



XXIV ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – XXIV ENANCIB

ISSN 2177-3688

GT 7 – Produção e Comunicação da Informação em Ciência, Tecnologia & Inovação

EM DIREÇÃO À CIENTOMETRIA FORENSE: O USO ESTRATÉGICO DOS ESTUDOS MÉTRICOS DA INFORMAÇÃO NA INVESTIGAÇÃO DE PRÁTICAS ESPÚRIAS NA CIÊNCIA

TOWARDS FORENSIC SCIENTOMETRICS: THE STRATEGIC USE OF QUANTITATIVE SCIENCE STUDIES IN INVESTIGATING FRAUDULENT PRACTICES IN SCIENCE

Karen Santos-d'Amorim – Universidade Federal de Pernambuco (UFPE)

Raimundo Nonato Macedo dos Santos – Universidade Federal de Pernambuco (UFPE)

Modalidade: Trabalho Completo

Resumo: A Cientometria Forense está emergindo como uma disciplina investigativa que integra domínios, ações e competências para identificar padrões de fraudes e comportamentos questionáveis, contribuindo para a criação de indicadores em diferentes etapas da produção do conhecimento científico, com ênfase na integridade da informação científica. Nesse contexto, este estudo investiga o uso estratégico dos Estudos Métricos da Informação (EMI) no âmbito da Cientometria Forense, ao analisar a dinâmica de artigos retratados na América Latina. Baseia-se em uma triangulação metodológica, que inclui técnicas bibliométricas, análise de citação e altmetria para identificar vulnerabilidades e oportunidades nesse campo de estudo e analisa 375 artigos retratados, com 9.617 metadados associados, evidenciando a predominância de retratações devido a má conduta. Demonstra o potencial dos EMI nos níveis macro, meso e micro desse domínio e destaca a importância da colaboração entre universidades, agências de fomento e periódicos científicos para garantir a confiabilidade das informações científicas e aumentar a confiança do público na ciência.

Palavras-chave: integridade em pesquisa; má conduta científica; artigos retratados.

Abstract: Forensic Scientometrics is emerging as an investigative discipline that integrates domains, actions, and competencies to identify patterns of fraud and questionable practices, contributing to the development of indicators at different stages of scientific knowledge production, with a focus on the integrity of scientific information. In this context, this study investigates the strategic use of Science Quantitative Studies within Forensic Scientometrics by analyzing the dynamics of retracted articles in Latin America. It is based on a methodological triangulation that includes bibliometric techniques, citation analysis, and altmetrics to identify vulnerabilities and opportunities in this domain. The study analyzes 375 retracted articles, with 9,617 associated metadata entries, highlighting the predominance of retractions due to misconduct. It demonstrates the potential of Science Quantitative Studies at the macro, meso, and micro levels of this domain and emphasizes the importance of collaboration among universities, funding agencies, and scientific journals to ensure the reliability of scientific information and strengthen public trust in science.

Keywords: research integrity; scientific misconduct; retracted articles.

1 INTRODUÇÃO

Objetos de estudos intrínsecos à Ciência da Informação, como a comunicação científica, bibliometria, cientometria e altmetria, no âmbito dos Estudos Métricos da Informação (EMI), têm, ao longo do tempo, galvanizado a atenção de cada vez mais pesquisadores para este promissor domínio (Cheng *et al.*, 2024). É consensual que os EMI fornecem importantes resultados sobre a configuração de campos científicos, domínios, tópicos, além de servirem como um complemento fundamental em outras pesquisas e diversas aplicações em Ciência e Tecnologia (Glänzel; Schoepflin, 1994; Grácio, 2020; Leydesdorff, 2001; Santos; Kobashi, 2009).

No âmbito dos EMI, uma ampla atenção acadêmica sobre o crescimento constante na taxa de retratação de artigos nas últimas duas décadas tem suscitado a aplicação recorrente de métodos e técnicas bibliométricas e cientométricas, corroborando com o avanço das discussões sobre integridade e boas práticas na pesquisa científica, e atributos adjacentes, como, práticas de pesquisa questionáveis, má conduta e melhores práticas de correção da memória científica (Fanelli, 2013; Zuckerman, 2020). Desde 2001 o aumento no número de artigos retratados tem sido notado (Fang; Steen; Casadevall, 2012), tendo o ano de 2023 representado um marco no recorde de retratações (Van Noorden, 2023). Ainda que em valores relativos, artigos retratados sejam ocorrências pouco frequentes, seu aumento contínuo e a complexidade dos desdobramentos associados vão além dos dados estatísticos meramente quantitativos.

Neste cenário, embora o campo dos Estudos Métricos da Informação tenha registrado um crescimento expressivo nas últimas duas décadas, impulsionado notadamente pelo uso de softwares de análise de publicações e de elementos como redes de autoria e citação (Elizondo; Kaltenbrunner, 2024), a compreensão de práticas espúrias, como plágio (Alzahrani; Salim; Abraham, 2012), manipulação de imagens (Bik; Casadevall; Fang, 2016), fábricas de papers (Santos-d’Amorim *et al.*, 2024b), ou cartéis de citação (Fister Junior.; Fister; Perc, 2016), também tem avançado significativamente.

Não obstante o contexto apresentado, a análise da dinâmica das publicações científicas fraudulentas ou errôneas, bem como das redes de citação de ciência falsa, ainda recebe pouca atenção e recursos humanos dos pesquisadores da própria Ciência da Informação local. Por exemplo, investigações sobre ética em pesquisa, integridade em

pesquisa ou más condutas na Ciência — termos distintos, mas correlatos—, que impactam a produção de informações científicas confiáveis, ainda são incipientes, embora alguns estudos possam ser notados em periódicos consolidados (Irizaga *et al.*, 2018; Meschini; Francelin, 2020; Rocha; Andrade, 2023; Santos-d’Amorim *et al.*, 2021a; 2021b; Santos-d’Amorim, 2022; 2024a).

