação começa a crescer de forma exponencial, sendo que 10,3% dela foi publicada nos primeiros 30 anos e 67,2% dela foi publicada nos últimos 10 anos. Essa literatura está crescendo de forma exponencial com uma taxa de 12,2% ao ano, duplicando seu volume a cada seis anos, e não alcançou ainda seu ponto de saturação.

Palavras-chave: Análise estatística da literatura. Comunicação científica. Literatura científica brasileira. Organização e Representação do Conhecimento.

Abstract

The objective of this paper is to carry out an inferential statistical analysis of the growth of literature published by Brazilian researchers on Knowledge Organization and Representation (KOR) from 1972 to 2021. It presents a review of studies on growth of literature and explores

proposed models to measure this growth. Articles published in academic journals, book chapters, and papers presented at conferences that dealt with some aspects of Knowledge Organization and Representation were collected in a comprehensive list of databases and websites, and a database was created in EndNote X8. The collected literature was analyzed considering the exponential growth model and the epidemic model. It was found 3,630 articles produced by 2,535 different researchers, most of which were congresses papers (53.3%), followed by journal papers (35.2%), and book chapters (5%). Portuguese was the language of 89.5% of the texts, with publications in English, Spanish, and French. Considering only authors, 20 are the most productive and account for 5% of the total number of authors researching and publishing in the field. It is since the year 2000 that the publication begins to grow exponentially, 10.3% of this literature was published in the first 30 years and 67.2% was published in the last 10 years. This literature is growing exponentially at a rate of 12.2% per year, doubling its volume every six years, and has not reached yet its saturation point.

Keywords: *Scientific communication. Statistical analysis of the literature. Brazilian scientific literature. Knowledge Organization and Representation.*

Introdução

As pesquisas acadêmicas sobre o desenvolvimento dos diversos campos do conhecimento buscam refletir o avanço da ciência nos contextos em que esses se desenvolvem, assim como fornecer *inputs* para os processos de avaliação das organizações responsáveis pelo planejamento educativo. Orientam-se a traçar um estudo metódico, reflexivo e sistemático de sua formação e desenvolvimento sedimentados em seus produtos intelectuais: a literatura acadêmica publicada específica do campo. Neste trabalho, aborda-se o estudo da literatura publicada sobre a Organização e Representação do Conhecimento (ORC) no Brasil como um espaço de pesquisas e reflexões produzidas por um coletivo de pesquisadores produtores da literatura acadêmica próprio dos saberes e práticas desse campo. O objetivo geral da ORC é possibilitar o acesso ao conhecimento contido nas informações. No contexto da Ciência da Informação (CI), são esquemas de representação do conhecimento e/ou instrumentos que fazem a tradução dos conteúdos dos documentos a um esquema estruturado que representa esses conteúdos. Isso se faz com a finalidade de organizar a informação e facilitar a recuperação dos assuntos contidos nos documentos. É devido a essas interações que a ORC é definida como:

[...] a ciência que ordena a estruturação e sistematização dos conceitos, de acordo com suas características, que podem ser definidas como elementos de herança do objeto, e a aplicação dos conceitos e classes dos conceitos ordenados pela indicação de valores, dos referentes dos objetos ou assuntos. A partir dessa organização do conhecimento criam-se ferramentas que apresentam a interpretação organizada e estruturada do objeto, chamado de SOC (Carlan; Medeiros, 2011, p. 55).

É possível constatar que a área da ORC está:

[...] em pleno vigor dado suas diversidades conceituais teóricas e metodológicas em discussão. Portanto, se por um lado, falta à área de Organização do Conhecimento uma consolidação científica, de outro, é inegável sua diversidade conceitual e o impacto de seus resultados para a organização do conhecimento de outras áreas científicas. Nesse contexto, é absolutamente fundamental à comunidade científica conhecer seu domínio (Fujita, 2008, p. 5).

Conhecer o campo é também conhecer como cresce a literatura produzida nesse campo, e conhecer esse crescimento é de vital importância para entender o desenvolvimento de um campo científico em particular. Um campo científico designa um espaço de produção cultural que ocupa uma posição em uma estrutura social determinada e se distingue de outros espaços culturais por seus interesses, estruturas, regras e formas de construir e impor uma hierarquia e prestígio. É um

campo sujeito à concorrência entre os que interatuam nesse espaço, por tanto, um campo está sempre em permanente construção e reconfiguração. Por outro lado, não há campo científico sem literatura publicada, já que está explícito o estado atingido pelos campos científicos, mesmo porque os concorrentes desse campo se alimentam dessa literatura para a reprodução do próprio campo. A ORC forma parte da CI, e, para existir e ser considerada como “ciência”, deve tornar público os resultados de suas indagações; portanto, este artigo insere-se na interlocução desses dois campos de conhecimento e práticas acadêmicas: CI e ORC no Brasil. É nesse espaço que estão inseridos os agentes e instituições que produzem, reproduzem ou difundem a literatura publicada sobre a ORC no Brasil.

O propósito desta pesquisa é fazer uma análise estatística inferencial do crescimento da literatura publicada por pesquisadores brasileiros sobre ORC desde seu início, marcado pela data da primeira publicação identificada em 1972 até dezembro de 2021, um período longo o suficiente para observar se a literatura acadêmica ligada ao campo da ORC cresceu. Questiona-se a forma desse crescimento, buscando gerar o indicador da taxa de crescimento, assim como o indicador da taxa de duplicação dessa literatura. Até onde é do conhecimento dos autores desta pesquisa, a literatura publicada por brasileiros sobre ORC não tem sido pesquisada para produzir indicadores orientados a construir essa taxa de crescimento ou sequer a sua taxa de duplicação. Procura-se, então, encontrar e construir esses dois indicadores. Será explorado também se essa literatura produzida entrou em uma etapa epidêmica e quando essa fase se inicia. Pesquisar a produção científica nas diversas áreas do conhecimento e fazer visível a ciência analisada, a partir de indicadores estatísticos, procura uma explicação histórico-social nos pontos sinalizados pelos dados estatísticos e analisa o contexto em que nasceram. Os números por si só têm pouco valor histórico, mas adquirem sentido na medida em que produzem significados a partir da análise do contexto histórico-social em que foram produzidos.

Marco Teórico

Propõe-se a bibliometria como uma ferramenta capaz de medir e facilitar a análise da informação armazenada e que:

[...] embora seu uso acarrete algumas dificuldades, muita informação útil [...] como a transmissão de ideias, crescimento e tendências de uma disciplina científica pode ser obtida por meio da análise estatística [...]. Por tanto, a bibliometria pode ser usada como um filtro de informação, podendo ser de grande utilidade para o pesquisador no levantamento do estado da arte do seu tema de pesquisa (Mugnaini, 2003, p. 48).

Essa mesma afirmativa é verdadeira para a estatística, especialmente a estatística inferencial, que será utilizada nesta pesquisa, pois a análise estatística probabilística e inferencial pode ser aplicada como um filtro da informação coletada e ser de grande utilidade na análise do estado da arte de um determinado assunto de pesquisa.

Não se pretende fazer um estudo quali-quantitativo (Miranda; Souza, 2021; Marques; Gonçalves, 2021, Santa Anna; Dias; Maculan, 2019) da literatura publicada sobre ORC, tampouco fazer uma pesquisa descritiva de abordagem quali-quantitativa (Oliveira, 2021) nem uma análise quali-quantitativa (Vieira; Oliveira; Cabral, 2020), sequer uma “análise bibliométrica” (Tartarotti; Fujita, 2016; Santos, 2015), nem ainda um “estudo bibliométrico” (Sales; Gaudêncio; Neves, 2018), nem uma “análise métrica” da produção de documentos sobre a ORC (Castro; Oliveira, 2016), mesmo porque esses tipos de estudos são primários e fazem uso da estatística descritiva

unidimensional e, portanto, não deveriam ser considerados, rigorosamente, como estudos bibliométricos, informétricos ou cientométricos. Para serem considerados estudos deste último tipo, deve-se trabalhar ao menos com duas variáveis correlacionadas, como no caso da lei de Lotka (autores *versus* produção de artigos), da lei de Bradford (revistas *versus* produção de artigos) e da lei de Zip (palavras *versus* a frequência de ocorrência dessas palavras em um texto).

Neste trabalho, pretende-se fazer uma análise estatística bidimensional inferencial relacionada com o crescimento dessa literatura e produzir os indicadores pertinentes. Quando se analisa o crescimento da literatura, postula-se uma correlação entre o tempo, medido em anos (variável independente), e o volume acumulado da literatura produzida segundo os anos, medida em unidades produzidas (variável dependente), seguindo um modelo explicitado na forma de uma equação estatística. Essa equação estatística é garantia para que uma pesquisa seja considerada pesquisa bibliométrica, infométrica ou cientométrica, que são apenas manifestações da estatística inferencial ou probabilística.

