26º ENANCIB
Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação
A Amazônia na sociedade global em rede: assimetrias e potencialidades informacionais, tecnológicas e ambientais
Belém, Pará, 09 a 13 de novembro de 2026
Universidade Federal do Pará

ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
ISSN 2177-3688

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

CAMPO DA SAÚDE NA OPENALEX: ORGANIZAÇÃO TEMÁTICA E REDES INSTITUCIONAIS DA PRODUÇÃO BRASILEIRA

HEALTH FIELD IN OPENALEX: *THEMATIC ORGANIZATION AND INSTITUTIONAL NETWORKS OF BRAZILIAN PRODUCTION*

Rogéria Veronese – Universidade Federal de São Carlos (UFSCar)
Raphael Augusto dos Santos – Universidade Federal de São Carlos (UFSCar)
Ariadne Chloe Mary Furnival – Universidade Federal de São Carlos (UFSCar)
Leandro Innocentini Lopes de Faria – Universidade Federal de São Carlos (UFSCar)
Maria Laura Machado Cangiani – Universidade Federal de São Carlos (UFSCar)

Modalidade: Trabalho Completo

Resumo: Este estudo analisa a produção brasileira em acesso aberto sobre Saúde Pública e Saúde Coletiva indexada na OpenAlex, entre 2015 e 2025, com o objetivo de compreender sua organização temática e suas redes institucionais de colaboração. A pesquisa possui natureza quantitativa e exploratória, com apoio de procedimentos bibliométricos e análise de redes sociais. O corpus foi composto por 28.260 artigos recuperados na OpenAlex e processados no VOSviewer. Os resultados mostram composição temática heterogênea, marcada pela convivência entre abordagens clínicas, epidemiológicas, sociais e ambientais da saúde. A rede de coautoria institucional revela concentração em centros consolidados, como Fiocruz, USP, UFRJ, UFBA e UFPE, indicando que a circulação científica permanece associada a instituições com maior capacidade de articulação colaborativa. Sob a perspectiva da Ciência da Informação, o estudo evidencia que bases abertas de metadados não funcionam apenas como fontes para estudos métricos, mas participam da organização da visibilidade científica. Conclui-se que a Ciência Aberta amplia o acesso aos registros bibliográficos, mas não elimina automaticamente assimetrias na produção e circulação do conhecimento científico.
Palavras-chave: OpenAlex; saúde coletiva; saúde pública; bibliometria; comunicação científica.

Abstract: This study analyzes Brazilian open access scientific production on Public Health and Collective Health indexed in OpenAlex between 2015 and 2025, aiming to understand its thematic organization and institutional collaboration networks. The research is quantitative and exploratory, supported by bibliometric procedures and social network analysis. The corpus comprised 28,260 articles retrieved from OpenAlex and processed in VOSviewer. The results show a heterogeneous thematic composition, marked by the coexistence of clinical, epidemiological, social, and environmental approaches to health. The institutional co-authorship network reveals concentration around consolidated centers, such as Fiocruz, USP, UFRJ, UFBA, and UFPE, indicating that scientific circulation remains associated with institutions with greater collaborative capacity. From the perspective of Information Science, the study shows that open metadata databases do not function merely as sources for metric studies, but also participate in the organization of scientific visibility. It concludes that Open Science expands access to bibliographic records, but does not automatically eliminate asymmetries in the production and circulation of scientific knowledge.
Keywords: OpenAlex; collective health; public health; bibliometrics; scholarly communication.

1 INTRODUÇÃO

A Ciência da Informação oferece uma chave importante para compreender a comunicação científica porque desloca a atenção do conteúdo publicado para as condições que tornam esse conteúdo registrável, validável e reconhecível. Nesse sentido, a ciência não se organiza apenas pelo acúmulo de descobertas, mas por sistemas que definem como os resultados circulam e quais produções adquirem legitimidade. Silva, Avelino, Souza e Vogel (2025) destacam que os canais formais de publicação dependem do rigor editorial e da avaliação por pares para produzir credibilidade. No entanto, essa mesma estrutura também pode estabelecer fronteiras de aceitação que dificultam a entrada de temas emergentes, instituições periféricas ou perspectivas menos estabilizadas no campo científico. Assim, analisar a comunicação científica exige observar não apenas o que é publicado, mas também os mecanismos que regulam sua visibilidade.

Esse problema se torna mais evidente com a expansão das bases abertas e das infraestruturas digitais de indexação. Embora o acesso aberto tenha ampliado a disponibilidade da produção científica registrada, sua existência não garante, por si só, uma circulação mais equilibrada do conhecimento (Galvino; Rosa; Oliveira, 2020). A abertura depende das formas pelas quais os registros são descritos, classificados e recuperados nas bases que os organizam. Por isso, as desigualdades não desaparecem quando a informação se torna acessível; elas podem permanecer inscritas nos próprios sistemas que distribuem relevância e visibilidade. É nesse ponto que a Ciência Aberta precisa ser examinada

criticamente: mais do que ampliar o acesso, importa compreender quem se torna visível, em quais condições e por meio de quais mediações informacionais.

No campo da Saúde Coletiva, a discussão sobre comunicação científica assume uma dimensão particular, porque o conhecimento produzido nessa área interfere diretamente na forma como problemas sanitários são compreendidos e enfrentados. Luz (2009) mostra que a Saúde Coletiva não se organiza a partir de uma única matriz disciplinar, mas pela articulação entre diferentes modos de interpretar o processo saúde-doença. Essa característica impede que sua produção científica seja lida apenas por critérios biomédicos ou epidemiológicos. Para este estudo, esse ponto é decisivo: compreender a literatura da área exige observar também como seus temas são registrados, classificados e tornados visíveis nas infraestruturas de informação científica.