É nesse contexto difuso que surge a proposta de estabelecer uma disciplina prática e interdisciplinar, visando avaliar o impacto de comportamentos antiéticos na comunidade científica, na integridade do registro da pesquisa e na confiança do público na ciência de forma mais articulada, a Cientometria Forense (McIntosh; Vitale, 2024). Tanto do ponto de vista formal quanto coloquial, o termo forense é apensado à Cientometria por sugerir algo que é por si só investigativo, a partir da utilização de metodologias ótimas, englobando princípios da Bibliometria, da Ciência e da Metaciência. Devido a seu cunho metodológico investigativo compromete-se em identificar padrões e consequências de ações, sejam estas intencionais ou não, que impactam significativamente a construção de estoques de informação científica confiáveis, a memória científica e suas ramificações—como base de dados e sistemas de informação científica—, além das repercussões dessas ações na percepção de confiabilidade da ciência pelo público e na própria governança do campo científico.

Em vista do exposto, este estudo tem como objetivo apresentar a aplicabilidade integrativa dos Estudos Métricos da Informação no âmbito da Cientometria Forense, a partir do estudo da dinâmica das retratações de artigos científicos na América Latina e suas implicações. Esta pesquisa parte da triangulação metodológica de métodos e técnicas de análise: a bibliometria, análise de citação e altmetria, cujo desenho faz-se salutar à identificação de vulnerabilidades, oportunidades e lições aprendidas relacionadas às boas práticas, à integridade em pesquisa e ao entendimento de comportamentos espúrios em diversas etapas do processo de obtenção e comunicação do conhecimento científico, de forma mais abrangente. Esta pesquisa faz parte de um estudo abrangente derivado da tese de doutorado de Santos-d’Amorim (2024a).

2 UMA DISCIPLINA AGREGADORA DE DOMÍNIOS, AÇÕES E COMPETÊNCIAS

A Cientometria Forense emerge como uma disciplina interdisciplinar focada em diversas abordagens e práticas para garantir a integridade e a segurança na pesquisa científica (McIntosh; Vitale, 2024). Estas abordagens incluem investigações de padrões de confiança e desconfiança, desenvolvimento de pedagogia especializada, aprimoramento de metodologias

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

estatísticas e computacionais, e avaliação do impacto das más condutas na ciência e na percepção pública. Por sua inerente complexidade, dez componentes são sugeridos como linhas de ações para esta disciplina, sempre com foco para a integração de estudos teóricos com a prática (McIntosh; Vitale, 2024). São eles:

a) Investigações

- Estudos de caso, pesquisas observacionais e investigações formais.
- Análise de padrões de confiança e desconfiança na comunicação científica.
- Detecção de anomalias e irregularidades que indiquem más condutas.
- Mensuração do efeito dessas práticas na integridade e segurança científica e na confiança do público.

b) Pedagogia

- Criação de cursos, workshops e recursos didáticos.
- Foco na integridade e segurança na pesquisa e comunicação científica responsável.
- Princípios e usos da Cientometria Forense para pesquisadores, estudantes e profissionais.

c) Desenvolvimento de Metodologias

- Aprimoramento de técnicas estatísticas e computacionais.
- Desenvolvimento de algoritmos e softwares para análise de publicações científicas e redes de citações.
- Identificação de potenciais infrações éticas, como manipulação de dados e plágio.

d) Pesquisa

- Estudo sistemático de fenômenos, teorias e questões pertinentes ao campo.
- Execução de experimentos, coleta e interpretação de dados.
- Formulação de teorias e divulgação dos achados em revistas científicas.

e) Avaliação de Impacto

- Análise do impacto da má conduta, retratações e desinformação.
- Efeitos nos padrões de citação e na percepção pública da ciência.
- Consequências da disseminação de resultados incorretos.

f) Desenvolvimento de Teorias

- Criação de modelos teóricos e estruturais.
- Análise das causas, motivações e processos de má conduta na pesquisa.
- Exploração das consequências para o ambiente científico.

g) Interdisciplinaridade

- Estímulo a parcerias entre disciplinas como Ética, Sociologia, Psicologia, Ciência da Informação, Negócios, Direito, Inteligência Artificial e Ciência da Computação.
- União de diferentes visões e especialidades para tratar problemas complexos de integridade na pesquisa e comunicação científica.

h) Desenvolvimento de Políticas

- Participação na criação de políticas e diretrizes.
- Fortalecimento da integridade na pesquisa e prevenção de más condutas.
- Colaboração com partes interessadas e organizações pertinentes.

i) Ética e Padrões

- Definição de padrões éticos e códigos de conduta.
- Alinhamento com diretrizes de organizações como o *Committee on Publication Ethics* (COPE) e o *International Committee of Medical Journal Editors* (ICMJE).

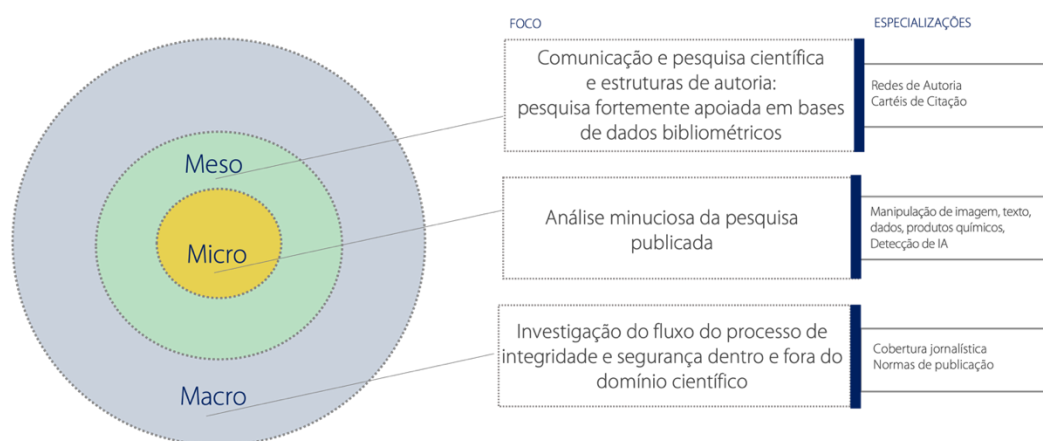
j) Desenvolvimento Profissional

- Promoção de conferências, oficinas, publicações e oportunidades de networking.

