

Estado da publicação: O preprint não foi publicado em outro meio.

Letramento informacional e de dados com Wikipédia e Wikidata sobre distúrbios da comunicação humana

Hector Gabriel Corrale de Matos, Sarah Rodrigues Thanis Guarrido, Fernanda Zucki, Kátia de Freitas Alvarenga, Lilian Cássia Bórnia Jacob

<https://doi.org/10.1590/SciELOPreprints.17732>

Submetido em: 2026-08-30

Postado em: 2026-09-17 (versão 1)

(AAAA-MM-DD)

A moderação deste preprint recebeu o(s) endosso(s) de:

- Ana Paula Berberian (ORCID: <https://orcid.org/0000-0001-7176-7610>)

**Letramento informacional e de dados com Wikipédia e
Wikidata sobre distúrbios da comunicação humana**

**Information and data literacy with Wikipedia and Wikidata
on human communication disorders**

Hector Gabriel Corrale de Matos
Universidade de São Paulo, Faculdade de Odontologia de Bauru
ORCID: <https://orcid.org/0000-0002-2649-370X>

Sarah Rodrigues Thanis Guarrido
Universidade de São Paulo, Faculdade de Odontologia de Bauru
ORCID: <https://orcid.org/0000-0003-0789-9832>

Fernanda Zucki
Universidade Federal de Santa Catarina
ORCID: <https://orcid.org/0000-0002-5473-9762>

Kátia de Freitas Alvarenga
Universidade de São Paulo, Faculdade de Odontologia de Bauru
ORCID: <https://orcid.org/0000-0002-7847-3225>

Lilian Cássia Bórnica Jacob
Universidade de São Paulo, Faculdade de Odontologia de Bauru
ORCID: <https://orcid.org/0000-0003-1947-7506>

RESUMO. Objetivo: Este estudo examinou como atividades de edição na Wikipédia e no Wikidata podem integrar o letramento informacional e o letramento de dados na formação de profissionais da saúde sobre distúrbios da comunicação. Método: Estudo misto, descritivo e transversal. Foram analisados dados de participação de uma amostra de 251 estudantes de graduação e pós-graduação em ciências da saúde no Brasil em oito atividades de edição na Wikipédia e no Wikidata (quatro virtuais e quatro presenciais). Realizou-se uma análise qualitativa de 30 artefatos de edição pública produzidos pelos estudantes na Wikipédia e no Wikidata, fundamentada no Modelo para o Letramento Informacional no Ensino Superior (Framework for Information Literacy for Higher Education) e no modelo Letramento–Habilidade–Conhecimento (Literacy–Ability–Knowledge). Foram empregados testes estatísticos exploratórios (Exato de Fisher, Wilcoxon e regressão binomial negativa) para avaliar as relações entre o tipo de evento e os indicadores de participação. Resultados: Não houve diferença entre os formatos presencial e virtual. A análise qualitativa indicou que a edição na Wikipédia eliciou práticas de avaliação de evidências, de tradução de evidências em textos de comunicação em saúde e de atribuição de fontes de referência. Observou-se que a edição no Wikidata eliciou a modelagem e o raciocínio sobre as relações estruturais dos dados. Conclusão: Atividades de edição na Wikipédia e no Wikidata oferecem pontos de entrada para a integração do letramento informacional e do letramento de dados na educação em saúde sobre distúrbios da comunicação humana.

Descritores: Enciclopédias como Assunto, Conjuntos de Dados como Assunto, Literacia para a Saúde, Ensino Superior, Comunicação em Saúde

ABSTRACT. Purpose: This study examined how Wikipedia and Wikidata editing activities can embed information literacy and data literacy within the training of healthcare professionals regarding communication disorders. Method: A mixed-methods, descriptive, and cross-sectional study. Participation data were analyzed from a sample of 251 undergraduate and graduate health science students in Brazil across eight Wikipedia and Wikidata editing activities (four virtual and four in-person). A qualitative analysis of 30 public editing artifacts produced by the students on Wikipedia and Wikidata was conducted, grounded in the Framework for Information Literacy for Higher Education and the Literacy–Ability–Knowledge model. Statistical tests (Fisher's Exact Test, Wilcoxon test, and negative binomial regression) were employed to evaluate relationships between event type and participation indicators. Results: There was no difference between the in-person and virtual formats. Qualitative analysis indicated that Wikipedia editing elicited practices of evidence evaluation, translation of evidence into health communication texts, and attribution of reference sources. It was observed that editing Wikidata elicited modeling and reasoning about the data's structural relationships. Conclusion: Wikipedia and Wikidata editing activities offer entry points for embedding information literacy and data literacy within health education concerning communication disorders.

Keywords: Encyclopedias as Topic, Datasets as Topic, Health Literacy, Higher Education, Health Communication

INTRODUÇÃO

Os distúrbios da comunicação geram impactos quantificáveis na saúde pública mundial (1). A Organização Mundial da Saúde projeta que a marca de 2,5 bilhões de pessoas com perda auditiva será atingida até 2050 (2). Esse cenário exige estratégias validadas para a disseminação de informações científicas (3). A tradução do conhecimento em linguagem acessível e a promoção do letramento em saúde são apontadas como ações complementares à disseminação efetiva de evidências científicas (4,5).