Neste contexto, já existem algumas aproximações do tipo bibliométrico; por exemplo, Kobashi e Santos (2007) analisaram as teses brasileiras produzidas em duas áreas científicas: CI, no período de 1997-2001, e Energia Nuclear, nos anos 1976-2005. Na CI, estudaram a representação temática das dissertações, encontrando uma grande variação na atribuição dos descritores e carecendo de normalização nos termos de indexação utilizados. Para o caso da Energia Nuclear, traçaram um mapa baseado nos assuntos abordados pelas teses e dissertações, encontrando que essas estavam dispersas em 33 categorias de assuntos. Segundo os autores:

O estudo mostrou que os mapas científicos, oriundos da exploração bibliográfica de bases de dados, requerem organização prévia e consistente das informações, com base em políticas explícitas e instrumentos terminológicos adequados à representação da informação (tesauros e ontologias). O estudo também mostrou que evidenciar as relações multidimensionais entre os elementos dos registros bibliográficos é essencial para descobrir as relações sociais e cognitivas que se estabelecem dentro de um campo científico e representá-las por meio da visualização gráfica (Kobashi; Santos, 2007, p. 293).

Com o objetivo de identificar as dificuldades encontradas na descrição física e o conteúdo das fontes utilizadas quando se realizam estudos bibliométricos, Café e Bräscher (2008), utilizando a base de dados LICl e consultando também a Biblioteca Digital de Teses e Dissertações (BDTD), ambas do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), efetuaram um levantamento de 59 teses e dissertações e realizaram entrevistas com especialistas do campo da bibliometria. Afirmam que as dificuldades se concentram, principalmente, na padronização dos títulos dos periódicos e dos nomes dos autores, assim como na diversidade de representações dos assuntos; por tanto, “[...] padrões de organização da informação são necessários em pesquisas bibliométricas para que documentos que possuam características similares possam ser recolhidos em fontes primárias e secundárias e para que se possa atingir confiabilidade na análise dos dados” (Café; Bräscher, 2008, p. 54). Procurando apresentar um panorama da pesquisa brasileira no campo da ORC, Alvarenga e Silva (2010) realizaram análise estatística descritiva elemental do conteúdo, assuntos e citações das publicações apresentadas no Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB) de 2008 e 2009 pelo Grupo de Trabalho 2 da Associação Nacional de Pesquisa em Ciência da Informação (ANCIB).

A revista *Scire: Representación y Organización del Conocimiento* (Espanha) é uma publicação semestral de caráter interdisciplinar orientada à divulgação de pesquisas sobre representação,

normalização, tratamento e recuperação e comunicação da informação e do conhecimento. Tomando como bases as publicações desse periódico, Alves, Grácio e Oliveira (2011) analisaram:

[...] as instituições e países mais produtivos, as temáticas que mais têm contribuído para os debates e desenvolvimento do saber no tema, bem como construir e analisar a rede de colaboração institucional de todas as instituições que fizeram colaboração e calcular os indicadores de densidade e centralidade de grau, no período de 2006 a 2010 (Alves; Grácio; Oliveira, 2011, p. 34).

Encontrou-se um total de 48 instituições, sendo 17 as mais produtivas, com pelo menos 2 artigos publicados. As temáticas mais presentes na revista, em ordem decrescente de frequência, foram: recuperação da informação; cibermetria; bibliotecas públicas; indexação; catalogação temática; linguagem documental; metadados; classificação; bibliotecas universitárias; gestão da informação. A rede de colaboração institucional mostrou destaque de quatro sub-redes espanholas, entre as quais ocorrem parcerias com a Universidade Federal de Mato Grosso. Observou-se, ainda, grande frequência de coautorias intrainstitucionais na Universidade Estadual Paulista (Unesp), na Universidade de Salamanca e na Universidade de Extremadura.

Coletando informações de três bases de dados (EBSCO, Scopus e *Web of Science*), Zancanaro *et al.* (2013) mapearam o perfil e as características das publicações internacionais relacionadas à memória organizacional e às ontologias. Identificaram 39 trabalhos escritos por 85 autores, vinculados a 40 instituições de 18 países diferentes. Nesses trabalhos, os autores utilizaram 845 referências e 117 palavras-chave, com destaque para *Knowledge management*, *e-Learning* e *Organizational learning*. Portanto,

[...] a memória organizacional pode ser entendida como uma área que faz a intersecção com outras áreas de pesquisa, tais como gestão do conhecimento, aprendizagem organizacional e engenharia do conhecimento. Dentre essas intersecções, estão os trabalhos que estudam as técnicas da inteligência artificial, como, por exemplo, as ontologias que têm sido desenvolvidas e aplicadas para a compreensão, o uso e o reuso do conhecimento para a tomada de decisão nas organizações (Zancanaro *et al.*, 2013, p. 59).

Já Castro e Oliveira (2016) analisaram a produção científica sobre as linguagens de indexação, divulgada nos Anais dos Encontros Nacionais de Pesquisa em Ciência da Informação e Biblioteconomia no período de 2012 a 2015. Encontraram 73 artigos produzidos por 102 autores diferentes e identificaram as instituições e os autores que mais publicaram nos anais do ENANCIB. Mostraram também que as principais linguagens de indexação pesquisadas foram as listas de cabeçalhos de assuntos, os tesouros e os sistemas de classificação bibliográfica; porém, os assuntos mais abordados foram indexação, ORC, tesouros, organização da informação, recuperação da informação e metadados. Amorim Neto e Lima (2016), coletando informações da Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (BRAPCI) no período de 1972 a 2012, estudaram o assunto Organização do Conhecimento (OC) na CI, descobrindo que esse assunto se encontra em crescimento, já que “53,2 % da produção brasileira sobre o domínio de OC na CI foi publicada entre 2006 e 2012 e que os demais 46,8 % dos artigos foram publicados nos 34 anos anteriores” (Amorim Neto; Lima, 2016, p. 101).

Figueiredo e Almeida (2017), procurando compreender o contexto da produção científica sobre ontologias no Brasil, analisaram os artigos científicos indexados pela BRAPCI. A busca nessa base de dados retornou 278 trabalhos, excluindo 105 artigos e analisando um total de 173 artigos científicos; nesse volume de literatura, notaram um crescimento no número de publicações a partir

do ano de 2007, com pico do número de artigos publicados em 2009 (21 artigos). Contudo, não indicam o período da pesquisa e nem a taxa desse crescimento.

Como se pode concluir pela literatura revisada anteriormente, não há em nenhum desses estudos a intenção de analisar o crescimento da literatura sobre a temática ORC publicada no país desde seus inícios e nem produzir os indicadores correspondentes, apesar de ser evidente que existe interesse por explorar esse campo. Uma segunda observação se refere ao período da pesquisa: em alguns casos, os estudos não indicam o período de estudo e, em outros, os períodos são limitados a poucos anos, por tanto não são representativos das áreas estudadas. Um estudo deste tipo deve cobrir a extensão completa do fenômeno estudado. Nesse sentido, esta pesquisa pretende assinalar pautas para futuras pesquisas deste tipo, procurando responder as seguintes perguntas básicas que orientam esta pesquisa: a literatura publicada sobre ORC no Brasil está em crescimento ou já alcançou seu ponto de saturação? E se está em crescimento, qual é a forma desse crescimento? Qual é sua taxa de crescimento atual e qual é sua taxa de duplicação? Essa literatura já entrou a uma etapa epidêmica? E se assim fosse, quando entrou nessa etapa epidêmica e qual é sua tendencia futura?

Para alcançar o objetivo proposto, este artigo está organizado em cinco partes. Na primeira, apresenta-se uma introdução ao assunto ORC, detalhando o problema e formulando as perguntas de pesquisa. Na segunda, oferece-se o referencial teórico sobre os estudos de crescimento da literatura, revisando exaustivamente a literatura publicada e explorando os modelos propostos para medir esse crescimento. Na terceira, descrevem-se a metodologia empregada e a aplicação do modelo de crescimento escolhido, com ênfase na forma de coleta dos dados e na forma de medição dos dados. Na quarta parte, expõem-se os resultados. Por fim, são oferecidas as conclusões e discussões dos resultados obtidos.

Procedimentos Metodológicos

Cada um dos artigos publicados em periódicos acadêmicos, capítulos de livros e trabalhos, apresentados em congressos que tratassem alguns aspectos da ORC e fossem publicados até dezembro de 2021, foram tomados como unidades de análise. Não foram considerados teses, monografias e literatura cinzenta, por ser comum seu versionamento em artigos de periódicos ou apresentações em congressos, e nem livros, pois as bases de dados pesquisadas não incluem referências bibliográficas a seu respeito.

Uma vez definido o assunto a ser pesquisado (ORC), definiu-se a estratégia de busca inicial. Dividiu-se o assunto em duas Palavras-Chave (PC): OC e representação do conhecimento. Isso foi feito como um ensaio, e por isso utilizamos a janela temporal de 2021 a 2022.

Em uma primeira aproximação, seguiu-se o seguinte roteiro:

1) Como primeira aproximação, usaram-se essas duas PC para recuperar a informação inicial na BRAPCI. Tomou-se apenas uma mostra de 200 registros iniciais.

2) Buscando “organização da informação” no título, PC e resumo da BRAPCI, desde 2021 até 2022, foram recuperados 152 registros.

3) Buscando “representação do conhecimento” no título, PC e resumo da BRAPCI, de 2021 até 2022, foram recuperados 69 registros.

4) Todos esses 221 registros foram enviados ao EndNote X8 para construir uma base de dados bibliográfica (BDB) específica inicial sobre ORC no Brasil.

5) Foram eliminados 21 registros duplicados.