É nesse ponto que a *OpenAlex* se torna relevante para a análise. A base reúne metadados de publicações, autores, instituições e temas de pesquisa, permitindo observar relações que dificilmente seriam percebidas em uma leitura isolada dos artigos (Priem; Piwowar; Orr, 2022). No entanto, essa possibilidade analítica não deve ser tomada como uma operação neutra. Como indicam Van Dijck, Poell e De Waal (2018) e Couldry e Mejias (2019), plataformas digitais também organizam hierarquias, pois definem quais registros são recuperados, como são classificados e quais relações se tornam mais visíveis. Assim, o uso da *OpenAlex* neste estudo não é apenas uma escolha metodológica; é também uma forma de examinar como a própria infraestrutura participa da construção da visibilidade científica.

A questão central deste estudo nasce da distância entre a promessa da Ciência Aberta e o modo como a produção científica efetivamente se torna visível. A abertura amplia o acesso aos registros, mas não elimina as desigualdades que atravessam a comunicação científica. Ao contrário, em certas condições, as próprias bases digitais podem reafirmar a centralidade de instituições já consolidadas e tornar menos perceptíveis agendas que circulam fora dos núcleos dominantes. Por isso, a produção brasileira em Saúde Pública e Saúde Coletiva precisa ser examinada não apenas pelo volume de publicações, mas pelas formas como é registrada e organizada na *OpenAlex*.

Diante desse problema, formula-se a seguinte questão de pesquisa: de que maneira a organização temática e as redes institucionais configuram a produção científica brasileira em Saúde Pública e Saúde Coletiva indexada na *OpenAlex*? Para respondê-la, este estudo tem

como objetivo analisar a produção nacional em acesso aberto nesses campos, no período de 2015 a 2025, considerando duas dimensões articuladas. A primeira, baseada na coocorrência de termos, permite observar como os temas se aproximam, quais vocabulários adquirem maior centralidade e que matrizes de conhecimento estruturam o corpus. A segunda, baseada na coautoria institucional, permite examinar como a produção circula entre instituições, quais centros assumem posições mais conectadas e como se configuram as relações de colaboração. Essas duas dimensões foram adotadas porque, em conjunto, permitem compreender a produção científica tanto pelo modo como é tematicamente organizada quanto pelas redes que sustentam sua circulação.

2 REFERENCIAL TEÓRICO

A Saúde Coletiva brasileira constitui um campo científico historicamente institucionalizado, cuja formação não pode ser confundida com a simples continuidade da Saúde Pública tradicional. Barata (2022) mostra que sua emergência se vincula a movimentos anteriores, como a medicina preventiva e a medicina social, mas também se diferencia deles ao afirmar a saúde como objeto coletivo, social e politicamente situado. Nunes (1994) contribui para essa compreensão ao organizar a constituição do campo em momentos históricos sucessivos, nos quais o projeto preventivista, a medicina social e a consolidação da Saúde Coletiva expressam diferentes modos de compreender a relação entre saúde, sociedade e prática científica. Essa passagem entre projetos distintos não ocorreu apenas no plano conceitual, pois envolveu também a formação de uma comunidade acadêmica, a criação de programas de pós-graduação e a consolidação de espaços próprios de circulação científica.

Nesse processo, Barata (2022) destaca o papel da Associação Brasileira de Saúde Coletiva (Abrasco), dos congressos da área, dos periódicos especializados e dos programas de pós-graduação na consolidação de uma comunidade científica com critérios próprios de reconhecimento. Mendonça e Carvalho (2025) reforçam essa leitura ao mostrarem que a Abrasco passou a reunir pesquisadores, docentes, estudantes e profissionais em torno da produção de conhecimento voltada ao fortalecimento do SUS e ao diálogo com a sociedade.

A partir dessa constituição histórica, a Saúde Coletiva passa a se distinguir menos pela criação de um novo rótulo disciplinar do que pela mudança no modo de compreender o processo saúde-doença. Barata (2022) mostra que o campo emerge de movimentos anteriores, mas se afirma ao tensionar os limites da medicina preventiva e da Saúde Pública

tradicional, especialmente quando desloca a saúde da esfera estritamente médica para o plano das relações sociais. Essa inflexão também é destacada por Mendonça e Carvalho (2025), ao associarem a formação da Saúde Coletiva à crítica do paradigma biomédico, cuja centralidade na clínica, no hospital e na prática curativa individual se mostrou insuficiente para responder aos problemas sanitários vividos pela população brasileira.

Nesse ponto, Laurell (1982) oferece uma formulação decisiva ao compreender a saúde e a doença da coletividade como expressão do processo social, e não como eventos isolados do corpo biológico. Nesse sentido, a crítica ao paradigma biomédico não opera apenas como revisão interna da área da saúde, mas como fundamento para a constituição de um campo científico próprio. Ao deslocar o processo saúde-doença para o plano das relações sociais, a Saúde Coletiva passa a disputar quais problemas devem ser reconhecidos como objetos legítimos de conhecimento e quais formas de intervenção podem ser sustentadas pela produção científica.

A consolidação de um campo científico não se realiza apenas pela elaboração de conceitos ou pela existência de instituições próprias; ela também depende das formas pelas quais o conhecimento produzido passa a circular publicamente. É nesse sentido que Witter (1997) compreende a produção científica como o modo pelo qual universidades e instituições de pesquisa se fazem presentes no “saber-fazer-poder ciência”. A formulação é relevante porque desloca a publicação do plano meramente documental para o plano da legitimação: produzir ciência implica inscrever resultados em espaços reconhecidos, nos quais o conhecimento possa ser comunicado, avaliado e incorporado à memória do campo. Kunsch (2015) contribui para essa leitura ao lembrar que os resultados da atividade científica se socializam em congressos, periódicos e outros formatos de publicação, isto é, por meio de canais que tornam o conhecimento partilhável e reconhecível por uma comunidade. Mugnaini (2006), por sua vez, chama a atenção para a qualidade das informações presentes na produção bibliográfica e para os mecanismos de validação que sustentam sua credibilidade.