Intervenções baseadas em evidências para o letramento em saúde, incluindo as mediadas por plataformas digitais, constituem formas de tradução do conhecimento (6). A Wikipédia e o Wikidata facilitam a difusão de informações sobre condições de saúde na internet em escala global (7,8). A Wikipédia é uma enciclopédia que organiza conteúdo textual livre em verbetes, nos quais os assuntos são inter-relacionados por meio de hiperligações internas (wikilinks). Cada verbete é estruturado em seções descritivas compostas por parágrafos, elementos multimídia, metadados (como categorias e infocaixas) e citações a fontes secundárias verificáveis (7). Wikidata é uma base de dados que organiza dados estruturados em redes de conhecimento, nas quais conceitos (itens) são inter-relacionados por meio de propriedades. Essa estrutura é legível por humanos e por máquinas. Cada item é descrito por afirmações compostas por propriedades, valores, qualificadores opcionais e fontes de referência (8).

Campanhas de saúde pública e atividades educacionais já utilizaram o alcance dessas plataformas para promover a saúde auditiva (10). O uso e a tradução de referências científicas em informações acessíveis consolidam a Wikipédia como um repositório de acesso aberto e uma fonte de informação verificável (7), enquanto o Wikidata organiza

dados científicos em redes de representação interoperáveis e abertas (8). A manutenção dessas plataformas é viabilizada pela colaboração de voluntários, em especial de estudantes, profissionais e pesquisadores (10).

Nesse contexto, o letramento informacional e o letramento de dados despontam como competências complementares para formação de profissionais da saúde (11,12,13). O modelo de análise Framework for Information Literacy for Higher Education (Marco de Referência para o Letramento Informacional no Ensino Superior) da Association of College and Research Libraries (ACRL) fundamenta teoricamente as práticas de letramento informacional (14). Outro modelo conceitual é o Literacy-Ability-Knowledge (LAK, Letramento-Habilidade-Conhecimento), uma estrutura que permite descrever competências disciplinares em três dimensões (letramento, habilidade e conhecimento) (15).

As pesquisas atuais frequentemente presumem o desenvolvimento automático de competências no uso de ferramentas digitais (7,16). A comunicação em saúde é considerada parte da formação de profissionais (16,17,18). Entretanto, a literatura carece de avaliações empíricas que diferenciem as exigências cognitivas da produção textual dos modelos de organização da informação sobre temas de saúde (13,19). Estudos sobre letramento de dados têm se concentrado prioritariamente em áreas como Ciência, Tecnologia, Engenharia e Matemática (STEM, Science, Technology, Engineering, and Mathematics), com menos investigações em outras disciplinas (13,15,19). Os dados disponíveis refletem um viés de produção científica concentrada em contextos distintos do brasileiro são provenientes, em sua maioria, de contextos internacionais (9,10). Não foram identificados. A literatura não apresenta estudos de mapeamento dessas competências no contexto de formação de profissionais em

distúrbios da comunicação humana (9,10). Nesse sentido, a Wikipédia, modelo de escrita colaborativa, e o Wikidata, sistema de estruturação de dados, podem servir como instrumentos de análise de competências em letramento informacional e em dados (7,8)

Postula-se que a curadoria estruturada em atividades de edição oportuniza práticas distintas e complementares entre estudantes de saúde: a Wikipédia atua como mediadora do letramento informacional, focado na avaliação textual e na argumentação, enquanto o Wikidata mobiliza o letramento de dados por meio da modelagem semântica, independentemente da modalidade ser presencial ou virtual.

Diante disso, este estudo examinou como as atividades de edição na Wikipédia e no Wikidata podem integrar o letramento informacional e o letramento de dados na formação de profissionais da saúde sobre distúrbios da comunicação.

MÉTODO

Delineamento do Estudo

Este estudo utilizou um método misto, descritivo e transversal. Foram analisados dados quantitativos de participação agregados por evento e uma amostra de artefatos de edição pública produzidos em oito atividades de edição (editatonas na Wikipédia e datatonas no Wikidata). As atividades foram realizadas no segundo semestre de 2025. O relato foi orientado pelas recomendações STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) para o componente observacional (métricas de participação agregadas por evento) e pelo guia TIDieR (Template for Intervention Description and Replication) para a descrição padronizada das atividades baseadas na edição da Wikipédia e do Wikidata.

Aspectos Éticos

Todos os dados utilizados são públicos e abertos e, assim, de acordo com a Resolução n.º 510/2016 do CNS, não foi necessário submeter este estudo à apreciação ética. Trata-se de métricas agregadas e de históricos de edição disponíveis sob pseudônimo nas plataformas (sob licenças CC BY-SA 4.0 na Wikipédia e CC0 no Wikidata). Para garantir o anonimato dos participantes do Programs & Events Dashboard, os artefatos de edição amostrados foram anonimizados, sem menção a nomes de usuário, títulos de verbetes ou identificadores do Wikidata.

Contexto e Casuística

As atividades foram realizadas em instituições de ensino superior no Sul do Brasil (n = 2), no Sudeste do Brasil (n = 4) e no Nordeste do Brasil (n = 2), selecionadas por amostragem de conveniência. A participação foi voluntária em cada atividade. Foram considerados participantes elegíveis estudantes de graduação e pós-graduação com, no máximo, 25 edições prévias autorrelatadas na Wikipédia ou no Wikidata. As atividades foram realizadas no contexto de uma iniciativa de difusão do conhecimento sobre distúrbios da comunicação humana na Wikipédia e no Wikidata no segundo semestre de 2025. A participação foi majoritariamente composta por estudantes de cursos de saúde (Fonoaudiologia, Medicina, Odontologia e Psicologia). Participantes de outras áreas (Engenharia e Ciências da Computação) vinculados a projetos transdisciplinares em saúde digital também participaram das atividades. Ao longo das oito atividades, participaram 251 estudantes (o número de participantes por atividade variou de 7 a 81). Quatro atividades foram realizadas de forma virtual, por videoconferência (com a ferramenta Jitsi Meet), e outras quatro, presencialmente, nas unidades de ensino das universidades participantes.