- Atualização contínua sobre avanços no campo e desenvolvimento de habilidades e conhecimentos.

Para tanto, considera-se haver diferentes focos e habilidades necessárias às partes interessadas nesta disciplina, estando distribuídos em três níveis – macro, meso e micro, como mostrado na Figura 1.

Figura 1 – Focos e especializações da Cientometria Forense



Fonte: Santos-d'Amorim (2024a)

Acredita-se, portanto, que a constituição de uma disciplina com o objetivo primeiro de manter estoques de informação científica confiáveis visando tanto os desdobramentos internos na própria Ciência e externos, como a confiança pública no conhecimento científico fiável, é salutar e premente.

3 PROCEDIMENTOS METODOLÓGICOS

Esta pesquisa descritiva parte da análise de artigos retratados no âmbito de autores afiliados a instituições de ensino e pesquisa da América Latina possibilitando a investigação de práticas espúrias na Ciência e seus desdobramentos. Estrutura-se em três etapas de análise, cuja triangulação metodológica de métodos e técnicas de análise perpassa a bibliometria, análise de citação e altmetria.

Com a finalidade de estruturar um corpus final abrangente e de mitigar vieses inerentes a levantamentos restritos, buscou-se a integração de distintas bases de dados. Para a definição dos *corpora* de entrada, os registros bibliográficos referentes a artigos retratados foram filtrados pelos países da América Latina, primariamente no conjunto de dados indexados na *Retraction Watch*, obtido por meio de acordo para uso de dados, firmado para fins específicos deste desta pesquisa. A versão utilizada deste conjunto de dados na data de análise contava com 38.579 itens, constituindo-se da base de dados mais abrangente de

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

artigos retratados disponível. Para mineração dos dados de interesse nesse amplo conjunto de dados, utilizou-se o *OpenRefine* (versão 3.7.4), software de código aberto. Neste conjunto de dados fez-se possível a identificação de 326 registros associados aos seguintes países mostrados na Tabela 2, cujo refinamento foi feito entre 14 e 17 de agosto de 2023. Em sequência, levantamentos complementares foram realizados entre fevereiro e setembro de 2023 em três bases de dados de conhecida relevância no campo científico: SciELO, *Web of Science* (WoS) e *Scopus*, conforme detalhado na Tabela 01.

Tabela 01 – Fontes de informação, estratégias de busca utilizadas e quantidade de dados de entrada

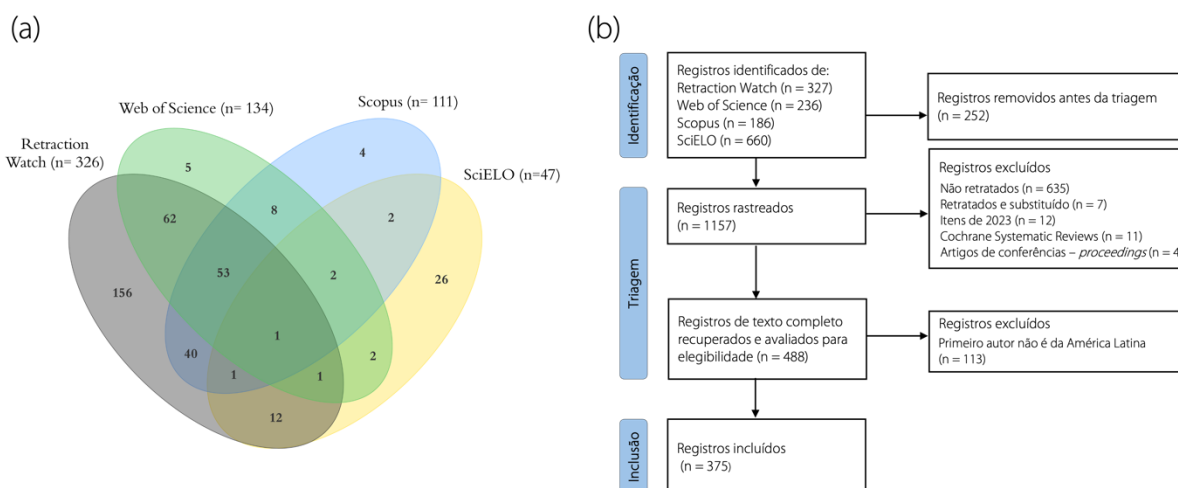
Fonte de Informação	Estratégia de busca	<i>n</i>
Retraction Watch	Foram utilizados softwares de mineração de dados para identificar e agrupar os registros de interesse disponibilizados pela RW.	327
SciELO	"artigo retratado" OR "retratação" OR "retratado" OR "retracted article" OR "retraction" OR "retracted" OR "artículo retractado" OR "retractación" OR "retractado" OR "nota de retratação" OR "aviso de retratação" OR "retraction notice" OR "carta de retractación")	660
WoS	DT=(Retraction OR Retracted Publication)	236
Scopus	PUBYEAR > 1800 AND PUBYEAR < 2023 AND (LIMIT-TO (SRCTYPE , "j")) AND (LIMIT-TO (AFFILCOUNTRY , "Brazil") OR LIMIT-TO (AFFILCOUNTRY , "Mexico") OR LIMIT-TO (AFFILCOUNTRY , "Colombia") OR LIMIT-TO (AFFILCOUNTRY , "Chile") OR LIMIT-TO (AFFILCOUNTRY , "Argentina") OR LIMIT-TO (AFFILCOUNTRY , "Peru") OR LIMIT-TO (AFFILCOUNTRY , "Venezuela") OR LIMIT-TO (AFFILCOUNTRY , "Uruguay") OR LIMIT-TO (AFFILCOUNTRY , "Ecuador") OR LIMIT-TO (AFFILCOUNTRY , "Dominican Republic") OR LIMIT-TO (AFFILCOUNTRY , "Bolivia") OR LIMIT-TO (AFFILCOUNTRY , "Cuba") OR LIMIT-TO (AFFILCOUNTRY , "Costa Rica") OR LIMIT-TO (AFFILCOUNTRY , "Honduras")) AND (LIMIT-TO (DOCTYPE , "tb"))	186