6) Nos 200 registros iniciais, foi analisada a frequência de aparecimento das PC, e notou-se que havia algumas variações nas PC empregadas. Por exemplo: organização da informação; representação da informação; sistemas de OC; sistemas de organização da informação.

7) Já conhecidas as PC mais utilizadas no campo da ORC, decidiu-se que as PC que seriam usadas na busca final na BRAPCI, assim como em outras bases de dados, seriam seis: OC; representação do conhecimento; organização da informação; representação da informação; sistemas de OC; e sistemas de organização da informação.

Com relação à estratégia de busca propriamente dita, a janela temporal foi maior (1972-dezembro de 2021). Considerando os ajustes após o ensaio, as ações foram:

1) Definidas as PC de busca, elas foram novamente utilizadas na busca definitiva no BRAPCI, agora desde 1972 até dezembro de 2021. Os resultados foram:

- a) OC = 1.181 registros
- b) Representação do conhecimento = 550 registros
- c) Organização da informação = 2.050 registros
- d) Representação da informação = 1.083 registros
- e) Sistemas de OC = 199 registros
- f) Sistemas de organização da informação = 342 registros.

2) Todos os registros identificados foram baixados e enviado ao EndNote X8 para formar uma espécie de BDB especializada em ORC no Brasil.

3) Considerou-se como produção brasileira a todos os autores desses documentos afiliados a uma instituição localizada em algum espaço geográfico do Brasil. Essa afiliação também foi registrada na BDB no EndNote X8.

Para a organização final dos registros bibliográficos na BDB, normalizaram-se os nomes dos autores, os títulos dos periódicos, os nomes dos congressos e os eventos similares. Em seguida, foram eliminados os registros duplicados. Vale lembrar que o EndNote X8 oferece estratégias para tal padronização e remoção de duplicatas. Com as modificações pertinentes à mesma estratégia, padronização foi aplicada nas buscas nas bases de dados bibliográficas e sítios web a seguir listados, selecionados por oferecem conteúdos de ORC de autores filiados à instituições brasileiras: BRAPCI (<https://www.brapci.inf.br/>); FEBAB (<http://repositorio.febab.org.br/>); ANCIB (<https://ancib.org/>); Encontros EDICIC (<https://edicic.org/encuentros-y-eventos/encuentros-edicic/>); CIFORM SALVADOR BAHIA (<http://www.ciform-antiores.ufba.br/>); ISKO-BRASIL (<https://isko.org.br/eventos/>); ISKO *España-Portugal* (<http://www.iskoiberico.org/>); ISKO-Espanha (<http://www.iskoiberico.org/congresos/getafe-1995/>); *International Society for Knowledge Organization* (<https://www.isko.org/>); Repositório da Universidade de Lisboa (<https://repositorio.ul.pt/>); Dialnet – *España* (<https://dialnet.unirioja.es/>); Google Acadêmico (<https://scholar.google.com/>); *WORLDCAT* (<https://www.worldcat.org/>); Catálogo da Biblioteca do Congresso dos Estados Unidos (<https://catalog.loc.gov/>); *University of Washington Libraries* (<https://www.lib.washington.edu/>); *Pascal and Francis Bibliographic Databases* (<https://pascal-francis.inist.fr/home/>); *Hispanic American Periodicals Index – HAPI* (<https://hapi.ucla.edu/>), Redalyc (<https://www.redalyc.org/>); Scopus (<https://www.scopus.com/>); *Web of Science* (<https://www.webofscience.com/>); e *Academic Search Complete* (<https://www.ebsco.com/products/research-databases/academic-search-complete>). A seguir, estão os resultados por base de dados:

- BRAPCI = 5.405 registros

- FEBAB = 680
- ANCIB = 2.180
- Encontros EDICIC = 540
- CIFORM SALVADOR BAHIA = 340
- ISKO-BRASIL = 520
- ISKO *España-Portugal* = 243
- ISKO-*Espanha* = 148
- *International Society for Knowledge Organization* = 168
- Repositório da Universidade de Lisboa = 250
- Dialnet - *España* = 450
- Google Acadêmico = 3.300
- *WORLDCAT* = 250
- Catálogo da Biblioteca do Congresso dos Estados Unidos = 158
- *University of Washington Libraries* = 1.810
- *Pascal and Francis Bibliographic Databases* = 250
- *Hispanic American Periodicals Index* - HAPI = 125
- Redalyc = 630

Destaca-se que foi realizada também uma revisão minuciosa dos documentos identificados na busca, prestando atenção especial às citações feitas no documento revisado. Cada referência relacionada à ORC foi verificada na base de dados construída em Ednote X8, e incorporada caso não tivesse sido identificada nas buscas anteriores nas bases de dados bibliográficas. As referências duplicadas foram eliminadas, mantendo-se apenas uma referência não repetida. Finalmente, foi produzida uma bibliografia que lista um total de 3.630 referências contendo artigos de periódicos, capítulos de livros e comunicações apresentadas em congressos e eventos similares. Essa bibliografia de 3.630 referências, identificadas como publicadas e/ou produzidas entre 1972 e 2021, constitui o universo desta pesquisa. Ressalta-se que a busca foi possível, pois há uma estabilidade na nomenclatura dos termos utilizados como PC na área ORC ao longo do tempo. A literatura coletada cobre um longo período de 50 anos suficientes para esperar que seja mostrada alguma forma de crescimento. Porém, é impossível afirmar se o crescimento dessa literatura assumirá uma forma linear, exponencial, de poder, ou de Gompertz, porém espera-se que seja exponencial, uma vez que essa forma de crescimento é a mais comum encontrada em estudos de literaturas que abrangem longos períodos, como neste caso (Sharma; Gupta; Kumar, 2002; Hadagali; Gavisiddappa, 2015; Urbizagástegui-Alvarado, 2016; Urbizagástegui; Restrepo-Arango, 2017).

Modelo de Crescimento Exponencial

O crescimento exponencial representa o aumento da população analisada em uma proporção fixa em cada unidade de tempo, e é expresso em porcentagens, como uma taxa de crescimento constante e com limite de crescimento indefinido. O modelo não fornece apenas uma taxa média de crescimento, mas também uma taxa de duplicação, ou seja, uma taxa ou proporção em que o tamanho da população estudada se duplica. Geralmente, o crescimento exponencial assume uma forma côncava inicial na sua representação gráfica. Segundo Egghe e Ravichandra Rao (1992), essa função é representada matematicamente como:

$$C(t) = CO e^{at}$$

Mas pode ser reescrita como:

$$C(t) = c g^t$$

onde

$$c > 0, g > 1 \text{ e } t \geq 0$$

Ao estudar o crescimento da literatura, postula-se uma relação entre o tempo, medido em anos (variável independente), e o volume acumulado da literatura produzida, medida em unidades produzidas (variável dependente) em cada ano. Assume-se que essa relação bivariada pode ser modelável estatisticamente e produzir uma taxa de crescimento e uma taxa de duplicação da literatura produzida.

Modelo Epidêmico

Goffman (1966a) afirmava que um dos problemas mais intrigantes da ciência moderna era a natureza obscura de seu próprio crescimento e que poucos cientistas tinham pensado seriamente como sua ciência tinha chegado ao seu atual estado e qual seria seu desenvolvimento futuro. Assinalava também que as ideias que se dispersam em uma determinada população de científicos possuem propriedades epidêmicas, por tanto, essas ideias poderiam ser estudadas como se fossem uma epidemia (Goffman; Newill, 1964). Algumas pessoas são suscetíveis a certas ideias, e outras são resistentes, porém, uma vez que uma pessoa tenha sido infectada com uma ideia, essa mesma pessoa, depois de certo período de latência, pode transmitir essa ideia a outras pessoas, resultando em uma epidemia. Essas ideias são transmitidas através dos documentos publicados. Os detalhes concernentes à teoria do modelo epidêmico podem ser revisados nos artigos publicados por Goffman (1966a, 1966b, 1971), Goffman e Harmon (1971), e Goffman e Warren (1972). A população considerada é composta por: *infectados*: todos aqueles que possuem o material infectante. Nesse caso, considerou-se que um autor estava infectado na data em que publicou sua primeira publicação, devido à dificuldade de determinar o ano em que um autor começa ou deixa de pensar (escrever) sobre um assunto; *suscetíveis*: são todos aqueles que podem se infectar, desde que tenham tido contato com o material infectante, ou seja, as ideias ou documentos; e *removidos*: todos aqueles que não fazem mais parte da população: morreram ou mudaram de interesse. Nesse caso, um autor é considerado como removido um ano após a publicação de seu último documento.

Na sua forma mais simples, um processo epidêmico poder ser representado por um sistema de equações diferenciais:

$$\frac{dS}{dt} = \beta SI$$

$$\frac{dI}{dt} = \beta SI - \gamma I \quad (1)$$

$$\frac{dR}{dt} = \gamma I$$

Esse modelo assume que existe uma população suscetível de tamanho S , uma população infectada de tamanho I e uma população removida de tamanho R , na qual a remoção constitui um estado de desaparecimento que só pode ser alcançado passando por um estado de infecção. Supõe-se ainda que a infecção é transmitida por contato direto entre pessoas suscetíveis e infectadas com β igual à taxa de infecção e γ igual à taxa de remoção. A população total de $N = S + I + R$ permanece constante em relação ao tempo; não há mistura homogênea entre os membros de N , e o período de latência é igual ao período de incubação que é igual a zero; as infecções ocorrem continuamente ao longo do tempo, e um suscetível se torna infectado no momento que ocorre o contato efetivo.