Quando a produção científica passa a circular em ambientes digitais, sua presença pública deixa de depender apenas da publicação em si e passa a depender das estruturas que a descrevem. Modesto e Ramalho (2024) mostram que a comunicação científica envolve a divulgação, a avaliação e a validação dos resultados de pesquisa no interior de uma comunidade especializada. No entanto, com a expansão dos sistemas digitais, essa

comunicação passa a exigir formas mais precisas de tratamento da informação, pois os registros precisam ser localizados, relacionados e reutilizados em diferentes bases. É nesse ponto que os metadados assumem função central. Para os autores, o controle e a organização da produção científica impõem a adoção de padrões capazes de torná-la acessível e utilizável.

Essa função também aparece em Baptista e Ferneda (2016), para quem os metadados reduzem dificuldades de localização e recuperação da informação. Desse modo, os metadados não são elementos acessórios da comunicação científica digital; eles constituem a camada que permite transformar publicações dispersas em registros identificáveis, recuperáveis e passíveis de análise.

A partir dessa lógica de descrição estruturada, bases abertas como a *OpenAlex* passam a ocupar papel relevante na comunicação científica contemporânea. Macêdo, Schiessl e Shintaku (2025) caracterizam a plataforma como uma alternativa gratuita para estudos bibliométricos, capaz de complementar bases tradicionais ao ampliar o acesso a métricas de publicação e citação. Essa abertura, contudo, não se resume à gratuidade da ferramenta. Priem, Piwowar e Orr (2022) assinalam que a *OpenAlex* foi desenvolvido após a descontinuidade do *Microsoft Academic Graph* e se apresenta como uma infraestrutura voltada à transparência, à descoberta e à representação da pesquisa científica. Seu funcionamento se apoia na organização de entidades informacionais, como publicações, autores, instituições e temas, permitindo que a produção científica seja observada por meio das relações estabelecidas entre esses elementos.

Quando a produção científica passa a ser organizada em bases de metadados, torna-se possível analisá-la por meio de procedimentos bibliométricos e de redes. Bufrem e Prates (2005) indicam que a mensuração da informação científica depende do modo como o saber registrado é organizado e tornado passível de análise. Nessa perspectiva, os indicadores não se limitam a quantificar publicações; eles permitem observar formas de concentração, aproximação temática e circulação institucional. Van Eck e Waltman (2010), ao tratarem dos mapas bibliométricos, mostram que a visualização de redes pode revelar relações que permanecem pouco perceptíveis em leituras isoladas da produção científica. No caso de bases como a *OpenAlex*, essa possibilidade se amplia porque os registros são estruturados por entidades informacionais que relacionam trabalhos, autores, instituições e temas. Assim, a coocorrência de termos e a coautoria institucional podem ser compreendidas como duas

formas complementares de interpretar a configuração de um campo científico.

3 PROCEDIMENTOS METODOLÓGICOS

Esta pesquisa tem natureza quantitativa e exploratória, com apoio de procedimentos cientométricos e de análise de redes sociais. A opção por esse desenho metodológico decorre do próprio objetivo do estudo: mapear um volume amplo de registros científicos e observar, a partir deles, como se organizam as relações temáticas e institucionais da produção brasileira em Saúde Pública e Saúde Coletiva em uma base aberta de informação científica.

A fonte escolhida para a coleta foi a *OpenAlex*, uma infraestrutura aberta que reúne metadados sobre obras, autores, instituições, periódicos, conceitos e relações de colaboração. A escolha dessa base se justifica por sua ampla cobertura, pela disponibilidade de metadados relacionais e pela possibilidade de extração em larga escala. Esses aspectos favorecem análises voltadas à comunicação científica, especialmente quando o interesse recai sobre redes de temas, instituições e vínculos de colaboração.

O corpus foi definido a partir de quatro critérios de inclusão. Foram considerados apenas artigos científicos, publicados em acesso aberto, no período de 2015 a 2025, com pelo menos uma afiliação institucional vinculada ao Brasil nos metadados do registro. A participação brasileira foi, portanto, definida pela afiliação informada na própria base, e não pelo idioma do texto, pela nacionalidade dos autores ou pelo país de publicação do periódico. Com isso, puderam ser incluídos trabalhos com vínculos institucionais brasileiros publicados tanto em periódicos nacionais quanto internacionais.

A extração dos metadados ocorreu em 14 de fevereiro de 2026, por meio da API (*Application Programming Interface*) da *OpenAlex*. A busca foi aplicada aos campos de título e resumo (*title/abstract search*), com descritores em português e inglês relacionados à Saúde Pública e à Saúde Coletiva. A expressão utilizada foi: (“saúde pública” OR “saude publica” OR “saúde coletiva” OR “saude coletiva” OR “*public health*” OR “*collective health*”).

Depois da recuperação inicial, foram aplicados os filtros correspondentes ao recorte temporal, ao tipo documental, ao acesso aberto e à presença de afiliação institucional brasileira. Também foram realizados procedimentos de depuração para reduzir duplicidades e inconsistências nos registros. Ao final dessa etapa, o corpus analítico ficou composto por 28.260 artigos científicos, que constituíram a base empírica da investigação.

Os metadados recuperados passaram, então, por organização e tratamento. Essa etapa

envolveu a higienização dos dados, a conferência de registros inconsistentes e a harmonização de nomes de autores, instituições e termos. Buscou-se reduzir a dispersão provocada por siglas, acrônimos, traduções, variações ortográficas e diferentes formas de registro institucional. Para isso, foi elaborado um arquivo de tesauro customizado no *software VOSviewer*, utilizado para padronizar variantes linguísticas e reunir entradas equivalentes sob uma mesma forma de representação.