Fonte: Santos-d’Amorim (2024)

No âmbito da Etapa 1 da triangulação metodológica, os registros recuperados foram salvos em arquivos *comma separated values* (.CSV) e também compilados em um arquivo único, no formato de planilha *Microsoft® Excel.xlsx* (versão 16.76, de 2023, para Mac). Para a limpeza e mineração dos dados recuperados, e posteriormente das análises relacionais, o software *VantagePoint* (versão 11) foi utilizado, a fim de detectar inconsistências, valores ausentes nos itens e duplicidades.

Os critérios de inclusão adotados consistiram em verificar se os registros consistiam em artigos (retratados) publicados em periódicos científicos e se o primeiro autor pertencia a instituições localizadas nos países da América Latina. Em conclusão aos procedimentos de seleção, triagem, identificação de duplicidades (Fig. 01a), obteve-se como corpus final 375 registros, com 9.617 metadados associados, conforme sumarizado na Figura 01(b).

Figura 01 – (a) Diagrama de Venn de sobreposições em função das fontes de informação utilizadas e (b) fluxograma do processo de triagem, exclusão e inclusão de registros do corpus



Fonte: Elaborada pelos autores

No que se refere a etapa analítica dos motivos de retratação, considerando que um artigo pode ter sido retratado por mais de um motivo, todas as ocorrências indicadas nos respectivos avisos de retratação foram consideradas para cada registro. Portanto, esses avisos serviram como fonte de informação para a categorização das razões para retratação em cinco grupos: 1) Erros - Desvios que tipicamente ocorrem involuntariamente/acidentalmente, por negligências (Shamoo; Resnik, 2009); 2) Erros Editoriais - cometidos pela entidade publicadora (*publisher*), sem ingerência dos autores; 3) Más condutas - violações das boas práticas de pesquisas, de modo a distorcer os achados e os registros de pesquisa, prejudicando a integridade do processo de pesquisa e dos pesquisadores (Shamoo; Resnik, 2009; Resnik *et al.*, 2015); 4) Condutas questionáveis - Práticas contestáveis, discutíveis, duvidosas e problemáticas (Sijtsma; Veldkamp; Wicherts, 2016); e 5) Não categorizáveis - Itens sem informações suficientes.

Para a análise das áreas do conhecimento, foi utilizada a classificação primária do periódico, com base na Tabela de Áreas de Conhecimento/Avaliação da CAPES. Foram identificados o 1º nível (Grande Área) e o 4º nível (Especialidade) para cada artigo retratado.

Na Etapa 2, levantou-se o número de citações de cada artigo do corpus para caracterizar quantitativamente suas citações, agregando dados sobre a frequência das mesmas. Esse levantamento foi realizado entre 14 e 16 de março de 2024, utilizando a ferramenta de busca do Google Acadêmico. Essa escolha se deve ao fato de a plataforma não restringir citações a coleções indexadas em bases de dados específicas, proporcionando, assim, informações mais abrangentes e alinhadas à realidade. Conforme observado por Martín-Martín *et al.* (2018, p. 1160), apesar de desafios como “citações perdidas”, os dados

de citação do Google Acadêmico são, “essencialmente, um superconjunto da WoS e da Scopus, com cobertura extra substancial”, sendo uma ferramenta válida para análise de dados de citação.

Por fim, na Etapa 3, para obtenção das métricas de atenção on-line desses artigos foi utilizado o *Altmetrics Bookmarklet* da *Altmetric*, também entre 14 e 16 de março de 2024.

Por escolha metodológica, os números relativos aqui utilizados foram primariamente provenientes da Web of Science Core Collection (WoSCC), via acesso CAFE-CAPES, filtrando-se por “artigos” e “artigos de revisão” em “*document type*” e pelos países da América Latina aqui considerados para análise, no período correspondente aos números absolutos, isto é, de 2002 a 2022. No caso dos periódicos não indexados na WoSCC recorreu-se aos dados disponíveis na base de dados SciELO. Estatística descritiva foi utilizada para analisar os resultados obtidos.

4 RESULTADOS E DISCUSSÕES

Um total de 375 artigos retratados entre 2002 e 2022 compuseram o *corpus* de entrada desta pesquisa. Neste conjunto de documentos foram identificados 1.588 autores distribuídos entre 304 afiliações, 10 grandes áreas do conhecimento e 63 especialidades. Foram identificados 148 autores reincidentes, que, individualmente acumulam de dois (2) a 18 artigos retratados. Utilizando o método de contagem completa (Gauffriau, 2021), foram identificados 100 autores com até duas ocorrências de retratação.

Ao analisar a evolução temporal dessas retratações ao longo de 20 anos, observa-se que durante a série histórica há uma tendência de aumento tanto em números absolutos quanto em números relativos.

Foi possível identificar a prevalência de retratações por más condutas, em que foram observados 168 artigos retratados, o que corresponde a 44,8% dos artigos do corpus. Por conseguinte, erros editoriais aparecem em segundo lugar, com 77 ocorrências (20,5%), seguidos de erros inadvertidos dos autores (62 ocorrências, 16,5%), condutas questionáveis (33 ocorrências, 8,8%). Não foi possível categorizar 35 registros, dada as limitações de informação nos avisos de retratação. O aumento de registros de artigos retratados em função do tempo, e também o aumento por más condutas tem sido resultado observado com frequência em estudos anteriores, desde os seminais, como na investigação de Fang, Steen e Casadevall (2012) e em Bik, Casadevall e Fang (2016).