Resultados

Até dezembro de 2021, foram publicados 3.630 documentos produzidos por 2.535 pesquisadores diferentes. Foram 1.278 artigos de periódico (1.169 em português, 91 em inglês, 16 em espanhol e dois em francês), 2.166 apresentações em congressos (1.900 em português, 192 em inglês, 62 em espanhol e 12 em francês) e 186 capítulos de livros (181 em português, quatro em inglês e um em espanhol). A forma preferencial dos pesquisadores brasileiros sobre esse assunto é apresentar os resultados e/ou avanços de suas pesquisas em congressos e eventos similares (52,3%) e logo publicá-los em periódicos nacionais (35,2%) ou como capítulos de livros (5,0%). Nada mais natural que os comunicar no idioma português na qual foram socializados. Esses três tipos de documentos publicados em português representam 89,5% do total das publicações sobre ORC.

É curioso constatar que as apresentações em congressos sejam as preferidas nos três idiomas identificados. Provavelmente, alguns pesquisadores associam a participação nesses tipos de eventos nacionais e internacionais com o turismo local ou regional. Essa é uma prática estendida em todas as disciplinas acadêmicas. O turismo profissional inclui congressos, colóquios, seminários, feiras, exposições, reuniões e eventos similares que estão se convertendo em fontes de ingressos monetários, tanto que muitos países e cidades investem na construção de auditórios e centros de conferências, de forma a dotá-los dos meios essenciais para entrar na concorrência por esse lucrativo mercado. Para o caso espanhol, esse assunto é explorado por Espejo Marín e Ponce Sánchez (1999), para o caso mexicano, por Ancona Alcocer (2020), e para o caso cubano, por Mora-Sánchez, Lao-León e León-Robaina (2021).

Modelo de crescimento exponencial

A Figura 1 mostra o crescimento da produção de documentos segundo os anos. Com pequenas flutuações, essa literatura não cresce nos primeiros 29 anos. A partir de 2000, começa a crescer de forma exponencial, côncava no início e linear em um segundo momento, ainda que com flutuações e uma queda notória no ano 2020 e um aumento notório no ano 2021 (Figura 1).

A flutuação no ano 1990 deve-se a que, de 10 a 14 de setembro desse ano, realizou-se o segundo Simpósio Latino-americano de Terminologia em Brasília, no qual 61 pesquisadores brasileiros interessados na organização da informação e do conhecimento apresentaram trabalhos sobre os mais diversos assuntos linguísticos, indo desde análise documentária a exploração dos aspectos ideológicos, lexicográficos, pragmáticos e semânticos dos artigos científicos e a organização de banco de dados linguísticos, construção de tesouros e controle terminológico. A segunda flutuação ocorre em 1995, quando, em Getafe, Espanha, realizou-se o Segundo Congresso ISKO-Espanha, do qual participam 86 pesquisadores brasileiros discutindo a OC e vocabulários controlados para arquivos, aplicação da teoria da análise facetada e mapas conceituais para

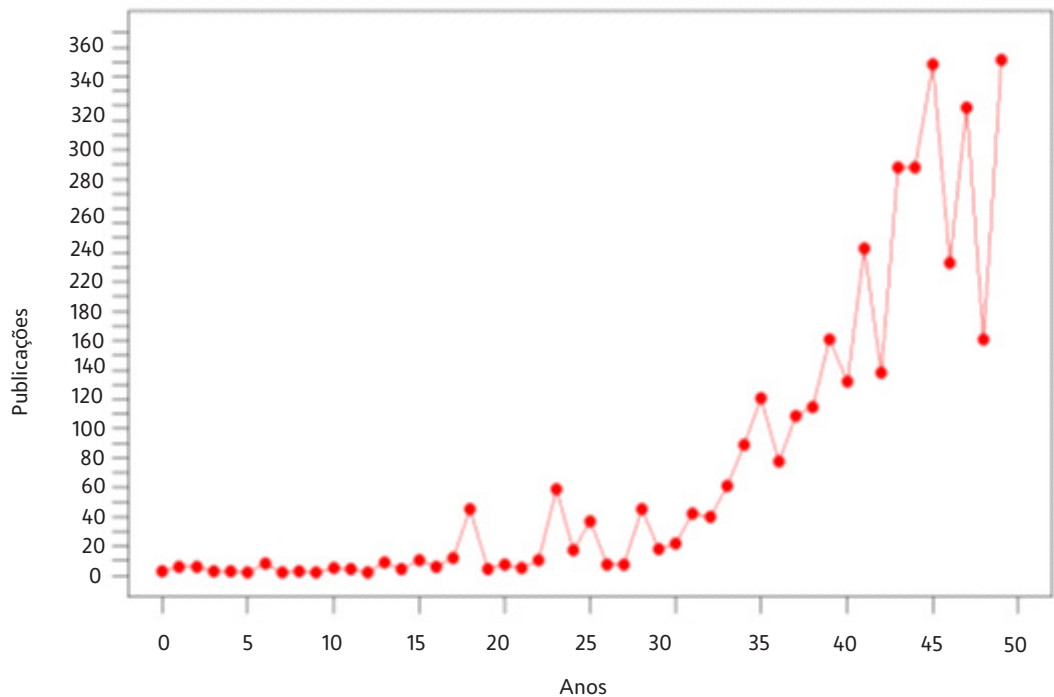


Figura 1 – Produção de documentos segundo os anos.

Fonte: Elaborada pelos autores (2022).

navegação em ontologia de domínios, políticas de indexação em bibliotecas universitárias, entre outros diversos assuntos. Também, nesse mesmo ano, acontece o segundo ENANCIB em Valinhos, São Paulo, onde se discutem os diversos aspectos das linguagens de indexação e linguagem documentária. A queda no volume de produção no ano 2020 explica-se pelos efeitos da pandemia de COVID-19 que paralisou a maioria dos eventos programados para esse ano e anos posteriores no mundo inteiro, retomando-se os encontros virtuais em 2021. Os dados foram logo acumulados, e mostram o traçado do crescimento da produção de documentos segundo os anos (Figura 2). Nota-se agora mais claramente que o crescimento dessa literatura começa no ano 2000, e que essa tem uma forma exponencial. Perguntamo-nos: por que essa literatura começa a crescer a partir desse ano e quais seriam os motivos ou detonadores desse crescimento?

Louzada (2002) menciona que, em 2000, criou-se um sistema nacional de fomento à pesquisa, com recursos dos fundos setoriais, estudando os efeitos desse novo sistema, encontrou que houve um aumento considerável de cursos de pós-graduação no Brasil. Observou também que o número de pessoas inseridas em cursos de pós-graduação em 2000 era calculado em 54 mil estudantes e que a cada ano se formavam cerca de 6.000 doutores no país. Afirma, ainda, que a evolução do contingente de alunos de pós-graduação parece ter papel relevante no aumento da produtividade científica/acadêmica brasileira, em função da coincidência no tempo destes dois fatores. Esse pode ser o caso dos cursos de pós-graduação em CI, onde se inserem os cursos sobre ORC no Brasil que teriam impulsionado essa produtividade. Por exemplo, Bufrem e Breda (2011) mencionam a existência de 72 cursos de CI distribuídos no país inteiro com 8 programas de pós-graduação (Costa; Castro, 2007); e Miranda e Souza (2021) mencionam 75 cursos de graduação em biblioteconomia e 40 cursos de pós-graduação *stricto sensu* em biblioteconomia e CI. Afirmando ainda que [...]

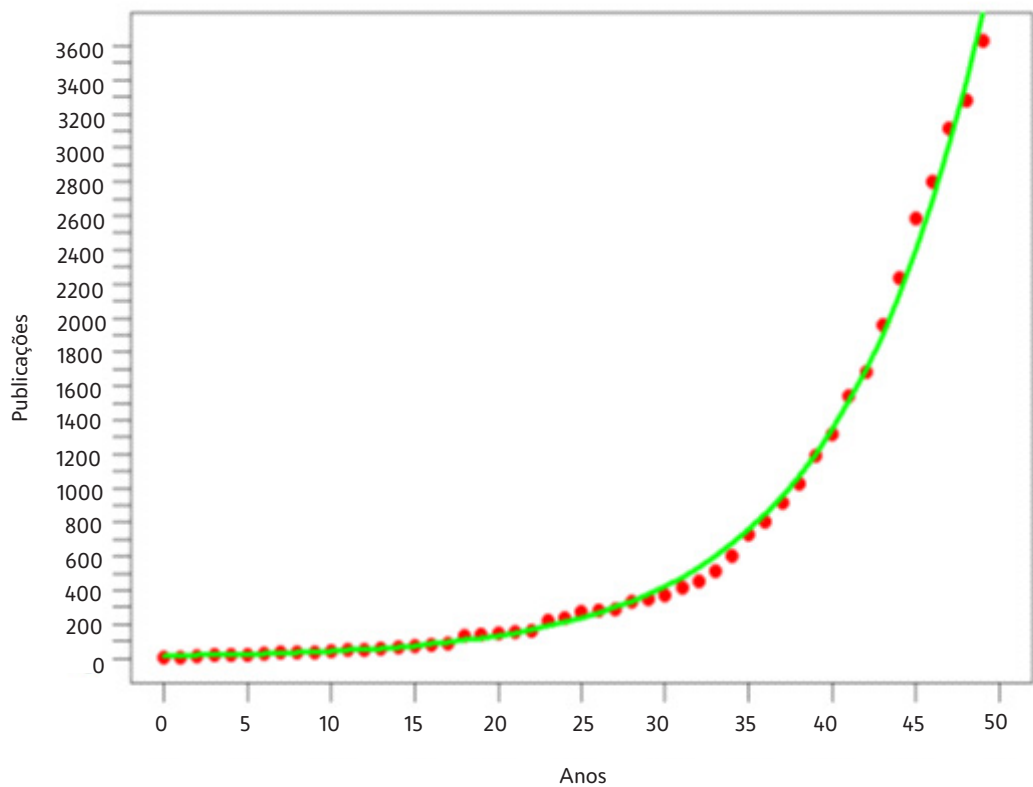


Figura 2 – Valores observados e estimados.