A análise foi realizada no *VOSviewer* em duas frentes. A primeira consistiu na análise de coocorrência de termos (*co-word analysis*), voltada à identificação dos principais domínios conceituais, aproximações semânticas e clusters temáticos presentes no corpus. Essa etapa permitiu observar quais temas aparecem com maior recorrência e como se relacionam no interior da produção analisada. A segunda frente correspondeu à análise de coautoria institucional (*co-authorship analysis*), orientada ao mapeamento das redes de colaboração entre instituições. Com essa etapa, foi possível identificar instituições com maior centralidade relativa, agrupamentos colaborativos e fluxos de cooperação científica. Assim, a análise articulou duas dimensões complementares: a organização temática da produção e sua configuração institucional.

Para evitar a saturação visual dos grafos, foram definidos limiares mínimos de ocorrência e de força de ligação (*link strength*). Essa escolha permitiu reduzir associações pouco expressivas e destacar os vínculos temáticos e institucionais mais consistentes do corpus. A leitura dos grafos considerou a distribuição espacial dos nós, a densidade das conexões, a espessura das arestas, a formação de clusters e a centralidade relativa dos termos e das instituições representadas.

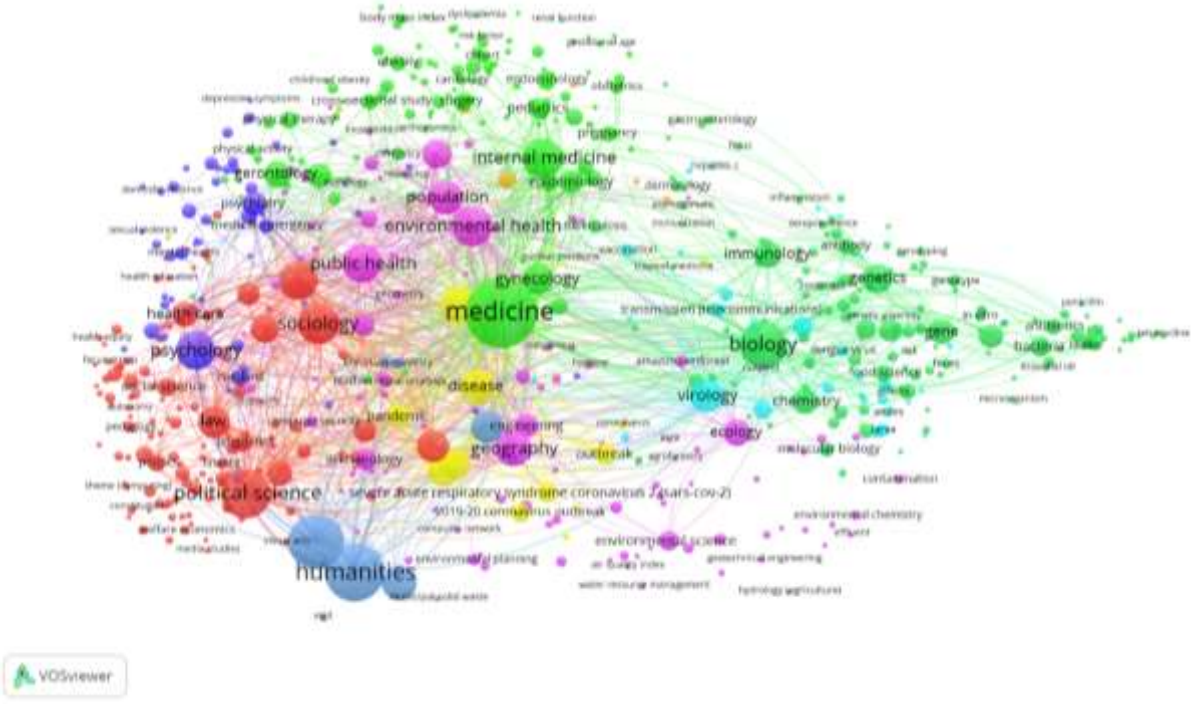
Cabe reconhecer, contudo, que a estratégia de busca adotada possui limites. Ao recuperar registros a partir dos termos “saúde pública”, “saúde coletiva”, “*public health*” e “*collective health*” nos campos de título e resumo, o estudo privilegia trabalhos que nomeiam explicitamente essas áreas nos metadados. Assim, pesquisas relacionadas ao Sistema Único de Saúde, à vigilância epidemiológica, à atenção básica, à promoção da saúde, às políticas públicas de saúde ou ao cuidado coletivo podem não ter sido recuperadas quando não empregaram esses descritores. Portanto, o corpus analisado não corresponde à totalidade da produção brasileira sobre Saúde Pública e Saúde Coletiva, mas ao conjunto de registros recuperáveis na *OpenAlex* a partir dos critérios definidos.

Também é necessário considerar que os resultados dependem da qualidade, completude, atualização, desambiguação institucional e consistência taxonômica dos metadados disponíveis na *OpenAlex* na data da coleta. Por essa razão, os grafos e indicadores produzidos devem ser lidos como representações situadas da produção indexada, e não como retratos absolutos do campo. Essa limitação é própria de estudos métricos baseados em grandes bases abertas e orienta a interpretação crítica dos resultados.

4 RESULTADOS E DISCUSSÃO

A análise dos 28.260 artigos científicos em acesso aberto sobre Saúde Pública e Saúde Coletiva, indexados na *OpenAlex*, permitiu observar duas dimensões principais da produção: a organização temática do corpus e as redes de colaboração institucional. A primeira dimensão foi examinada por meio da análise de coocorrência de termos (*co-word analysis*), com o objetivo de identificar aproximações conceituais e agrupamentos temáticos. A segunda voltou-se à análise de coautoria institucional (*co-authorship analysis*), permitindo observar os fluxos de colaboração e a centralidade relativa das instituições envolvidas.

Figura 1 - Rede de coocorrência de conceitos gerada no VOSviewer



Fonte: elaboração própria (2026).