Procedimento das Atividades

Cada atividade seguiu um protocolo de orientação de uma hora abordando normas de edição na Wikipédia e no Wikidata, incluindo políticas editoriais para conteúdo biomédico (notoriedade, ponto de vista neutro e verificabilidade). A prática de citação foi ensinada utilizando identificadores persistentes, como DOIs ou identificadores do

PubMed, que permitem rastrear uma citação de forma inequívoca até sua fonte. Em seguida, os participantes dedicaram de duas a quatro horas à edição na Wikipédia e no Wikidata, em conteúdos sobre distúrbios da comunicação humana e temas de saúde relacionados. Foi oferecido suporte de pesquisadores em distúrbios da comunicação e de voluntários experientes nos projetos Wikimedia (editores com experiência entre 2 e 4 anos na Wikipédia e na Wikidata). Os materiais de treinamento foram adaptados do currículo da Wiki Education Foundation (<https://outreachdashboard.wmflabs.org/training>) e das páginas de orientação do Meta-Wiki e da Wikiversidade (projetos de organização da comunidade educacional e técnica da Wikipédia e do Wikidata). Antes de cada atividade, a equipe criou listas temáticas de verbetes na Wikipédia e de itens no Wikidata utilizando o Wikidata Query Builder (<https://query.wikidata.org/querybuilder/>), o PetScan (<https://petscan.wmcloud.org/>) e o PagePile (<https://pagepile.toolforge.org/>). Os participantes foram orientados a localizar evidências revisadas por pares nos bancos de dados PubMed/MEDLINE, Scopus, SciELO e BVS-PAHO (Biblioteca Virtual em Saúde/Organização Pan-Americana da Saúde) antes de iniciarem as edições.

Coleta de Dados

Os dados de participação de todas as oito atividades (identificadas neste estudo como W1 a W8) foram exportados do Wikimedia Programs & Events Dashboard (<https://outreachdashboard.wmflabs.org/>) em novembro de 2025, em arquivos CSV, e tabulados no LibreOffice Calc. Foram derivadas quatro medidas de nível de evento: (i) número de participantes registrados e treinados; (ii) número de editores ativos (participantes com pelo menos uma edição rastreada); (iii) número de edições na

Wikipédia e no Wikidata; e (iv) número de referências adicionadas. O Wikimedia Programs & Events Dashboard reporta métricas agregadas por evento, não vinculadas a dados individuais. Por essa razão, a análise quantitativa opera no nível do evento ($n = 8$).

Análise Estatística

A proporção de participantes treinados que se tornaram editores ativos foi comparada entre formatos por meio do teste exato de Fisher (razão de chances, OR) e do teste do qui-quadrado; a mediana, por evento, da razão edições/editor foi comparada entre formatos por meio do teste de Wilcoxon (Mann-Whitney U). Para modelar o total de edições por evento com ajuste pelo número de editores, utilizou-se uma regressão binomial negativa com offset. Todos os testes foram bicaudais ($\alpha = 0,05$), conduzidos no R v.4.3.1. Dado o pequeno número de eventos, os testes têm caráter estritamente exploratório.

Modelos de Análise de Qualitativa

Foram utilizados dois modelos para a análise qualitativa: (i) Modelo para o Letramento Informacional no Ensino Superior da ACRL (Framework for Information Literacy for Higher Education) (14) forneceu seis conceitos: Busca como Exploração Estratégica (Searching as Strategic Exploration), Autoridade é Construída e Contextual (Authority Is Constructed and Contextual), Criação da Informação como um Processo (Information Creation as a Process), Informação Tem Valor (Information Has Value), Pesquisa como

Investigação (Research as Inquiry) e Produção Acadêmica como Conversação (Scholarship as Conversation); (ii) modelo Letramento–Habilidade–Conhecimento (LAK, Literacy–Ability–Knowledge) (15), originalmente desenvolvido para avaliar competências na educação em engenharia. O modelo LAK oferece três níveis (letramento, habilidade e conhecimento), originalmente descritos como disposições profissionais, capacidade aplicada e domínio teórico no contexto da educação em engenharia (15). Esses três níveis foram redefinidos neste estudo para analisar como os participantes demonstraram competências de letramento informacional e de dados na edição da Wikipédia e do Wikidata, o que diverge da aplicação original do modelo. Essa redefinição é apresentada como uma adaptação exploratória (15).

Amostragem e Análise de Artefatos

Foi selecionada, por amostragem aleatória estratificada por plataforma, uma amostra de 30 artefatos de edição pública (15 da Wikipédia e 15 do Wikidata) a partir dos históricos públicos de edição; considerou-se um artefato cada conjunto de edições de um mesmo participante em uma página da Wikipédia ou em um item do Wikidata. Excluíram-se edições marcadas pelo próprio autor como edição menor (minor edit), recurso nativo da Wikipédia/Wikidata utilizado para sinalizar correções ortográficas, gramaticais ou de formatação sem alteração de conteúdo substancial. Esse critério assegurou que os artefatos analisados refletissem contribuições de conteúdo para análise.