Fonte: Dados da pesquisa (2022).

[...] a pesquisa em OC no Brasil é oriunda principalmente dos trabalhos de conclusão de cursos de bacharelado, licenciatura, especialização, mestrado, doutorado e pós-doutorado em Biblioteconomia e Ciência da Informação e os resultados destas pesquisas são publicados em Anais de Congresso, como o CBBB, ENANCIB, SNBU, ISKO-Brasil dentre outros, como artigos em periódicos científicos e capítulos de livros (Miranda; Souza, 2021, p. 241).

Segundo as afirmações de Bräscher (2014, p. 8), a OC no Brasil:

[...] tem uma trajetória consolidada pela atuação de diversos grupos de pesquisa e pesquisadores em programas de pós-graduação da área de ciência da informação. No âmbito institucional, o Grupo de Trabalho 2 (GT2) da Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ANCIB) promove, desde 1994, o debate sobre teorias, metodologias e práticas relacionadas à organização do conhecimento em arquivos, museus, bibliotecas e outros ambientes informacionais.

Os cursos de CI distribuídos no país, os grupos de pesquisa e a atuação da ANCIB podem ser fatores detonadores desse crescimento da produção de documentos a partir de 2000, ou todas essas variáveis podem ter se combinado para produzir o crescimento dessa literatura.

A Figura 2 mostra ainda as linhas coincidentes do ajuste do modelo exponencial pelo método de regressão não linear proposto por Egghe e Ravichandra Rao (1992), obtendo-se um “R²” igual a 0,997, a um nível de significância de 0,005, com 48 graus de liberdade, o que indica que há

apenas 0,3% de possibilidade de erro na estimativa dos valores esperados. O valor estimado de c foi igual a 13,285049, e o de g , igual a 1,122381. Com esses valores conhecidos, pode-se estabelecer a seguinte equação como capaz de prever o crescimento exponencial da literatura publicada sobre a ORC no Brasil:

$$C(t) = 13,285049 \times 1,122381^t$$

Essa equação indica que essa literatura cresce a uma taxa de 12,2% ao ano e dobra de tamanho a cada 6 anos.

Resumindo, para se ter uma ideia do crescimento da literatura sobre ORC no Brasil, só 10,3% dessa literatura foi publicada nos primeiros 30 anos; no entanto, 67,2% dela foi publicada nos últimos 10 anos. Essa literatura está crescendo de forma exponencial com uma taxa de 12,2% ao ano, e duplica seu volume a cada seis anos. Essa literatura ainda não alcançou seu ponto de saturação, e tem um longo período de desenvolvimento futuro pela frente.

Modelo epidêmico

Já foi dito que, no período estudado, encontraram-se 3.630 publicações produzidas por 2.535 pesquisadores diferentes. A Tabela 1 mostra a taxa da população dos documentos publicados e dos produtores das publicações organizados segundo os quinquênios estudados desde 1972 até 2021. As taxas estimadas são todas positivas, porém a entrada dessa literatura em um período epidêmico ocorre a partir do ano 2000, quando mais autores entram no campo e começa uma produção massiva de documentos sobre esse assunto. Nos quinquênios anteriores, essa produção parece ter estado em uma situação de latência incipiente, não atingindo ainda um salto quantitativo no volume de produção de documentos. Na passagem do sexto quinquênio para o sétimo, a taxa dos documentos publicados praticamente se duplica, e a duplicação continua até o nono quinquênio, que sofreu os efeitos da epidemia de COVID-19, resultando na alteração dessa taxa de crescimento epidêmico.

A Figura 3 mostra, de forma mais explícita, essa entrada em um estado epidêmico da população comprometida com a produção de literatura sobre ORC no Brasil, assim como o volume da literatura produzida sobre esse assunto. Claramente, a partir do ano 2000, ocorrem esse salto quantitativo e a entrada dessa população em um estado epidêmico. Em outras palavras, foram

Tabela 1 – Taxa das publicações e dos autores segundo os quinquênios.

Quinquênios	Quinquênios	Número de documentos	Taxa de documentos	Número de autores	Taxa de autores	Taxa de produtividade
1972-1976	1	21	4,2	25	5,0	0,84
1977-1981	2	17	3,4	20	4,0	0,85
1982-1986	3	24	4,8	36	7,2	0,67
1987-1991	4	77	15,4	77	15,4	1,00
1992-1996	5	98	19,6	111	22,2	0,88
1997-2001	6	114	22,8	99	19,8	1,15
2002-2006	7	254	50,8	227	45,4	1,12
2007-2011	8	584	116,8	367	73,4	1,59
2012-2016	9	1049	209,8	655	131,0	1,60
2017-2021	10	1392	278,4	918	183,4	1,52
Total		3630		2535		

Fonte: Elaborada pelos autores (2022).

necessários quase 30 anos para essa população entrar em um estado epidêmico e se iniciar um crescimento acelerado tanto dos autores quanto das publicações produzidas. Os fatores detonantes desse processo já foram assinalados anteriormente.

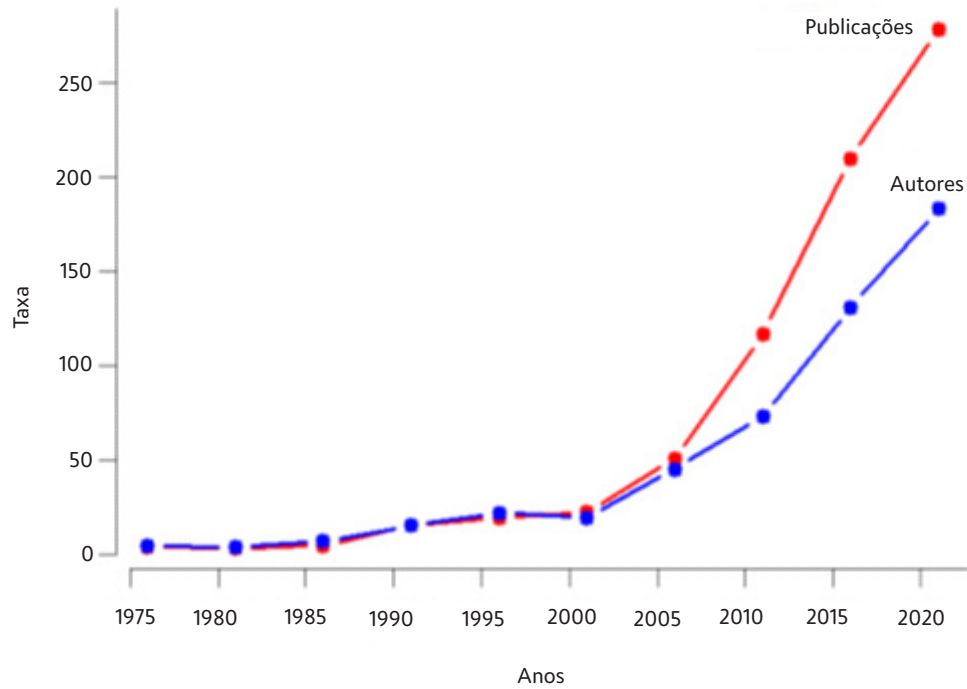


Figura 3 – Modelo de crescimento epidêmico.
Fonte: Dados da pesquisa (2022).

A Figura 4 mostra o traçado da distribuição das taxas de publicação versus os autores agrupados em intervalos de cinco anos. A proporção das publicações em relação aos autores produtores dessa literatura para toda a população, abrangendo cinquenta anos, foi $2.535/3.630 \approx 0,7$. Um corte no ponto em que a taxa da produção é de um documento por autor mostra que, depois do ano 2000, essa taxa de produtividade está abaixo de um documento por cada autor.

O número médio de publicações por autor parece ser constante, abaixo de uma publicação ao longo do tempo. Isso significa que o campo da ORC cresce, não porque os autores produzem mais artigos, mas porque mais autores entram no campo. Essa entrada no campo é facilitada pelos cursos de formação e especialização em ORC existentes no país.