A Figura 1 mostra que a produção brasileira em Saúde Pública e Saúde Coletiva

indexada na *OpenAlex* não se organiza em torno de um único núcleo temático. A rede apresenta uma composição heterogênea, na qual termos de diferentes tradições científicas aparecem em relação, ora aproximando dimensões clínicas e epidemiológicas, ora articulando problemas sociais, ambientais e institucionais da saúde. Esse resultado é coerente com a leitura de Luz (2009), para quem a Saúde Coletiva se constitui em um campo marcado pela convivência de diferentes racionalidades científicas. Também dialoga com Paim e Almeida Filho (1998), ao indicar que a área não pode ser compreendida como simples extensão da Saúde Pública tradicional ou como aplicação direta das ciências biomédicas. Desse modo, a rede não apenas distribui termos: ela torna visível a própria composição plural do domínio analisado.

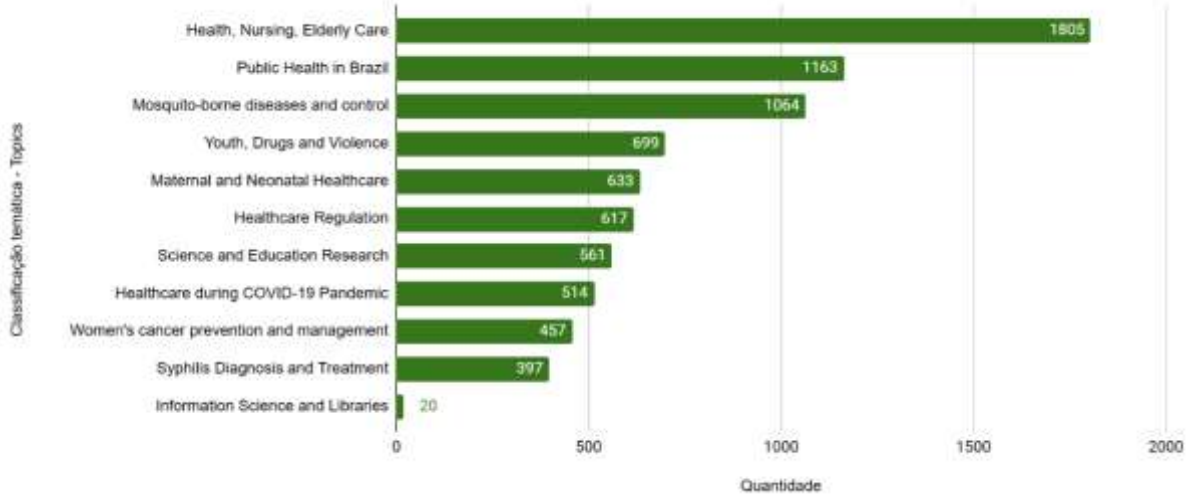
Essa pluralidade, contudo, não significa equilíbrio entre as matrizes que compõem o campo. A centralidade de medicine indica que a matriz biomédica permanece como referência dominante na representação da produção analisada. Esse dado é importante porque revela uma diferença entre a complexidade epistemológica da Saúde Coletiva e o modo como essa complexidade aparece nos metadados de uma infraestrutura global. Embora o campo incorpore debates sociais e políticos, as categorias clínicas tendem a ocupar posições mais estáveis e visíveis nos sistemas de classificação. Assim, o grafo sugere que a abertura da base não elimina as hierarquias já incorporadas aos vocabulários científicos; ao contrário, pode torná-las mais perceptíveis.

A presença de termos como *sociology*, *political science* e *humanities* tensiona essa centralidade ao indicar que a produção também se organiza a partir de leituras sociais da saúde. Esses termos mostram que o corpus não está restrito à doença como ocorrência biológica, mas incorpora discussões sobre instituições, desigualdades, políticas públicas e condições coletivas de vida. No entanto, sua posição na rede sugere que essas abordagens disputam espaço com categorias mais consolidadas nos sistemas internacionais de indexação. Para a Ciência da Informação, esse ponto é decisivo: a forma como os termos aparecem, se aproximam ou permanecem periféricos revela como os metadados participam da construção da visibilidade científica.

As zonas intermediárias da rede, nas quais aparecem conceitos como *public health* e *environmental health*, permitem observar os pontos de passagem entre essas matrizes. Esses termos funcionam como conectores entre abordagens clínicas, sociais e ambientais,

mostrando que a organização temática do corpus depende de relações semânticas produzidas pela própria infraestrutura de dados. Nesse sentido, a Figura 1 confirma a utilidade da análise de coocorrência, mas também evidencia seus limites: ela permite visualizar aproximações entre termos, sem explicar sozinha as condições históricas e institucionais que produziram tais aproximações. Cabe à Ciência da Informação justamente interpretar essa mediação, compreendendo que descrição, classificação e recuperação não são operações neutras, mas formas de organizar a presença de um domínio científico em uma base aberta.

Gráfico 1 - Dez principais classificações temáticas da produção científica em Saúde Pública e Saúde Coletiva



Fonte: elaboração própria (2026).

O Gráfico 1 aprofunda a leitura da rede ao indicar quais classificações temáticas concentram maior número de registros no corpus. A categoria *Health*, *Nursing*, *Elderly Care*, com 1.805 artigos, ocupa posição de destaque e confirma o peso das abordagens clínicas e epidemiológicas na produção analisada. Esse resultado não surpreende, considerando a centralidade histórica dos problemas assistenciais e sanitários na organização da pesquisa em saúde no Brasil. No entanto, ele também revela algo importante do ponto de vista informacional: os vocabulários vinculados à matriz biomédica parecem encontrar maior estabilidade nos sistemas de classificação científica.

Essa concentração precisa ser lida com cautela. Como indicam Van Dijck, Poell e De Waal (2018), plataformas digitais não apenas armazenam registros, mas interferem na forma como eles se tornam visíveis. No caso da *OpenAlex*, as categorias atribuídas aos artigos ajudam a organizar o campo, mas também produzem uma determinada imagem dele. Assim, quando

as classificações clínicas aparecem com maior força, não se observa apenas uma característica interna da Saúde Pública e da Saúde Coletiva; observa-se também o modo como essa produção é enquadrada por uma infraestrutura global de metadados.