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

A Tabela 02 sintetiza os dados obtidos por meio das fontes de coleta utilizadas, apresentando também os valores proporcionais à produção científica total indexada na *Web of Science* (WoS) para cada país, o que possibilitou a mensuração da taxa de retratação.

Tabela 02 – Síntese dos dados de artigos retratados

Posição		País	Artigos retratados	Produção WoS ¹	Taxa ²
Relativa	Absoluta				
01	10	Paraguai	01	2.676	3,73
02	10	Guatemala	01	3.384	2,95
03	01	Brasil	236	870.677	2,71
04	04	Colômbia	25	96.361	2,59
05	09	Panamá	02	8.270	2,41
06	10	Bolívia	01	4.960	2,01
07	06	Peru	06	30.158	1,98
08	07	Equador	05	26.569	1,88
09	02	México	46	282.690	1,62
10	03	Argentina	29	185.579	1,56
11	08	Cuba	03	21.618	1,38
12	08	Venezuela	03	23.945	1,25
13	05	Chile	16	158.607	1,00
14	10	Costa Rica	01	14.035	0,71
		Total	375		

¹Levantamento feito considerando os documentos (artigos e artigos de revisão) indexados na WoSCC entre 2002-2022, a partir de dados coletados em 06 de fevereiro de 2024; filtros temporais e de documentos aplicados vide metodologia. ²Por 10.000 publicações.

Fonte: elaborada pelos autores

A partir da análise das áreas do conhecimento foi possível identificar que dentre as dez grandes áreas identificadas, as Ciências da Saúde (n= 169) possuem o maior número de artigos retratados. Já em termos percentuais, as áreas com maior proporção de artigos retratados são as Ciências Humanas, com uma taxa de 314.24 artigos retratados para cada 10.000 artigos publicados, as Ciências Sociais Aplicadas, com uma taxa de 68.57, seguidas da Linguística, Letras e Artes, com 59.12.

4.1 Artigos retratados continuam sendo citados

Ao recuperar os dados de citações desses artigos foi possível identificar que 285 artigos retratados (76,0%) receberam ao menos uma citação, totalizando 9.662 citações. Observou-se uma distribuição assimétrica, na qual alguns trabalhos acumulam um número elevado de citações. Um exemplo disso é o artigo retratado mais citado, com 1.041 citações. A heterogeneidade nas citações de artigos retratados motivou uma investigação mais detalhada, focada nas particularidades identificadas. A Tabela 03 apresenta os artigos retratados que receberam 200 ou mais citações.

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

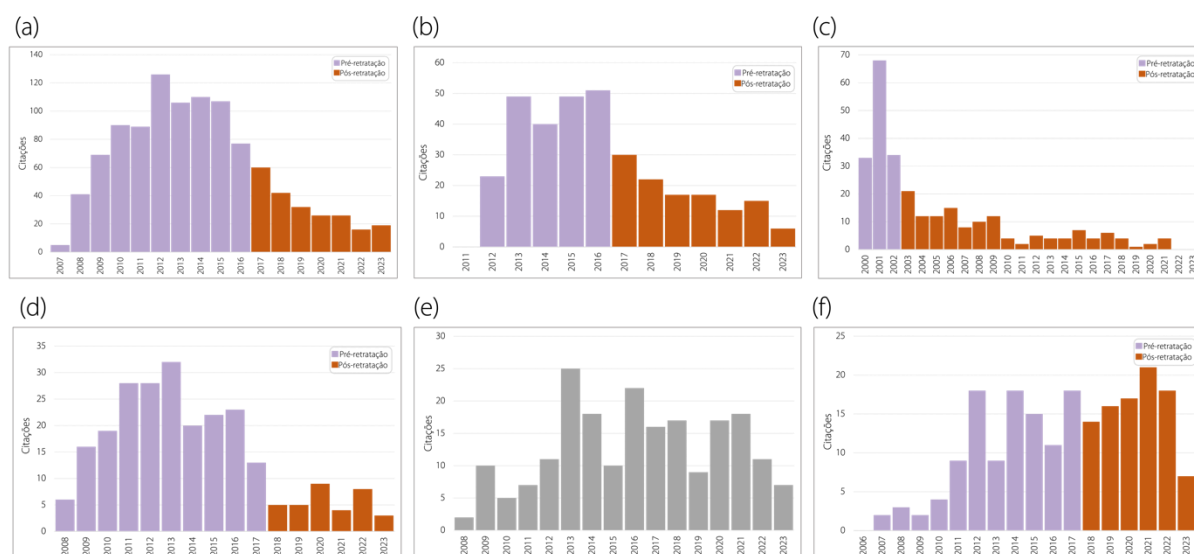
Tabela 03 – Artigos retratados mais citados

(ID) Digital Object Identifier	Quantidade de citações	Área do Conhecimento	Razão(ões) para retratação
(a) 10.2337/db06-1595	1.041	Saúde	Suspeita de falsificação/ manipulação de imagens
(b) 10.1371/journal.pbio.1001212	331	Biológicas	Suspeita de falsificação/ manipulação de imagens
(c) 10.1038/35001001	272	Biológicas	Erros em dados; Erros em análises
(d) 10.2337/db07-0573	241	Saúde	Suspeita de falsificação/ manipulação de imagens
(e) 10.1590/S0034-70942007000100006	205	Saúde	Não informado
(f) 10.1590/1413-82712017220315	202	Humanas	Direitos autorais

Fonte: elaborada pelos autores

As análises de citação pré-retratação e pós-retratação dos artigos retratados mais citados do corpus (≥ 200 citações) mostram que todas essas publicações continuaram sendo citadas mesmo após sua retirada. Essa tendência é ilustrada na Figura 02, na qual os IDs fornecidos na tabela anterior estão associados às letras de (a) a (f) exibidas.