Autores mais produtivos

Procurou-se identificar os autores mais produtivos sobre ORC nesses 50 anos de pesquisa. Os indicadores de produção utilizados até agora têm sido baseados na contagem crua e nua do número total de publicações realizadas pelos pesquisadores. Para identificar esses autores mais produtivos, normalmente, seria aplicada a lei de Lotka utilizando o modelo do poder inverso generalizado, porém esse modelo não leva em consideração o tempo (medido em anos) em que os autores vêm atuando em uma área de pesquisa determinada. Em outras palavras, o modelo

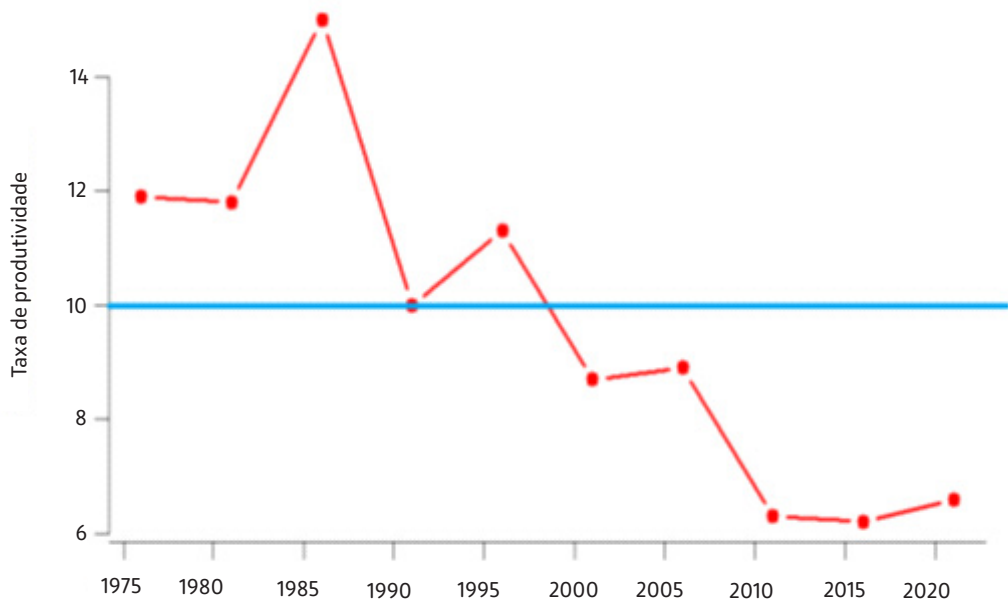


Figura 4 – Taxa de produtividade dos autores por quinquênios.
Fonte: Dados da pesquisa (2022).

privilegia aqueles autores que têm mais anos de participação na pesquisa sobre qualquer assunto. Isto é, o modelo faz competir em igualdade de condições, por exemplo, um autor que já tem 40 anos de carreira no campo com aquele que apenas tem um ou dois anos de participação nas pesquisas do mesmo campo. Para evitar esses vieses, consideramos que um autor entrou no campo no ano de sua primeira publicação e permaneceu no campo até o ano de sua última publicação. Nesse período de permanência medido em anos, esse autor produziu uma determinada quantidade de documentos. Essa quantidade de documentos publicados durante o período de permanência no campo, dividida pelo número de anos que o autor permaneceu no campo, produz o “índice de produtividade” dos autores proposto nesta pesquisa. Ainda assim, compreendemos que um autor atinge uma maturidade média de produção depois de cinco anos de permanência contínua no campo; portanto, neste caso, consideramos só os autores que têm permanecido na produção de documentos sobre ORC com um mínimo de cinco anos. Mais uma vez, destaca-se que a consolidação dos nomes dos autores foi possível pelos recursos oferecidos pelo EndNote X8, acompanhada por uma revisão cuidadosa de seus nomes.

A Tabela 2 mostra esse índice e o número de autores identificados segundo seus índices de produtividade. Nota-se como somente 20 autores são os mais produtivos quando se usa o índice de produtividade estimado maior que 0,5, e esses 20 autores representam 5% do total de autores pesquisando e publicando no campo. Ressalta-se que a ausência de um ou dois trabalhos de um determinado pesquisador pouco ou nada alteraria seu “índice de produtividade, considerando o volume do corpus construído.

Note-se também como quase 95% dos autores permanecem em uma espécie de espaço intermediário ou de fronteira; são pequenos produtores; e se agrupam em um espaço de dispersão, lembrando a lei de dispersão da literatura de Bradford. A lista desses autores mais produtivos segundo seu índice de produtividade pode ser vista na Tabela 3.

Tabela 2 – Número de autores segundo o índice de produtividade.

Índice de produtividade	Número de autores	%
6,00 – 6,99	2	0,08
5,00 – 5,99	1	0,04
4,00 – 4,99	4	0,16
3,00 – 3,99	13	0,51
2,00 – 2,99	120	4,73
1,00 – 1,99	1918	75,66
0,00 – 0,99	477	18,82
Total	2535	100,00

Fonte: Elaborada pelos autores (2022).

Tabela 3 – Taxa de produtividade dos autores com ao menos 10 anos de permanência no campo.

Autores	Publicações	Anos no campo	Taxa de produtividade
Fujita, Mariângela Spotti Lopes	232	33	7,03
Guimarães, José Augusto Chaves	139	31	4,48
Campos, Maria Luiza de Almeida	144	34	4,24
Kobashi, Nair Yumiko	119	34	3,50
Barros, Thiago Henrique Bragato	43	12	3,58
Sales, Rodrigo de	46	14	3,29
Lima, Gercina Ângela Borém de Oliveira	81	26	3,12
Pinho, Fábio Assis	48	16	3,00
Almeida, Maurício Barcellos de	53	18	2,94
Albuquerque, Ana Cristina de	37	13	2,85
Almeida, Carlos Cândido de	69	26	2,65
Souza, Renato Rocha	45	17	2,65
Martínez-Ávila, Daniel	67	26	2,58
Cervantes, Brígida Maria Nogueira	46	18	2,56
Lara, Marilda Lopes Ginez de	79	32	2,47
Tognoli, Natália Bolfarini	31	12	2,58
Moreira, Walter	60	26	2,31
Barros, Camila Monteiro de	23	10	2,30
Campos, Linair Maria	35	16	2,19
Boccatto, Vera Regina Casari	39	18	2,17
Souza, Rosali Fernandez de	55	26	2,12
Moura, Maria Aparecida	52	25	2,08
Silva, Márcio Bezerra da	24	11	2,18
Madio, Telma Campanha de Carvalho	22	10	2,20
Moraes, João Batista Ernesto de	50	25	2,00

Fonte: Elaborada pelos autores (2022).

Essa tabela apresenta somente aqueles autores que atingiram um índice de produtividade maior que 2,00 e permaneceram no campo por, pelo menos, dez anos, e pode-se interpretar que, em média, cada um dos autores produz a cada ano essa quantidade de publicações sobre ORC. Por exemplo, Mariângela Spotti Lopes Fujita produz uma média de 7,03 documentos por ano (um documento a cada um mês e três semanas); José Augusto Chaves Guimarães produz uma média

de 4,48 documentos por ano (um documento cada dois meses e meio); e Maria Luiza de Almeida Campos produz uma média de 4,24 por ano (um documento cada dois meses e meio) e assim sucessivamente para todos os autores listados na Tabela 3.

Essa taxa de produtividade mostra também como os pesquisadores da ORC no Brasil têm sido impactados pelo fenômeno que está se denominando “produtivismo”. Afirma-se que “o meio universitário foi invadido pela lógica gerencialista, o que gera muitas distorções, pois, em vez de produzir conhecimento, estamos enlatando sardinha em forma de papers” (Alcadipani, 2011, p. 275). Outros autores o chamam de “publicacionismo” para “[...] caracterizar essa prática de encaminhar à publicação análises aligeiradas, fragmentadas, repetidas, que pouco ou nada contribuiriam ao conhecimento, tendo como único móbil o propósito de multiplicar os indicadores de produção” (Kuhlmann Junior, 2015, p. 841). Os pesquisadores da ORC no Brasil não são imunes a esse impacto do produtivismo, e embora estejam cientes da importância do sistema de avaliação da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), consideram esse sistema de avaliação como um dos principais motivos para a manutenção do produtivismo acadêmico e como a “qualidade e o trato teórico das pesquisas são por vezes comprometidas por conta dos parâmetros quantitativos de avaliação que fazem com que o número de publicações seja mais valorizado que a relevância das publicações em si” (Costa; Barbosa Filho, 2021, p. 187).

Discussão e Conclusão

Analisando os padrões da organização da informação necessários para que os documentos que possuam características similares possam ser recolhidos em fontes primárias e secundárias e possam garantir confiabilidade na análise dos dados, Café e Bräscher (2008) afirmam que as dificuldades se concentram principalmente na padronização dos títulos dos periódicos e dos nomes dos autores, assim como na diversidade de representações dos assuntos, como descrição do conteúdo dos documentos. Isto é, os dados precisam de organização consistente com base em políticas explícitas e instrumentos terminológicos adequados à representação da informação, em especial tesouros e ontologias (Kobashi; Santos, 2007). Essas mesmas dificuldades foram encontradas e enfrentadas nesta pesquisa: há a necessidade do controle da “qualidade” dos dados incorporados nas bases de dados bibliográficas, especialmente no que se refere aos itens apontados anteriormente. Isso também é verdade para os editores dos periódicos e os indexadores desses documentos. Nota-se a não existência de um vocabulário controlado da ORC e suas áreas associadas e, como resultado, a descrição do conteúdo dos documentos é uma perfeita desorganização da informação.