A presença reduzida da categoria *Information Science and Libraries*, com 20 registros, deve ser interpretada nesse mesmo movimento. O dado não autoriza concluir que a Ciência da Informação seja irrelevante para o tema, mas mostra que sua presença aparece pouco explicitada na classificação temática do *corpus*. Essa diferença é importante: uma área pode não ocupar posição quantitativamente expressiva em determinada taxonomia e, ainda assim, oferecer instrumentos decisivos para interpretar como a produção científica é registrada, indexada e recuperada.

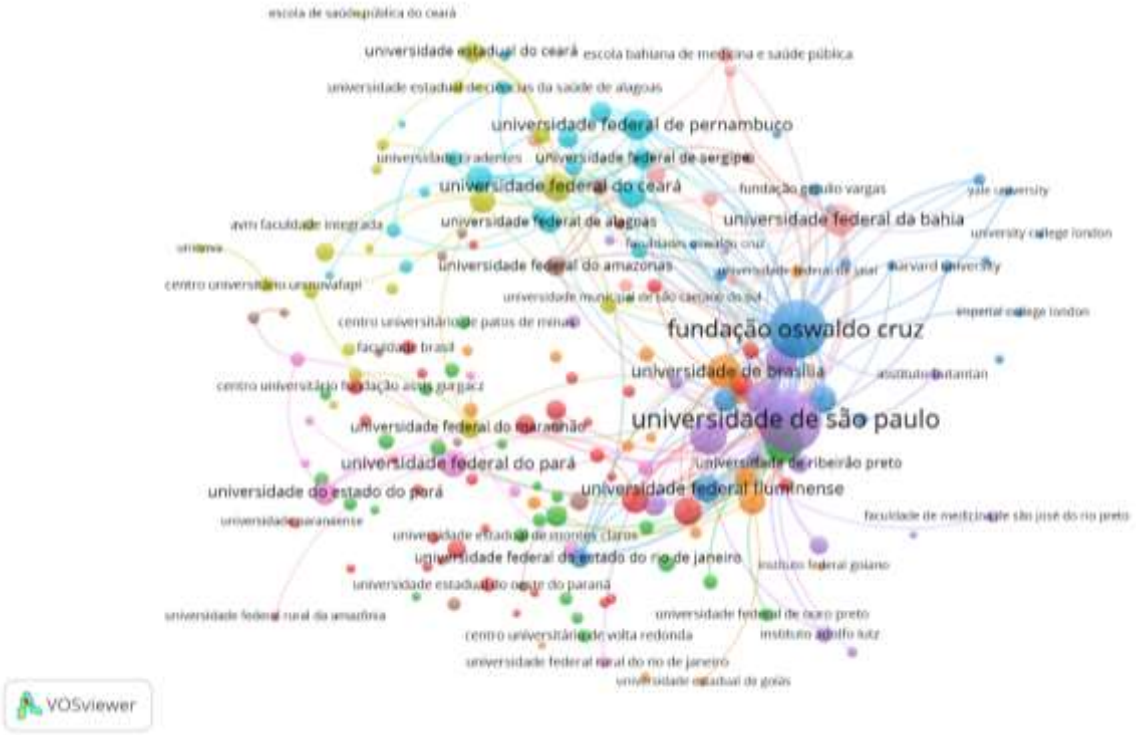
Desse modo, o Gráfico 1 reforça a hipótese de que a abertura dos dados não resolve, por si só, os problemas de representação da ciência. Suber (2012) e Albagli (2015) ajudam a compreender que o acesso aberto só ganha força quando os registros podem ser efetivamente organizados e apropriados. No caso analisado, a questão não é apenas saber quais áreas aparecem mais, mas compreender quais condições tornam certos temas mais classificáveis, mais recuperáveis e mais reconhecíveis do que outros. É nesse ponto que a leitura bibliométrica se aproxima da Ciência da Informação: os números deixam de ser apenas indicadores de volume e passam a revelar efeitos de organização, seleção e visibilidade.

A rede de coautoria mostra que a circulação científica em Saúde Pública e Saúde Coletiva se estrutura em torno de instituições já consolidadas na pesquisa em saúde, como Fundação Oswaldo Cruz (Fiocruz), Universidade de São Paulo (USP), Universidade Federal do Rio de Janeiro (UFRJ), Universidade Federal da Bahia (UFBA) e Universidade Federal de Pernambuco (UFPE). Essa centralidade não decorre apenas do volume de artigos vinculados a essas instituições, mas de sua capacidade de estabelecer conexões com diferentes grupos e centros de pesquisa. No grafo, portanto, a posição ocupada por esses nós indica uma forma de autoridade relacional: são instituições que não apenas produzem, mas também organizam parte importante dos fluxos de colaboração no corpus analisado.

Esse resultado pode ser interpretado à luz do Efeito Mateus, formulado por Merton (1968), segundo o qual o prestígio científico tende a produzir novas vantagens para agentes e instituições já reconhecidos. No caso analisado, a centralidade da Fiocruz e de universidades historicamente fortes na pesquisa em saúde sugere a permanência desse mecanismo de

acumulação. Instituições com maior tradição, infraestrutura e inserção em redes colaborativas tendem a ampliar sua própria visibilidade, enquanto centros com menor capital científico permanecem em posições menos articuladas. O dado, assim, não deve ser lido apenas como indicador de produtividade, mas como sinal de uma distribuição desigual das condições de circulação do conhecimento.

Figura 2 - Rede de coautoria entre instituições gerada no VOSviewer



Fonte: elaboração própria (2026).

A Ciência Aberta torna essa dinâmica mais observável, mas não a corrige automaticamente. Bases como a *OpenAlex* permitem visualizar redes extensas de colaboração e acompanhar a circulação institucional da produção científica. Entretanto, a existência de uma infraestrutura aberta não redistribui, por si só, financiamento, prestígio, capacidade técnica ou inserção internacional. O que a abertura oferece, nesse caso, é uma superfície de observação: ela permite enxergar com mais nitidez as hierarquias que já estruturam a ciência brasileira.