A estratificação por plataforma foi adotada para garantir representação equivalente entre as duas plataformas. A codificação foi realizada por um único pesquisador e prosseguiu em três etapas: (1) leitura de cada artefato com atribuição de rótulos

indutivos de ação; (2) agrupamento em categorias de competência pelo LAK; (3) mapeamento em categorias conceituais da ACRL. Os 30 artefatos constituem um subconjunto das edições totais registradas nas oito atividades. As categorias descrevem práticas observáveis, não ganhos de aprendizagem medidos por meio de avaliações. A análise indica as práticas que as atividades propiciam, mas não constitui evidência direta do desenvolvimento individual de competências.

Disponibilidade de Dados

A relação de atividades extraída do Programs & Events Dashboard e os arquivos das análises estatísticas encontram-se depositados publicamente em um repositório no Open Science Framework (<https://doi.org/10.17605/OSF.IO/3DWHT>). Dados da análise de artefatos não estão disponíveis para garantir o anonimato dos participantes das atividades, sem menção a nomes de usuário, títulos de verbetes ou identificadores do Wikidata. Informações complementares podem ser disponibilizadas mediante solicitação.

RESULTADOS

A análise organiza-se em duas vertentes. A primeira (Tabela 1), de caráter quantitativo, descreve o volume de produção e o perfil de engajamento dos participantes segundo o formato da atividade. A segunda (Tabela 2), de caráter qualitativo, mapeia as práticas observadas nas contribuições dos estudantes a partir de artefatos de edição pública, interpretadas com base nos modelos ACRL e LAK.

A Tabela 1 apresenta os resultados das oito atividades ($n = 8$; 251 participantes). As atividades presenciais corresponderam a 91,2% (718/787) de todas as edições no Wikidata e a 57,3% (118/206) das edições na Wikipédia, embora representassem apenas 75,3% (189/251) do total de participantes. A mediana, por atividade, da razão edições/editor foi de 2,6 para atividades virtuais e de 5,5 para atividades presenciais, sem diferença estatisticamente significativa (Wilcoxon, $p = 0,68$).

Tabela 1. Dados de participação nas oito atividades de edição na Wikipédia e no Wikidata (W1–W8) por formato (virtual ou presencial), por tema principal relacionado e por métricas de participação. Taxa de participação (editores/participantes totais da atividade). Dados consolidados do Wikimedia Programs & Events Dashboard de novembro de 2025.

Atividade	Tema princial	Participantes	Editores	Taxa de participação	Wikidata (dados)	Wikipedia (texto)
W1	Saúde Coletiva	7	5	71,4%	61	0
W2	Saúde Auditiva	26	22	84,6%	8	45
W3	Saúde Coletiva	11	4	36,4%	0	4
W4	Motricidade Orofacial	18	14	77,8%	0	39
Total (formato virtual)		62	45	72,6%	69	88
W5	Fonoaudiologia/Saúde Digital	38	34	89,5%	115	0
W6	Fonoaudiologia/Saúde Digital	32	29	90,6%	107	63
W7	Fonoaudiologia/Saúde Digital	38	33	86,8%	0	55
W8	Difusão Científica	81	41	50,6%	496	0
Total (formato presencial)		189	137	72,5%	718	118
Total (virtual e presencial)		251	182	72,5%	787	206

A análise estatística não revelou associação significativa entre formato e proporção de editores ativos ($\chi^2(1, N = 251) = 0,0002, p = 0,988$; teste exato de Fisher, $p = 1,00$, OR = 1,00, IC 95% 0,51–2,05). O total de edições por atividade, ajustado pelo número de editores, não apresentou diferença significativa (IRR = 1,26, IC 95% 0,42–3,78, $p = 0,67$). Os participantes adicionaram 689 referências científicas ao longo das oito atividades.

A Tabela 2 resume as categorias de competência que emergiram da codificação dos 30 artefatos amostrados, descritivas das práticas observadas e não generalizáveis a todos os participantes.

Tabela 2. Mapeamento das competências dos participantes com base nos modelos LAK e ACRL. As categorias de competência emergiram da análise de 30 artefatos de contribuição dos participantes nas atividades (15 da Wikipédia, 15 do Wikidata).

Processos de Produção de Informação das Ações dos Estudantes	Competência analisada com aplicação do modelo LAK	Estrutura de Letramento Informacional Analisada com modelo ACRL
Wikipedia		
Busca e seleção de evidências científicas em bases de dados bibliográficas	Avaliação de Evidências: Localiza fontes revisadas por pares e avalia se são confiáveis e atualizadas.	- Pesquisa como Exploração Estratégica (Searching as Strategic Exploration) - Autoridade é Construída e Contextual (Authority Is Constructed and Contextual)
Síntese de achados técnicos em uma linguagem clara, neutra e verificável para o grande público	Tradução do Conhecimento: Composição de texto que explica evidências científicas, mantendo um tom neutro e uma escrita acessível.	- Criação da Informação como um Processo (Information Creation as a Process) - Informação Tem Valor (Information Has Value)
Inserção e formatação de referências e metadados de fontes verificáveis	Atribuição de Fontes: Uso de citações padronizadas e de identificadores persistentes que permitem aos leitores rastrear a origem das afirmações.	- Informação Tem Valor (Information Has Value) - Autoridade é Construída e Contextual (Authority Is Constructed and Contextual)
Identificação de lacunas, vieses ou linguagem estigmatizante em artigos existentes e proposição de revisões	Curadoria de Conteúdo: Revisa o conteúdo publicado e sinaliza formulações problemáticas, substitui termos estigmatizantes por linguagem clinicamente descritiva e aborda a cobertura irregular.	- Pesquisa como Investigação (Research as Inquiry) - Autoridade é Construída e Contextual (Authority Is Constructed and Contextual)
Discussão e negociação de edições com pares (páginas de discussão da Wikipédia, mentoria síncrona durante	Produção Colaborativa: Justifica as escolhas de fontes, negocia a redação de conteúdos, responde às interações e revisa textos	- Conhecimento como Conversação (Scholarship as Conversation)