Figura 02 – Análise de citações pré- e pós-retratação dos artigos mais citados (≥ 200 citações)



Fonte: Santos-d'Amorim (2024a)

Analisados em função das áreas do conhecimento, estes dados mostraram diferenças nas médias e nas somas dessas citações. Por exemplo, as Ciências Biológicas e as Ciências da Saúde se destacam com as maiores médias de citações nesses artigos retratados, atingindo cerca de 34 e 31, respectivamente. Em contraste, as áreas com as menores médias de citações são Linguística, Letras e Artes e Ciências Sociais Aplicadas, de 2,3 e 8,3.

Razões variadas, como (i) má sinalização de uma retratação, (ii) inconsistência de metadados em bases de dados e sites oficiais de periódicos, (iii) falta de atenção dos autores

citantes, (iv) falta de leitura do texto original, (v) uso de cópias impressas em bibliotecas, (vi) uso inconsciente da versão de registro salva localmente em máquinas pessoais, (vii) autoarquivamento sem a atualização e marca d'água de "retratado" em redes sociais acadêmicas, como *ResearchGate*, *Academia.edu*, ou em sites piratas como *Sci-Hub*, e (viii) até mesmo o uso proposital de autores que tiveram artigos retratados (autocitação) para aumentar seus rendimentos bibliométricos de citação (Bucci, 2019; Hsiao; Schneider, 2022; Suelzer *et al.*, 2021; Silva; Bornemann-Cimenti, 2017) denotam os motivos pelos quais artigos retratados continuam sendo citados.

4.2 Artigos retratados chamam atenção em plataformas digitais

A análise altmétrica dos dados permitiu identificar uma média de 37,7 de menções nos artigos retratados do *corpus*, em que 41,3% itens (n= 155) alcançaram pontuações que variaram de 1 a 9.534. Entre os dez artigos com as maiores pontuações altmétricas, sete mostraram ampla visibilidade em plataformas de mídia social, destacando-se no X (Ex-Twitter) e no Facebook, com uma inclinação particular para o X. Esses artigos retratados pertencem à área das Ciências da Saúde e receberam atenção adicional também por meio de menções em veículos de notícias, blogs, na *Wikipédia*, no *Reddit*, no *YouTube* e em registros de pedidos de patentes.

Os artigos de destaque incluem pesquisas sobre a COVID-19, realizadas em cooperação com pesquisadores de outros países. Demais assuntos com maior propensão a serem mencionados on-line incluem, mas não se limitam a tratamentos em saúde e bem-estar e o uso de medicamentos.

Como mostrado na Tabela 04, dentre as 22.337 menções identificadas em ambientes on-line no *corpus* da pesquisa, a plataforma X (antigo Twitter) liderou o ranking de menções, com 21.648 ocorrências (96,9%) – incluindo postagens originais e compartilhamentos, seguidos de outras plataformas.

Tabela 04 – Sumário de menções em plataformas via Altmetric

Plataforma	Quantidade de menções totais	%
X (antigo Twitter)	21.648	96,91
Facebook	251	1,12
Veículos de notícias	148	0,66
Blogs	109	0,49

(continua)

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

		(conclusão)
Reddit	50	0,22
Vídeos	37	0,17
Patentes	31	0,14
Sites de revisão por pares pós-publicação	29	0,13
Páginas da Wikipédia	12	0,06
Google + (descontinuado)	10	0,05
Documentos de diretrizes	06	0,03
Tópicos de perguntas e respostas	03	0,01
Destaques de pesquisa	03	0,01

Fonte: elaborada pelos autores (2024)

Da mesma forma que os dados de citação, o que se observa é que o potencial de menções em ambientes on-line parece depender de fatores, como o tema da pesquisa. Por exemplo, temas que influenciam a opinião pública, como vacinas, podem fomentar maior atenção e a formação de grupos de pessoas dentro e fora do ecossistema científico que se posicionam contra vacinas. Somado a isso, tem sido cada vez mais evidenciado que o ambiente do ex-Twitter tem proporcionado maior amplificação algorítmica a temas anticiência e antidemocráticos (Bradshaw, 2023; Ginossar *et al.*, 2023; Unlu *et al.*, 2024). A situação pode se agravar ainda mais quando as motivações pessoais e ideológicas dos proprietários de plataformas se cruzam com o vasto poder de disseminação em massa, amplificado pelas tecnologias que fundamentam esses ambientes on-line.

5 CONSIDERAÇÕES FINAIS

A análise de artigos retratados decorrentes de autores vinculados a instituições da América Latina demonstra um aumento constante de retratações motivadas por más condutas. Esses dados reforçam a necessidade de uma abordagem investigativa integrada, como proposto pela Cientometria Forense, que, no âmbito dos Estudos Métricos da Informação (EMI) inclui componentes como investigações formais, o aprofundamento analítico das redes de citação e o impacto da desinformação científica, especialmente em plataformas digitais, por meio da altmetria.

Os resultados deste estudo destacam a crescente relevância da Cientometria Forense como uma disciplina fundamental para unir pesquisadores e suas competências no enfrentamento dos desafios relacionados à má conduta científica e à preservação da integridade da informação científica. Nesse contexto, os EMI desempenham um papel-chave,

estabelecendo uma relação simbiótica nesse domínio nos três níveis analíticos propostos – macro, meso e micro.

Ainda, a análise das citações pós-retratação demonstrou que, mesmo após a retratação dos artigos, muitos continuam a ser amplamente citados. O que compromete não apenas a confiança pública na ciência, como também perpetua a circulação de informações errôneas, reforçando a necessidade de intervenções mais rigorosas e de políticas claras que conectem a comunicação científica às melhores práticas de integridade.