A relação entre a rede temática e a rede institucional reforça que a Saúde Pública e a Saúde Coletiva devem ser compreendidas como domínios informacionais atravessados por disputas de representação e circulação. Os temas que aparecem com maior força no corpus

não estão separados das instituições que os produzem, financiam, publicam e fazem circular. Para a Ciência da Informação, essa articulação é central, porque mostra que indicadores bibliométricos não são apenas medidas de desempenho científico. Eles também expressam formas de organização da visibilidade, condicionadas por metadados, critérios de indexação e padrões de colaboração.

Desse modo, a Figura 2 amplia a interpretação dos resultados ao mostrar que a abertura informacional convive com concentração institucional. A *OpenAlex* permite mapear a produção brasileira em Saúde Pública e Saúde Coletiva em larga escala, mas a rede de coautoria revela que essa produção continua ancorada em centros com maior capacidade histórica de articulação científica. A contribuição da análise está justamente em evidenciar essa ambivalência: a Ciência Aberta amplia o acesso aos registros, mas a visibilidade e a circulação do conhecimento continuam dependentes das condições institucionais que organizam o campo científico.

5 CONSIDERAÇÕES FINAIS

Esta investigação analisou a produção brasileira em acesso aberto sobre Saúde Pública e Saúde Coletiva indexada na *OpenAlex*, no período de 2015 a 2025. Os resultados indicaram que esse domínio não se organiza de maneira homogênea. A rede de coocorrência evidenciou o peso de vocabulários clínicos e epidemiológicos, mas também revelou abordagens sociais e ambientais da saúde, confirmando a pluralidade da Saúde Coletiva quando observada por meio de seus metadados. A rede de coautoria revelou uma configuração institucional concentrada. Fiocruz, USP, UFRJ, UFBA e UFPE ocupam posições centrais, indicando que a circulação científica permanece associada a centros historicamente consolidados na pesquisa em saúde. Esse achado não expressa apenas produtividade, mas uma estrutura desigual de colaboração científica. A Ciência Aberta amplia o acesso aos registros, mas não redistribui automaticamente as condições institucionais que sustentam a produção do conhecimento.

A principal contribuição do estudo para a Ciência da Informação está em tratar a *OpenAlex* não apenas como fonte de dados, mas como infraestrutura que participa da construção da visibilidade científica. Ao organizar registros, atribuir categorias e relacionar instituições, a base interfere nos modos pelos quais a produção científica pode ser recuperada e interpretada. Desse modo, a análise bibliométrica e a análise de redes ultrapassam uma função meramente descritiva, pois permitem examinar como a comunicação científica é

mediada por sistemas de metadados.

A abertura torna os registros mais acessíveis, mas sua efetividade depende das condições que permitem organizar e interpretar esses registros. Quando tais condições são desiguais, a disponibilidade da informação não garante visibilidade equivalente entre áreas, temas e instituições. Assim, a democratização do conhecimento científico não decorre apenas da abertura das bases, mas também dos modos pelos quais a informação é descrita, recuperada e legitimada. Como limitação, reconhece-se que o estudo se concentrou em uma única base de dados e trabalhou com os metadados disponíveis na *OpenAlex* na data da coleta. Além disso, a expressão de busca adotada privilegiou registros que mencionavam explicitamente os termos "saúde pública", "saúde coletiva", "*public health*" ou "*collective health*" nos campos de título e resumo. Com isso, trabalhos relacionados ao Sistema Único de Saúde, à vigilância epidemiológica, à atenção básica ou a outras práticas de cuidado coletivo podem não ter sido recuperados quando não utilizaram esses descritores.

Para pesquisas futuras, recomenda-se comparar os resultados com outras bases de dados e refinar a estratégia de busca por meio de descritores complementares. Também se mostra pertinente aprofundar a análise das desigualdades regionais e das condições de representação da produção científica em saúde. Esses desdobramentos podem contribuir para compreender com maior precisão como a Saúde Pública e a Saúde Coletiva brasileiras são organizadas, recuperadas e tornadas visíveis em diferentes infraestruturas de comunicação científica.

REFERÊNCIAS

ALBAGLI, Sarita. Ciência aberta em questão. *In*: ALBAGLI, Sarita; MACIEL, Maria Lucia; ABDO, Alexandre Hannud (org.). **Ciência aberta, questões abertas**. Brasília: Ibict; Rio de Janeiro: Unirio, 2015. p. 9-25. Disponível em: https://livroaberto.ibict.br/handle/1/1060. Acesso em: 14 mar. 2026.

BAPTISTA, Fernando; FERNEDA, Edberto. O padrão brasileiro de metadados para teses e dissertações MTD-BR: aspectos e contribuições. *In*: SEMINÁRIO EM CIÊNCIA DA INFORMAÇÃO, 6., 2016, Londrina**. Anais** [...]. Londrina: UEL, 2016. Disponível em: https://www.uel.br/eventos/cinf/index.php/secin2016/secin2016/paper/viewFile/323/202. Acesso em: 10 set. 2026.

BARATA, Rita Barradas. O campo científico da saúde coletiva. **Saúde em Debate**, Rio de Janeiro, v. 46, n. 133, p. 473-486, abr./jun. 2022. Disponível em: https://doi.org/10.1590/0103-1104202213316. Acesso em: 07 set. 2026.

BUFREM, Leilah Santiago; PRATES, Yara. O saber científico registrado e as práticas de mensuração da informação. **Ciência da Informação**, Brasília, DF, v. 34, n. 2, p. 9-25, maio/ago. 2005. Disponível em: https://doi.org/10.18225/ci.inf.v34i2.1087. Acesso em: 3 maio 2026.

COULDRY, Nick; MEJIAS, Ulises Ali. **The costs of connection:** how data is colonizing human life and appropriating it for capitalism**.** Stanford: Stanford University Press, 2019.