editatonas/datatonas)	coletivamente.	
Verificação da compreensão dos pares antes da publicação e ciclo de reescrita	Paráfrase: Durante as atividades presenciais ou de mentoria, os editores parafraseiam o conteúdo para garantir a clareza; se o texto não estiver claro, é revisado antes da publicação.	- Criação da Informação como um Processo (Information Creation as a Process) - Pesquisa como Exploração Estratégica (Searching as Strategic Exploration)
Criação de hiperligações internas/contextuais na Wikipédia (wikilinks) para conteúdos relacionados	Construção do Conhecimento: Adição de hiperligações que expandem a navegação para o grande público e permitem explorar conceitos relacionados.	- Pesquisa como Exploração Estratégica (Searching as Strategic Exploration)
Wikidata		
Decomposição de um conceito científico em esquemas de dados relacionados	Modelagem Semântica: Cultiva o pensamento computacional e semântico ao estruturar dados complexos de saúde em componentes verificáveis.	- Criação da Informação como um Processo (Information Creation as a Process) - Pesquisa como Investigação (Research as Inquiry)
Adição de referências e identificadores (ex.: DOI e PubMed ID) para sustentar declarações de dados	Avaliação de Evidências: Definição da proveniência e da confiabilidade dos dados por meio da avaliação de evidências e do uso de metadados padronizados.	- Informação Tem Valor (Information Has Value) - Autoridade é Construída e Contextual (Authority Is Constructed and Contextual)
Exploração de informações em formato de dados para extrair e visualizar redes de conhecimento	Raciocínio de Dados: Ilustra o raciocínio algorítmico e a recuperação estratégica de dados.	- Pesquisa como Exploração Estratégica (Searching as Strategic Exploration) - Pesquisa como Investigação (Research as Inquiry)
Colaboração na modelagem de itens ou esquemas de dados por meio de crowdsourcing e discussão	Produção Colaborativa: Oportuniza a negociação, a construção de consenso e a adesão aos padrões da comunidade.	- Conhecimento como Conversação (Scholarship as Conversation)

DISCUSSÃO

Este estudo examinou como a edição na Wikipédia e no Wikidata pode integrar o letramento informacional e o letramento de dados na formação em distúrbios da comunicação.

Letramento usando Wikipédia e Wikidata

A Wikipédia demanda argumentação textual ancorada em fontes secundárias verificáveis, o que aciona a avaliação de evidências, a tradução do conhecimento científico para uma linguagem acessível ao público em geral e a atribuição de fontes de referência (7,18). Esse é um exercício distinto do exigido pelo Wikidata, uma diferenciação que reflete o modelo de letramento subjacente a cada plataforma (8,19). No Wikidata, a decomposição de conceitos em saúde oportuniza que os participantes decidam quais propriedades melhor representam cada conceito e como vinculá-las a fontes verificáveis. Essa sequência de decisões aproxima-se da lógica de construção de ontologias (modelos de representação do conhecimento que definem categorias de conceitos e suas relações) (20). O Wikidata permite representar conceitos como relações estruturadas em uma rede semântica, favorecendo o desenvolvimento de noções de modelagem e o raciocínio sobre as relações entre os dados (8). Essa complementaridade é consistente com a proposta de que o letramento de dados amplia, sem substituir, o letramento informacional: a Wikipédia pode favorecer a comunicação científica escrita; o Wikidata, o raciocínio sobre a estrutura do conhecimento (12,11,13,19).

Competências de Letramento

Embora a literatura documente atividades com a Wikipédia e o Wikidata na educação em saúde (16, 17), ainda há ausência de pesquisas empiricamente orientadas que confirmem se as atividades de edição refletem o desenvolvimento de competências de letramento informacional ou de letramento de dados (7). Os dados observacionais deste estudo permitem considerar que as atividades de edição podem possibilitar que essas competências sejam adquiridas não automaticamente, mas eliciadas de forma mediada e contextual. A análise qualitativa dos artefatos de edição identificou práticas observáveis (Tabela 2), tais como a avaliação de fontes, a síntese de evidências e a modelagem de estruturas de conhecimento. Esses eventos são indicativos das competências em práticas contextuais, desenvolvidas durante atividades de edição. Dessa forma, este estudo contribui com dados exploratórios que indicam que tarefas estruturadas na Wikipédia e no Wikidata podem elicitar competências de letramento informacional e de dados em estudantes da saúde (7,17,18). As competências analisadas basearam-se nos modelos ACRL (14) e LAK (15), servindo como rubrica descritiva para a análise de artefatos de edição de textos e de dados sobre distúrbios da comunicação e temas relacionados (17, 21). A adaptação do modelo LAK, originalmente desenvolvido para a educação em STEM, ao contexto de edição em saúde constitui uma contribuição metodológica exploratória. Da mesma forma, a aplicação dos seis conceitos do modelo ACRL a artefatos de edição, e não a textos acadêmicos convencionais, representa uma extensão do escopo original do modelo de análise (14). A contribuição da análise desenvolvida também se deve à expansão

inédita do escopo para o Wikidata, visto que a literatura apresenta um número maior de relatos focados apenas na Wikipédia (17), com menos investigações sobre o uso do Wikidata na educação em saúde (19).