Ao comparar esses achados com o estado da arte discutido, conclui-se que, embora a vigilância sobre a má conduta científica tenha aumentado, o sistema acadêmico ainda enfrenta desafios para reduzir sua tolerância a essas práticas. Um exemplo claro disso é a persistência de citações a artigos retratados, o que indica que, em alguns casos, pesquisadores continuam agindo de forma inadequada como autores, seja por negligência ou intencionalmente. Por fim, é salutar destacar a importância da colaboração entre universidades, agências de fomento e periódicos científicos para reforçar a confiabilidade da ciência e promover uma cultura de integridade.

REFERÊNCIAS

ALZHRANI, Salha; SALIM; Naomie; ABRAHAM, Ajith. Understanding plagiarism linguistic patterns, textual features, and detection methods. **IEEE**, New York, v. 42, n. 2, p. 133–149, 2012. DOI: <https://doi.org/10.1109/TSMCC.2011.2134847>

BIK, Elisabeth M.; CASADEVALL, Arturo; FANG, Ferric C. The prevalence of inappropriate image duplication in biomedical research publications. **mBio**, Washington, v. 7, n. 3, 2016. DOI: <https://doi.org/10.1128/mBio.00809-16>

BRADSHAW, Amanda S. Doctors speak up: exploration of hashtag hijacking by anti-vaccine advocates and the influence of scientific counterpublics on twitter. **Health Commun**, London, v. 38, n. 10, p. 2167–2177, 2023. DOI: <https://doi.org/10.1080/10410236.2022.2058159>

BUCCI, Enrico M. On zombie papers. **Cell Death Dis**, London, v. 10, p. 1–2, 2019. DOI: <https://doi.org/10.1038/s41419-019-1450-3>

CHENG, K.; LI, Zhiyong; SUN, Zaijie; GUO, Qiang; LI, Wanqing; LU, Yanqiu; QI, Shaoyan; SHEN, Zefeng; XIE, Ruijie; WANG, Yulin; WU, Zhenjiang; WU, Yinhu; WU, Chunchun; LI, Yuqiao; XIE, Yuchen; WU, Haiyang; LI, Cheng. The rapid growth of bibliometric studies: a call for international guidelines. **International Journal of Surgery**, London, v. 110, n. 4, p. 2446–2448, 2024. DOI: <https://doi.org/10.1097/JS9.0000000000001049>

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

ELIZONDO, Andrea Reyes; KALTENBRUNNER, Wolfgang. Navigating the Science System: Research Integrity and Academic Survival Strategies. **Sci Eng Ethics**, England, v. 30, 2024. DOI: <https://doi.org/10.1007/s11948-024-00467-3>

FANELLI, Daniele. Why growing retractions are (mostly) a good sign. **PLOS Med**, San Francisco, v. 10:e1001563, 2013. DOI: <https://doi.org/10.1371/journal.pmed.1001563>

FANG, Ferric C.; STEEN, R. Grant; CASADEVALL, Arturo. Misconduct accounts for the majority of retracted scientific publications. **PINAS**, Washington, v. 109, n. 42, p. 17028–17033, 2012. DOI: <https://doi.org/10.1073/pnas.1212247109>

FISTER JUNIOR, Iztok; FISTER, Iztok; PERC, Matjaž. Toward the discovery of citation cartels in citation networks. **Frontiers in Physics**, Lausanne, v. 4, 2016. DOI: <https://doi.org/10.3389/fphy.2016.00049>

GAUFFRIAUX, Marianne. Counting methods introduced into the bibliometric research literature 1970–2018: A review. **Quantitative Science Studies**, Cambridge, v. 2, n. 3, p. 932–975, 2021. DOI: https://doi.org/10.1162/qss_a_00141

GINOSSAR, Tamar; CRUICKSHANK, Iain. J.; ZHELEVA, Elena; SULSKIS, Jason; BERGER-WOLF, Tanya. Cross-platform spread: vaccine-related content, sources, and conspiracy theories in YouTube videos shared in early Twitter COVID-19 conversations. **Hum Vaccin Immunother**, Philadelphia, v. 18, n. 1, p. 1–13, 2022. DOI: <https://doi.org/10.1080/21645515.2021.2003647>

GLÄNZEL, Wolfgang; SCHOEPFLIN, Urs. Little scientometrics, big scientometrics and beyond? **Scientometrics**, Amsterdam, v. 30, n. 2–3, p. 375–384, 1994. DOI: <https://doi.org/10.1007/BF02018107>

GRÁCIO, Maria Cláudia Cabrini. Estudos métricos da informação. In: GRÁCIO, Maria Cláudia Cabrini. **Análises relacionais de citação para a identificação de domínios científicos: uma aplicação no campo dos Estudos Métricos da Informação no Brasil**. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2020. p. 19-75. Disponível em: <https://books.scielo.org/id/tx83k/pdf/gracio-9786586546125-02.pdf>. Acesso em: 17 abr. 2024.

HSIAO, Tzu-Kun; SCHNEIDER, Jodi. Continued use of retracted papers: Temporal trends in citations and (lack of) awareness of retractions shown in citation contexts in biomedicine. **Quant Sci Stud**, Cambridge, v. 2, n. 4, p. 1144–1169, 2022. DOI: https://doi.org/10.1162/qss_a_00155

IRIZAGA, Karen Ribeiro de Freitas; MONTEIRO-KREBS, Luciana; BETTIO, Maiara; ROCKEMBACH, Moisés. Questões éticas na comunicação científica. **Prisma.com**, Porto, v., n. 36, 2018. DOI: <http://dx.doi.org/10.21747/16463153/36a9>

LEYDESDORFF, Loet. **The challenge of Scientometrics: the development, measurement, and self-organization of scientific communication**. 2 ed. Leiden: DSWO Press, 2001. DOI: <https://doi.org/10.2139/ssrn.3512486>

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

MARTÍN-MARTÍN, Alberto; ORDUNA-MALEA, Enrique; THELWALL, Mike; LÓPEZ-CÓZAR, Emilio Delgado. Google Scholar, Web of Science, and Scopus: a systematic comparison of citations in 252 subject categories. **Journal of Informetrics**, Amsterdam, v. 12, n. 4, p. 1160–1177, 2018. DOI: <https://doi.org/10.1016/j.joi.2018.09.002>

MCINTOSH, Leslie D.; VITALE, Cynthia Hudson. Forensic scientometrics: an emerging discipline to protect the scholarly record. **arXiv**, [S. l.], 2024. Disponível em: <http://arxiv.org/abs/2404.00478>. Acesso em:??