O segundo problema é o período de pesquisa elegido para o estudo. Não é possível construir uma visão real e crível do fenômeno estudado se esse fenômeno é setorizado em pequenos períodos de tempos (Alvarenga; Silva, 2010; Alves; Grácio; Oliveira, 2011; Castro; Oliveira, 2016). Qualquer fenômeno social tem um tempo de aparecimento, continuidade, desenvolvimento e saturação, e se esse fenômeno não é estudado ao longo de sua construção e sua totalidade, não produziremos indicadores críveis, mas apenas serão mimetizados períodos não representativos do fenômeno. Portanto, é recomendável dar conta do fenômeno em sua totalidade, como neste caso: estudou-se um período de 50 anos, desde o aparecimento do fenômeno em 1972 até sua situação atual em dezembro de 2021. Dessa forma, os indicadores construídos são críveis e adequados. Por exemplo, Amorim Neto e Lima (2016, p. 101) encontraram “[...] que 53,2 % da produção brasileira sobre o domínio de OC na CI foi publicada entre 2006 e 2012 e que os demais 46,8 % dos artigos foram publicados nos 34 anos anteriores”. Nossos dados mostram que apenas 22,2% dessa literatura foi

publicada entre 2006 e 2012 e que, nos 34 anos anteriores, só se publicaram 14,2% dos documentos. Essas diferenças são o resultado da exaustividade do levantamento bibliográfico. Infelizmente, nenhum dos autores mencionados construiu os indicadores desse crescimento que nos permitisse uma comparação dos resultados pertinentes.

Essa tendência de fatiar o fenômeno em pequenos períodos produz também resultados enganosos ou parciais. A respeito dos autores mais produtivos em ORC, Pavanelli *et al.* (2015) listam 45 autores altamente produtivos nesse campo, sendo que 20 deles são brasileiros com mais de três publicações cada um. A lista dos autores “mais produtivos” dessa lista, comparada com os achados nesta pesquisa, mostra diferenças substanciais. Nesta pesquisa, encontraram-se 525 autores brasileiros produtores de três ou mais documentos, porém, na Tabela 4, apresentam-se os 25 autores que têm produzido mais de dez artigos cada em um período mínimo de dez anos. Também, Farías Blundell, Aguirre-Liguëra e Barité (2022) apresentam os autores que ocupam as 10 primeiras posições de acordo com sua produção e listam 19 autores brasileiros cujas publicações variam entre 33 e nove documentos. Para uma comparação dos resultados, só nesse nível da classificação da produção dos autores brasileiros nesta pesquisa, encontraram-se 122 pesquisadores brasileiros.

Para concluir, tanto o modelo exponencial quanto o modelo epidêmico indicam que essa literatura começou a crescer a partir do ano 2000. O modelo exponencial indica que essa literatura cresce a uma taxa de 12,2% ao ano, e duplica seu tamanho a cada 6,5 anos. No caso do modelo epidêmico, essa literatura entra em um estado epidêmico a partir do ano 2000 e cresce, não porque os autores produzem mais documentos, mas porque mais autores entram no campo. Essa entrada massiva de autores no campo é facilitada pelos programas de pós-graduação em CI, nos quais se oferecem cursos específicos sobre ORC.

Para finalizar, a base de dados em ORC do Brasil construída em EndNote X8 está em permanente atualização: uma análise da demografia dessa literatura e das redes de coautorias e coocorrência de palavras está aberta à exploração.

Referências

- Alcadipani, R. Resistir ao produtivismo: uma ode à perturbação acadêmica. *Cadernos Ebape*, n. 9, p. 1174-1178, 2011. Doi: <https://doi.org/10.1590/S1679-39512011000400015>. Acesso em: 10 out. 2022.
- Alvarenga, L.; Silva, D.L. Organização e representação do conhecimento na ciência da informação. *Tendências da Pesquisa brasileira em Ciência da Informação*, v. 3, n. 1, p. 47-84, 2010. Disponível em: <https://revistas.ancib.org/index.php/tpbci/article/view/179>. Acesso em 14 nov. 2022.
- Alves, B. H.; Grácio, M. C. C.; Oliveira, E. F. T. A produção científica da revista Scire: uma análise bibliométrica do período 2006/2010. In: Congresso Brasileiro em Organização e Representação Do Conhecimento, 1., 2011, Brasília, *Anais...* Brasília: UNB, 2011. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/134902>. Acesso em: 14 nov. 2022.
- Amorim Neto, M. R.; Lima, M. H. T.F. O domínio de organização do conhecimento na base BRAPCI: uma análise estatística. *Investigación Bibliotecológica*, v. 30, n. 70, p. 83-104, 2016. Doi: <https://doi.org/10.1016/j.ibbai.2016.10.005>. Acesso em: 10 out. 2022.
- Ancona Alcocer, M. C. Turismo de convenciones, una política pública, como fuente de ingreso, en Mérida, Yucatán, México. *Journal of Tourism and Heritage Research*, v. 3, n. 1, p. 18-28, 2020. Disponível em: <http://www.jthr.es/index.php/journal/article/view/157>. Acesso em: 14 nov. 2022.
- Bräscher, M. Organização do conhecimento. *Ciência da Informação*, v. 43, n. 1, p.8-10, 2014. Disponível em: <https://brapci.inf.br/index.php/res/download/52731>. Acesso em: 14 nov. 2022.
- Bufrem, L. S.; Breda, S. M. Presença da lógica no domínio da organização do conhecimento: aspectos interdisciplinares no currículo do ensino superior. *Perspectivas em Ciência da Informação*, v. 16, n. 1, p. 185-194, 2011. Doi: <https://doi.org/10.1590/S1413-99362011000100011>.

Café, L.; Bräscher, M. Organização da informação e bibliometria. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, p. 54-75, 2008. Doi: <https://doi.org/10.5007/1518-2924.2008v13nesp1p54>.

Carlan, E.; Medeiros, M. B. B. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. *RICI: Revista Ibero-americana de Ciência da Informação*, v. 4, n. 2, p. 53-73, 2011. Doi: <https://doi.org/10.26512/rici.v4.n2.2011.1675>

Castro, I. R.; Oliveira, M. Análise bibliométrica da produção científica sobre as linguagens de indexação publicada nos *anais* de congresso do Enancib no período de 2012 a 2015. *Biblionline*, v. 12, n. 3, p. 48-60, 2016. Disponível em: <https://periodicos.ufpb.br/ojs2/index.php/biblio/article/view/29876>. Acesso em: 14 nov. 2022.

Costa, L. F.; Barbosa Filho, E. T. O produtivismo acadêmico na Pós-Graduação stricto sensu em Ciência da Informação no Brasil. *Ciência da Informação em Revista*, v. 8, n.1, p. 165-190, 2021. Doi: <https://doi.org/10.28998/cirev.2021v8n1j>.

Costa, M. D.; Castro, G. Mapeamento de teses e dissertações sobre gestão do conhecimento em cursos de pós-graduação em ciência da informação no Brasil. In: Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação, 8., 2007, Salvador. *Anais [...]*. Salvador: UFBA, 2007. Disponível em: <https://brapci.inf.br/index.php/res/download/180291>. Acesso em: 14 nov. 2022.

Egghe, L.; Ravichandra Rao, I. K. Classification of growth models based on growth rates and its applications. *Scientometrics*, v. 25, n. 1, p. 5-46, 1992. Disponível em: <https://link.springer.com/content/pdf/10.1007/BF02016845.pdf>. Acesso em: 14 nov. 2022.

Espejo Marín, C.; Ponce Sánchez, M. D. Turismo de congresos y reuniones en la ciudad de Murcia. *Cuadernos de Turismo*, n. 3, p. 45-60, 1999. Disponível em: <https://dialnet.unirioja.es/descarga/articulo/900477.pdf>. Acesso em: 14 nov. 2022.

Fariás Blundell, V. L.; Aguirre-Liguëra, N; Barité, M. Aporte latinoamericano a la investigación en Organización del Conocimiento: una aproximación bibliométrica. *Bibliotecas. Anales de Investigación*, v. 18, n. 1, p. 1-21, 2022. Disponível em: <http://revistas.bnjm.cu/index.php/BAI/article/view/388>. Acesso em: 14 nov. 2022.

Figueiredo, F. C.; Almeida, F. G. Ontologias em ciência da informação: um estudo bibliométrico no Brasil. *Ciência da Informação*, v. 46, n. 1, p. 23-33, 2017. Doi: <https://doi.org/10.18225/ci.inf.v46i1.4011>. Acesso em: 10 out. 2022.

Fujita, M. S. L. Organização e representação do conhecimento no Brasil: análise de aspectos conceituais e da produção científica do ENANCIB no período de 2005 a 2007. *Tendências da Pesquisa Brasileira em Ciência da Informação*, v. 1, n. 1, p. 1-32, 2008. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/119329>. Acesso em: 22 nov. 2022.

Goffman, W. Mathematical approach to the spread of scientific ideas: the history of mast cell research. *Nature*, v. 212, n. 5061, p. 449-452, 1966b. Doi: <https://doi.org/10.1038/212449a0>. Acesso em: 10 out. 2022.