Influência do Formato das Atividades

A ausência de diferença estatisticamente significativa na conversão de participantes em editores ativos entre os formatos virtual e presencial é compatível com a hipótese de que a curadoria prévia das tarefas de edição e o suporte oferecido por pesquisadores e editores experientes desempenhem um papel mais relevante do que a presença física. Essa hipótese é compatível com relatos de migração de treinamentos presenciais para virtuais em programas de pesquisa, nos quais a estruturação da tarefa compensou a ausência de contato físico (22). O presente estudo estende essa observação ao contexto específico de editatonas e datatonas na saúde. O formato da atividade parece poder ser adaptado ao formato virtual, sem prejuízo da conversão de participantes em editores ativos.

A concentração de edições no Wikidata no formato presencial (Tabela 1) decorreu de um único evento (W8). É possível que a edição de dados no Wikidata (focada em modelagem de informação) precise de maior suporte presencial para contornar barreiras técnicas do que a redação de textos (20,23). A premissa de que o ambiente presencial pode facilitar atividades com Wikidata pode ser uma hipótese de análise. No entanto, a análise estatística não indicou diferenças significativas entre os formatos (IRR=1,26; IC 95%: 0,42 a 3,78; p=0,67).

Estudos anteriores já demonstraram que usuários iniciantes, sem experiência em programação, conseguem compreender e organizar dados estruturados (20). Esse achado reforça a ideia de que o Wikidata, apesar de parecer tecnicamente complexo, é uma plataforma acessível para estudantes da saúde, desde que recebam o suporte adequado (19,20).

A análise dos artefatos (Tabela 2) mostrou que os participantes conseguiram estruturar dados específicos da área da saúde, como taxas de prevalência e identificadores taxonômicos. Isso reforça que a organização de dados no Wikidata não exige formação tecnológica avançada (20). Esse desempenho está diretamente ligado ao nível de suporte oferecido. Um estudo anterior sobre saúde auditiva na Wikipédia já havia demonstrado que intervenções estruturadas geram contribuições mais profundas do que campanhas abertas sem acompanhamento (24). Embora esse estudo não tenha se concentrado em modalidades de ensino, seus achados reforçam a hipótese central desta pesquisa. Os dados indicam que atividades com curadoria de temas, orientação prévia e mentoria simultânea podem ser um fator para a qualidade das edições mais determinante do que o formato ser presencial ou virtual (16,22,24).

Implicações para Comunicação em Saúde

Os achados deste estudo dialogam diretamente com três eixos: a carga de doença dos distúrbios da comunicação e a centralidade do acesso à informação (1,2); a tradução do conhecimento para mitigar infodemias e promover o letramento em saúde (4,5); e o papel das plataformas digitais como veículos dessa tradução (7). Em relação à carga de doença, os resultados demonstram que estudantes da área da saúde produzem

conteúdo verificável sobre distúrbios da comunicação na Wikipédia e no Wikidata quando são apoiados por uma estruturação da atividade. Isso pode sugerir que a Wikipédia e o Wikidata podem ser incorporados à formação profissional como instrumento de tradução do conhecimento (7,18), contribuindo para ampliar o acesso a informações embasadas em evidências sobre distúrbios da comunicação humana (1,2). Quanto à mitigação de infodemias, a tradução do conhecimento para a comunicação em saúde requer práticas de avaliação de evidências e atribuição de fontes verificáveis (4,5). A adição de referências científicas pelos participantes ao longo das sessões indica um comportamento na prática de atribuição de autoridade. Este estudo contribui com evidências sobre a formação dos produtores de informação. As ações de comunicação em saúde baseiam-se não apenas na produção de conteúdo, mas também na disseminação em larga escala, no acesso aberto e no fortalecimento da competência crítica dos leitores para distinguir informação de qualidade da desinformação (7,25). Sobre o papel das plataformas digitais, os achados corroboram parcialmente a afirmação de que essas plataformas facilitam a tradução de evidências científicas (7,8), ao demonstrar que cada plataforma elicia práticas de letramento: a Wikipédia, argumentação textual e atribuição de autoridade; o Wikidata, modelagem semântica e raciocínio relacional (8,19). A centralidade dessas plataformas na tradução do conhecimento em saúde deve ser interpretada com o reconhecimento de que a produção de conteúdo é condição necessária, mas não suficiente, para a tradução efetiva (7,8). A literatura aponta ainda duas lacunas: a concentração dos modelos de letramento informacional e de dados em áreas STEM, com menos investigações em outras disciplinas (13,15,19), e a indicação de ausência de conteúdos sobre distúrbios da comunicação humana (9,10). Em relação à primeira lacuna, os achados indicam que