MESCHINI, Fabio Orsi; FRANCELIN, Marivalde Moacir. Produção científica brasileira sobre plágio: caracterização e alcance a partir da base Scopus. **Encontros Bibli**, Florianópolis, v. 25, p. 01–26, 2020. DOI: <https://doi.org/10.5007/1518-2924.2020.e70258>

RESNIK, David B.; NEAL, Talicia; RAYMOND, Austin; KISSLING, Grace E. Research misconduct definitions adopted by U.S. Research Institutions. **Account Res**, London, v. 22, n. 1, p. 14–21, 2015. DOI: <https://doi.org/10.1080/08989621.2014.891943>

ROCHA, E. S. S.; ANDRADE, D. R. S. Integridade científica nos periódicos de Ciência da Informação: análise de conteúdo das diretrizes para submissão de artigos. **TransInformação**, Campinas, v. 35, p. 1-12, 2023. DOI: <https://doi.org/10.1590/2318-0889202335e220017>

SANTOS, Raimundo Nonato Macedo; KOBASHI, Nair Yumiko. Bibliometria, cientometria e infometria: conceitos e aplicações. **Pesq. Bras. Ci. Inf.**, Brasília, DF, v. 2, n.1, p.155–172, 2009. Disponível em: <https://revistas.ancib.org/index.php/tpbci/article/view/174/174>. Acesso em: 16 jan. 2023

SANTOS-D'AMORIM, Karen. **Dinâmica das retratações de artigos científicos na América Latina e implicações na cultura de integridade em pesquisa e governança na Ciência**. 2024a. Tese (Doutorado em Ciência da Informação) – Programa de Pós-graduação em Ciência da Informação, Universidade Federal de Pernambuco, Recife, 2024a.

SANTOS-D'AMORIM, Karen; CORREIA, Anna Elizabeth Galvão Coutinho; MIRANDA, Májory Karoline Fernandes de Oliveira; SANTA-CRUZ, Petrus. Reasons and implications of retracted articles in Brazil. **TransInformação**, Campinas, v. 33, 2021b. DOI: <https://doi.org/10.1590/2318-0889202133e210001>

SANTOS-D'AMORIM, Karen; WANG, Ting; LUND, Brady; SANTOS, Raimundo Nonato Macedo. From plagiarism to scientific paper mills: a profile of retracted articles within the SciELO Brazil collection. **Ethics & Behavior**, London, v. 34, n. 1, p. 40–57, 2024b. DOI: <https://doi.org/10.1080/10508422.2022.2141747>

SANTOS-DAMORIM, Karen. Integridade na pesquisa no âmbito da conjuntura pandêmica de covid-19: um mapeamento bibliométrico. **AtoZ: Novas Práticas em Informação e Conhecimento**, Curitiba, v. 11, p. 1–11, 2022. DOI: <http://dx.doi.org/10.5380/atoz.v11.84325>

SANTOS-DAMORIM, Karen; SANTOS, Raimundo Nonato Macedo; MELO, Rinaldo Ribeiro. Retratações e citações pós-retratação na infodemia de covid-19: a academia está espalhando desinformação?. **Liinc em revista**, Rio de Janeiro, v. 17, n.1, 2021a. DOI: <https://doi.org/10.18617/liinc.v17i1.5593>

XXIV Encontro Nacional de Pesquisa em Ciência da Informação – XXIV ENANCIB
Vitória-ES – 04 a 08 de novembro de 2024

SHAMOO, Adil E.; RESNIK, David B. **Responsible conduct of research**. New York: Oxford University Press, 2009.

SIJTSMA, Klaas; VELDKAMP, Coosje L. S.; WICHERTS, Jelte M. Improving the conduct and reporting of statistical analysis in psychology. **Psychometrika**, North Carolina, v. 81, p. 33–38, 2016. DOI: <https://doi.org/10.1007/s11336-015-9444-2>

SILVA, Jaime A Teixeira.; BORNEMANN-CIMENTI, Helmar. Why do some retracted papers continue to be cited? **Scientometrics**, Amsterdam, v. 110, p. 365–370, 2017. DOI: <https://doi.org/10.1007/s11192-016-2178-9>

SUELZER, Elizabeth M.; DEAL, Jennife; HANUS, Karen; RUGGERI, Barbara E.; WITKOWSKI, Elizabeth. Challenges in identifying the retracted status of an article. **JAMA Network Open**, Chicago, v. 4, n. 6, 2021. DOI: <http://dx.doi.org/10.1001/jamanetworkopen.2021.15648>

UNLU, Ali; TRUONG, Sophie; SAWHNEY, Nitin; SIVELÄ, Jonas; TAMMI, Tuukka. Long-term assessment of social amplification of risk during COVID-19: challenges to public health agencies amid misinformation and vaccine stance. **J Comput Soc Sci**, Singapore, 2024. DOI: <https://doi.org/10.1007/s42001-024-00257-8>

VAN NOORDEN, Richard. More than 10,000 research papers were retracted in 2023 — a new record. **Nature**, London, v. 624, p. 479–481, 2023. DOI: <https://doi.org/10.1038/d41586-023-03974-8>

ZUCKERMAN, Harriet. Is the time ripe for quantitative research on misconduct in science? **Quantitative Science Studies**, Cambridge, v. 1, n. 3, p. 945–958, 2020. DOI: https://doi.org/10.1162/qss_a_00065