Goffman, W. Stability of epidemic process. *Nature*, v. 210, n. 5038, p. 786-787, 1966a. Disponível em: <https://www.nature.com/articles/210786a0>. Acesso em: 14 nov. 2022.

Goffman, W. A Mathematical method for Analyzing the Growth of a Scientific Discipline. *Journal of the Association for Computing Machinery*, v. 18, n. 2, p.173-185, 1971. Doi: <https://doi.org/10.1145/321637.321640>. Acesso em: 14 nov. 2022.

Goffman, W.; Harmon, G. Mathematical approach to the prediction of scientific discovery. *Nature*, v. 229, n. 5980, p. 103-104, 1971. Disponível em: <https://www.nature.com/articles/229103a0>. Acesso em: 14 nov. 2022.

Goffman, W.; Newill, V. A. Generalization of epidemic theory. *Nature*, v. 204, n. 4955, p. 225-228, 1964. Disponível em: <https://www.nature.com/articles/204225a0>. Acesso em: 14 nov. 2022.

Goffman, W.; Warren, K. S. The ecology of the medical literatures. *American Journal of the Medical Sciences*, v. 263, n.4, p. 267-273, 1972. Disponível em: <https://eurekamag.com/research/044/534/044534082.php/>. Acesso em: 14 nov. 2022.

Hadagali, G. S.; Gavisiddappa, A. Modeling the growth of neurology literature. *Journal of Information Science Theory and Practice*, v. 3, n. 3, p. 45-63, 2015. Doi: <https://doi.org/10.1633/JISTaP.2015.3.3.3>. Acesso em: 10 out. 2022.

Kobashi, N. Y.; Santos, R. N. M. Information organization and representation by graphic devices: interdisciplinary approach. In: Congreso Isko-España, 8., 2007, León. *Actas del [...]* León: Universidad de León, 2007. Disponível em: <https://dialnet.unirioja.es/descarga/articulo/2533501.pdf>. Acesso em: 14 nov. 2022.

Kuhlmann Junior, M. Produtivismo acadêmico, publicação em periódicos e qualidade das pesquisas. *Cadernos de Pesquisa*, v. 45, n. 158, p. 838-855, 2015. Disponível em: <https://publicacoes.fcc.org.br/cp/article/view/3597>. Acesso em: 14 nov. 2022.

Louzada, R. C. R. Sobre o crescimento da produtividade científica brasileira: anotações. *Revista Espaço Acadêmico*, ano II, n. 18, p. 1-3, nov. 2002. Disponível em: <https://periodicos.uem.br/ojs/index.php/EspacoAcademico/article/download/36401/751375139033/>. Acesso em: 14 nov. 2022.

Marques, I. H.; Gonçalves, R. B. Biblioterapia pediátrica: análise quali-quantitativa da produção em artigos, teses e dissertações brasileiras (1975-2019). *Revista ACB: Biblioteconomia em Santa Catarina*, v. 26, n. 1, p. 1-29, 2021. Disponível em: <https://revista.acbsc.org.br/racb/article/view/1729>. Acesso em: 14 nov. 2022.

Miranda, M. L. C.; Souza, R. F. As contribuições de Derek W. Langridge na organização do conhecimento: uma abordagem ao ensino e à pesquisa nos cursos de graduação e pós-graduação em biblioteconomia e ciência da informação no Brasil. *PontodeAcesso*, v. 15, n. 3, p. 228-248, 2021. Doi: <https://doi.org/10.9771/rpa.v15i3.47467>. Acesso em: 10 out. 2022.

Mora-Sánchez, J. A.; Lao-León, Y. O.; León-Robaina, R. Procedimiento para la gestión de eventos en hoteles cubanos. *Anuario de la Facultad de Ciencias Económicas y Empresariales*, n. 12, p. 344-362, 2021. Disponível em: <https://anuarioeco.uo.edu.cu/index.php/aeco/article/view/5196>. Acesso em: 14 nov. 2022.

Mugnaini, R. A bibliometria na exploração de bases de dados: a importância da Linguística. *Transinformacao*, v. 15, p. 45-52, 2003. Disponível em: <https://periodicos.puc-campinas.edu.br/transinfo/article/view/6386/4070>. Acesso em: 14 nov. 2022.

Oliveira, L. P. *Bases conceituais e formativas do tratamento temático da informação no Brasil*, 2021. 490 f. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, 2021. Disponível em: <https://repositorio.unesp.br/handle/11449/210936>. Acesso em: 14 nov. 2022.

Pavanelli, M. A. et al. Referentes teóricos basílicos na ISKO-Ibérico e suas contribuições para a Organização do Conhecimento (OC). In: Congreso ISKO España, 12.; Congreso ISKO España-Portugal, 2., 2015, Murcia. *Actas del [...]* Murcia: Universidad de Murcia, 2015. Disponível em: http://www.iskoiberico.org/wp-content/uploads/2015/11/69_Pavanelli.pdf. Acesso em: 14 nov. 2022.

Santa Anna, J.; Dias, C. C.; Maculan, B. C. M. S. M. Serviço de referência em unidades de informação: os resultados de pesquisa abordados em artigos de periódicos no Brasil. In: Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação-Febab, 28., 2019, Vitória. *Anais [...]*. São Paulo: FEBAB, 2019. Disponível em: <https://portal.febab.org.br/cbbd2019/article/view/2347>. Acesso em: 14 nov. 2022.

Santos, C. A. C. M. Organização e representação do conhecimento: bibliometria temática em artigos de periódicos brasileiros. *Revista Brasileira de Biblioteconomia e Documentação*, v. 11, n. especial, p. 640-653, 2015. Disponível em: <https://rbbd.febab.org.br/rbbd/article/view/494>. Acesso em: 14 nov. 2022.

Sales, O. M. M.; Gaudêncio, S. M.; Neves, D. A. B. Produção científica brasileira em organização e representação da informação: estudo bibliométrico nos periódicos Qualis A. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, v. 23, n. 53, p. 16-24, 2018. Doi: <https://doi.org/10.5007/1518-2924.2018v23n53p16>.

Sharma, P.; Gupta, B. M.; Kumar, S. Application of growth models to science and technology literature in research specialities. *DESIDOC Journal of Library & Information Technology*, v. 22, n. 2, p. 17-25, 2002. Doi: <https://doi.org/10.14429/djlit.22.2.3560>.

Tartarotti, R. C. D.; Fujita, M. S. L. Produção e colaboração científica em Organização e Representação do Conhecimento: análise bibliométrica do GT2 do ENANCIB no período de 2009 a 2014. *Em Questão*, v. 22, no. 3, p. 136-160, 2016. Doi: <https://doi.org/10.19132/1808-5245223.136-160>.

Urbizagástegui-Alvarado, R. Growth of literature on Bradford's law. *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información*, v. 30, n. 68, p. 51-72, 2016. Doi: <https://doi.org/10.1016/j.ibbai.2016.02.003>

Urbizagástegui, R.; Restrepo-Arango, C. Crecimiento de la literatura sobre bibliometría, informetría y cienciometría en el Brasil. *Revista Ibero-americana de Ciência da Informação*, v. 10, n. 1, p. 6-31, 2017. Doi: <https://doi.org/10.26512/rici.v10.n1.2017.2469>.

Vieira, V. O. A.; Oliveira, D. R.; Cabral, L. M. Análise dos trabalhos acadêmicos de pós-graduação de uma unidade acadêmica da UFRJ no âmbito da nova lei da biodiversidade brasileira-lei nº 13.123/2015. *Práticas em Gestão Pública Universitária*, ano 4, v. 4, no. 2, p. 178-204, 2020. Disponível em: <https://revistas.ufrj.br/index.php/pgpu/article/view/38649>. Acesso em: 14 nov. 2022.

Zancanaro, A. *et al.* Mapeamento da produção científica sobre memória organizacional e ontologias. *Perspectivas em Ciência da Informação*, 18, n. 1, p. 43-65, 2013. Disponível em: <https://brapci.inf.br/index.php/res/download/49322>. Acesso em: 14 nov. 2022.

Colaboradores

R. Urbizagástegui-Alvarado colaborou com a concepção e desenho, análise e interpretação dos dados, revisão e aprovação da versão final do artigo. M.J.M. Vogel colaborou com a análise e interpretação dos dados, revisão, normalização e aprovação da versão final do artigo.



Disponível em:

<https://www.redalyc.org/articulo.oa?id=384378477031>

Como citar este artigo

Número completo

Mais informações do artigo

Site da revista em redalyc.org

Sistema de Informação Científica Redalyc
Rede de Revistas Científicas da América Latina e do Caribe,
Espanha e Portugal
Sem fins lucrativos acadêmica projeto, desenvolvido no
âmbito da iniciativa acesso aberto

Rubén Urbizagástegui-Alvarado,

Michely Jabala Mamede Vogel

**Análise estatística da literatura publicada sobre
Organização do Conhecimento no Brasil**

**Statistical analysis of published literature on Knowledge
Organization in Brazil**

Transinformação

vol. 35, e227245, 2023

Pontifícia Universidade Católica de Campinas,

ISSN: 0103-3786

ISSN-E: 2318-0889

DOI: <https://doi.org/10.1590/2318-0889202335e227245>