os conceitos do modelo ACRL (14), especialmente "Autoridade é Construída e Contextual" e "Criação da Informação como um Processo", mostraram-se aplicáveis à análise de artefatos de edição em saúde, capturando práticas como a avaliação de fontes e a tradução de evidências. Em nossa aplicação, por exemplo, o modelo LAK (15) permitiu diferenciar habilidades de manipulação de informação (letramento) de habilidades e conhecimentos de estruturação de dados (habilidades e conhecimentos). Não obstante, a adaptação de um modelo desenvolvido para a educação em engenharia ao contexto de distúrbios da comunicação impõe questões conceituais: categorias oriundas do modelo original, como competência tecnológica e conhecimento de tecnologia da informação, podem não corresponder integralmente às competências envolvidas na edição sobre saúde, e a adequação do modelo a contextos fora da engenharia requer validação no domínio da saúde (15,19). Em relação à segunda lacuna, o estudo demonstrou que é viável integrar a edição sobre distúrbios da comunicação humana às atividades de letramento informacional e de dados. As referências científicas adicionadas (Tabela 1) e as práticas de edição observadas (Tabela 2) indicam que a participação no tema é possível quando mediada por suporte estruturado, sem que isso implique a ausência de barreiras. A existência dessa lacuna na literatura, sugerida por revisões sobre o uso de referências científicas na Wikipédia em tópicos de saúde (9) e por campanhas de promoção da saúde auditiva (10), é compatível com fatores como a complexidade da área, a divisão em subespecialidades (audiologia, motricidade orofacial) e a menor disponibilidade de revisões sistemáticas em comparação com condições de maior prevalência. A curadoria prévia e a mentoria síncrona parecem ter compensado essas barreiras neste contexto. Essa reflexão

reforça a importância de iniciativas que ampliem a cobertura de conteúdo sobre distúrbios da comunicação na Wikipédia e no Wikidata (9,10).

Limitações

Este estudo apresenta limitações que restringem a generalização dos achados. O pequeno número de eventos ($n = 8$) reduz o poder estatístico e amplia os intervalos de confiança das estimativas, tornando os testes estritamente exploratórios. A análise qualitativa exploratória realizada por um único pesquisador pode comprometer a confiabilidade da categorização. A redefinição das dimensões do LAK para este contexto diverge do modelo original e não foi empiricamente testada. A amostra de 30 artefatos, embora estratificada por plataforma, não permite inferir sobre a totalidade das 993 edições registradas. O estudo analisou apenas práticas observáveis em artefatos públicos, sem acesso aos dados individuais dos participantes, o que limita a compreensão dos processos cognitivos subjacentes às práticas de edição.

O desenho transversal impede afirmar que os participantes desenvolveram ou retiveram as competências identificadas de forma individual após a atividade; os resultados descrevem práticas eliciadas durante a intervenção mediada, não sua consolidação como competências retidas.

Por fim, o escopo deste estudo limitou-se à produção do conteúdo, sem avaliar o ciclo de existência das edições nem o impacto direto sobre o público leitor. A escala das edições realizadas não permite inferir um impacto mensurável na qualidade geral da informação sobre distúrbios da comunicação disponível nas plataformas. Além disso, o estudo não acompanhou métricas de acesso (visualizações) nem a persistência dessas

contribuições a longo prazo, o que inviabiliza saber se as informações produzidas foram consultadas pelo público ou se acabaram sendo revertidas pela comunidade de editores da Wikipédia ou do Wikidata.

Perspectivas Futuras

A incorporação de modelos de linguagem generativa aos fluxos de trabalho dos estudantes pode alterar a natureza das competências eliciadas pela edição na Wikipédia (26,27). Se estudantes passarem a utilizar esses modelos para gerar rascunhos de artigos enciclopédicos, a edição pode deixar de exercitar diretamente a avaliação de evidências e a síntese de textos, passando a exercitar preferencialmente a avaliação crítica de textos gerados artificialmente (27). Essa mudança não corresponde necessariamente a um retrocesso: a curadoria crítica de saídas de modelos pode constituir uma competência relevante de letramento informacional (25,28). Ao mesmo tempo, transfere o esforço cognitivo da avaliação de fontes primárias para a edição de superfície, reduzindo o contato direto com evidências verificáveis. Como a edição da Wikipédia é uma ferramenta estabelecida na educação em saúde, investigar de que modo a inteligência artificial altera essa função constitui uma agenda de pesquisa a ser explorada (16,26). Sugere-se a replicação deste estudo com abordagens metodológicas que permitam avaliar a retenção das competências ao longo do tempo, superando as limitações do desenho transversal. Pesquisas futuras com amostras maiores também poderão testar a validade da adaptação dos modelos ACRL e LAK para contextos de saúde, ampliando sua aplicação para além dos cenários STEM e dos contextos originais de desenvolvimento desses modelos (13,15). A necessidade de ampliar o

escopo dos modelos de letramento para acompanhar as transformações digitais e as demandas contemporâneas tem sido destacada na literatura (28). Essa aplicabilidade pode não se restringir a conteúdos sobre distúrbios da comunicação ou a temas de saúde. A estrutura de letramento informacional e de dados, baseada na Wikipédia e no Wikidata, pode orientar investigações semelhantes em outras áreas do conhecimento.

CONCLUSÃO

Este estudo mostrou que a edição na Wikipédia e no Wikidata pode integrar o letramento informacional e o letramento de dados na formação de profissionais da saúde sobre distúrbios da comunicação. As duas plataformas funcionam como pontos de entrada complementares: a Wikipédia mobilizou a avaliação de evidências, a tradução do conhecimento em linguagem acessível e a atribuição de fontes de referência; o Wikidata mobilizou a modelagem e o raciocínio sobre relações entre dados. O formato da atividade, presencial ou virtual, não pareceu influenciar os indicadores de participação.

A principal contribuição deste estudo é o mapeamento metodológico dessa integração em um contexto pouco explorado na literatura sobre letramento e na difusão do conhecimento sobre distúrbios da comunicação humana. A combinação das duas plataformas, em uma única intervenção, articulou a curadoria textual de evidências e a estruturação de dados para a representação do conhecimento. Este mapeamento pode servir de base descritiva para o desenho de intervenções educacionais que combinem letramento informacional e de dados na formação em saúde.