GALVINO, Cláudio Cesar Temóteo; ROSA, Maria Nilza Barbosa; OLIVEIRA, Bernardina Maria Juvenal Freire de. O movimento de Acesso Aberto e a Ciência Aberta: uma proposta de repositório de dados e memória na Universidade Federal de Alagoas. **Ciência da Informação em Revista**, Maceió, v. 7, n. 1, p. 34-45, jan./abr. 2020. Disponível em: http://dx.doi.org/10.28998/cirev.2020v7n1c. Acesso em: 14 mar. 2026.

KUNSCH, Margarida Maria Krohling. Os campos acadêmicos em Comunicação Organizacional e Relações Públicas no Brasil: caracterização, pesquisa científica e tendências. **Revista Internacional de Relaciones Públicas**, Málaga, v. 5, n. 10, p. 105-124, 2015. Disponível em: http://dx.doi.org/10.5783/RIRP-10-2015-07-105-124. Acesso em: 10 set. 2026.

LAURELL, Asa Cristina. La salud enfermedad como proceso social. **Revista Latinoamericana de Salud**, [*s. l.*], v. 2, p. 7-25, 1982. Disponível em: https://digitalrepository.unm.edu/lasm_pt/180/. Acesso em: 6 set. 2026.

LUZ, Madel Therezinha. Complexidade do campo da Saúde Coletiva: multidisciplinaridade, interdisciplinaridade e transdisciplinaridade de saberes e práticas: análise sócio-histórica de uma trajetória paradigmática. **Saúde e Sociedade**, São Paulo, v. 18, n. 2, p. 304-311, 2009. Disponível em: https://doi.org/10.1590/S0104-12902009000200013. Acesso em: 17 mar. 2026.

MACÊDO, Diego José; SCHIESSL, Ingrid Torres; SHINTAKU, Milton. O uso do OpenAlex nos estudos de métricas bibliométricas. **Biblios**: Journal of Librarianship and Information Science, [*s. l.*], n. esp., e015, 2025. Disponível em: https://doi.org/10.5195/biblios.2025.1268. Acesso em: 17 ago. 2026.

MENDONÇA, Paula Beatriz de Souza; CARVALHO, Raquel Baroni de. Estruturação do campo da Saúde Coletiva: a história recente da formação na pós-graduação. **Revista Baiana de Saúde Pública**, Salvador, v. 49, n. 1, jan./mar. 2025. Disponível em: https://rbsp.sesab.ba.gov.br/index.php/rbsp/article/view/4218. Acesso em: 7 set. 2026.

MERTON, Robert King. The Matthew Effect in Science. **Science**, [*s. l.*], v. 159, n. 3810, p. 56-63, 1968. Disponível em: https://doi.org/10.1111/j.1365-4362.1988.tb02325.x. Acesso em: 20 out. 2026.

MODESTO, Fernando; RAMALHO, Amanda. Metadados na comunicação científica: mapeamento de modelos e aplicações. *In*: MUGNAINI, Rogério; MODESTO, Fernando (ed.). **Monitoramento, infraestruturas e tecnologias**: novos desafios na comunicação científica.

Salvador: EDUFBA, 2024. p. 265-294. Disponível em: https://books.scielo.org/id/25y65/14. Acesso em: 7 set. 2026.

MUGNAINI, Rogério. **Caminhos para adequação da avaliação da produção científica brasileira: impacto nacional versus internacional**. 2006. 254 f. Tese (Doutorado em Cultura e Informação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2006. Disponível em: https://teses.usp.br/teses/disponiveis/27/27151/tde-11052007-091052/pt-br.html. Acesso em: 10 set. 2026.

NUNES, Everardo Duarte. Saúde coletiva: história de uma idéia e de um conceito. **Saúde e Sociedade**, São Paulo, v. 3, n. 2, p. 5-21, dez. 1994. Disponível em: https://doi.org/10.1590/S0104-12901994000200002. Acesso em: 8 set. 2026.

PAIM, Jairnilson S.; ALMEIDA FILHO, Naomar de. Saúde coletiva: uma "nova saúde pública" ou campo aberto a novos paradigmas? **Revista de Saúde Pública**, São Paulo, v. 32, n. 4, p. 299-316, 1998. Disponível em: https://doi.org/10.1590/S0034-89101998000400001. Acesso em: 3 mar. 2026.

PRIEM, Jason; PIWOWAR, Heather; ORR, Richard. **OpenAlex:** a fully-open index of scholarly works, authors, venues, institutions, and concepts. 2022. Disponível em: https://arxiv.org/abs/2205.01833. Acesso em: 16 fev. 2026.

SILVA, Maria Cecília Brandão da; AVELINO, Maria Eduarda Marçal; SOUZA, Manuela Nonaka Vargas de; VOGEL, Michely Jabala Mamede. A noção de comunicação científica e divulgação científica na Ciência da Informação brasileira: questão de estrutura ou de linguagem. *In*: FEISC - FÓRUM DE ESTUDOS EM INFORMAÇÃO, SOCIEDADE E CIÊNCIA, 6., 2025, Porto Alegre. **Anais** [...] Porto Alegre: Faculdade de Biblioteconomia e Comunicação/UFRGS, 2025.

SUBER, Peter. **Open access**. Cambridge: MIT Press, 2012.

VAN DIJCK, José; POELL, Thomas; DE WAAL, Martijn. **The platform society**: public values in a connective world. Oxford: Oxford University Press, 2018.

VAN ECK, Nees Jan; WALTMAN, Ludo. Software survey: VOSviewer, a computer program for bibliometric mapping. **Scientometrics**, Budapest, v. 84, n. 2, p. 523-538, 2010. Disponível em: https://doi.org/10.1007/s11192-009-0146-3. Acesso em: 5 maio 2026.

WITTER, Geraldina Porto. **Produção científica**. Campinas: Átomo, 1997.