Declaração de Ética: Todos os dados utilizados são públicos e abertos e, assim, de acordo com a Resolução n.º 510/2016 do CNS, não foi necessário submeter este estudo à apreciação ética.

Declaração de Disponibilidade de Dados: Os materiais de scripts e dados tabulados foram depositados no Open Science Framework (<https://doi.org/10.17605/OSF.IO/3DWHT>). Dados da análise de artefatos não estão disponíveis para garantir o anonimato dos participantes das atividades, sem menção a nomes de usuário, títulos de verbetes ou identificadores do Wikidata. Informações complementares podem ser disponibilizadas mediante solicitação.

Financiamento: Esta pesquisa foi apoiada pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Processo n.º 2024/05572-7 e Wikimedia Foundation Community Rapid Fund, Grant ID: 23481525.

Conflito de Interesses: Os autores declaram não haver interesses concorrentes de natureza pessoal ou financeira.

Contribuição dos Autores: Conceitualização: HG, LJ e FZ contribuíram para a concepção e o desenho do estudo. Curadoria de Dados: HG, LJ, FZ e SG foram responsáveis pela curadoria dos dados. Aquisição de Financiamento: LJ e HG obteve o financiamento para a realização do estudo. Investigação: HG, SG e FZ conduziram a investigação. Metodologia: HG, LJ e FZ desenvolveram a metodologia do estudo. Análise: HG, SG, FZ realizaram a análise formal dos dados. Validação: LJ e FZ realizaram a validação dos resultados. Supervisão: LJ e KF supervisionaram o desenvolvimento do estudo. Redação: HG e FZ

elaboraram a versão inicial do texto. Todos os autores contribuíram para a edição do texto e concordaram com a versão final do manuscrito.

Declaração de Uso de Inteligência Artificial: Uso da ferramenta Proton AG Lumo AI revisão textual. Os autores mantêm total responsabilidade pelo conteúdo e pela integridade do manuscrito.

REFERÊNCIAS

1. Cieza A, Causey K, Kamenov K, Hanson SW, Chatterji S, Vos T. Global estimates of the need for rehabilitation based on the Global Burden of Disease study 2019: A systematic analysis for the Global Burden of Disease Study. *Lancet*. 2020;396(10267):2006-2017. [https://doi.org/10.1016/S0140-6736\(20\)32340-0](https://doi.org/10.1016/S0140-6736(20)32340-0)
2. World Health Organization. *World report on hearing*. Geneva: WHO; 2021. <https://www.who.int/publications/i/item/world-report-on-hearing>
3. Turon H, Wolfenden L, Finch M, et al. Dissemination of public health research to prevent non-communicable diseases: A scoping review. *BMC Public Health*. 2022;22(1):1379. <https://doi.org/10.1186/s12889-023-15622-x>
4. Helmer SM, Matthias K, Mergenthal L, Reimer M, de Santis KK. Dissemination of knowledge from Cochrane Public Health reviews: A bibliographic study. *Syst Rev*. 2022;11(1):89. <https://doi.org/10.1186/s13643-023-02272-8>
5. Hicks A, Grant V, Jenkins C. From hierarchies of exclusion to participant-led inclusion: A qualitative research agenda for health information literacy. *Libr Inf Sci Res*. 2023;45(2):101-110. <https://doi.org/10.1016/j.lisr.2024.101295>
6. Yunusa I, Karaye IM. Evidence-based interventions for improving health and medication literacy in community and healthcare settings. In: *Health literacy in medicines use and pharmacy*. Amsterdam: Elsevier; 2023. p. 45-62. <https://doi.org/10.1016/B978-0-12-824407-4.00011-8>
7. Smith DA. Situating Wikipedia as a health information resource in various contexts: A scoping review. *PLoS One*. 2020;15(2):e0228786. <https://doi.org/10.1371/journal.pone.0228786>

8. Waagmeester A, Stupp G, Burgstaller-Muehlbacher S, et al. Wikidata as a knowledge graph for the life sciences. *eLife*; 2020;9:e52614. <https://doi.org/10.7554/eLife.52614>
9. Lewoniewski W, Węcel K, Abramowicz W. Understanding the Use of Scientific References in Multilingual Wikipedia across Various Topics. *Procedia Computer Science* 2023; 225: 3977–3986. <https://doi.org/10.1016/j.procs.2023.10.393>
10. Montilha AAP, Morata TC, Flor DÁ, et al. The Promotion of Hearing Health through Wikipedia Campaigns: Article Quality and Reach Assessment. *Healthcare*. 2023; 11: 1572. <https://doi.org/10.3390/healthcare11111572>
11. Calzada Prado J, Marzal M. Incorporating Data Literacy into Information Literacy Programs: Core Competencies and Contents. *Libri*. 2013;63(2): 123-134. <https://doi.org/10.1515/libri-2013-0010>
12. Koltay T. Data literacy: In search of a name and identity. *J Doc*. 2015;71(3):401-412. <https://doi.org/10.1108/JD-02-2014-0026>
13. Carlson J, Fosmire M, Miller CC, Nelson MS. Determining data information literacy needs: A study of students and research faculty. *Libr Acad*. 2011;11(2):629-657. <https://doi.org/10.1353/pla.2011.0022>
14. American Library Association. *Framework for information literacy for higher education*. Chicago: ACRL; 2015. <http://www.ala.org/acrl/standards/ilframework>
15. Zheng L, Ge D. Cultivating key disciplinary competencies among university students against the background of China's new engineering education initiative. *PLoS ONE* 2025; 20: e0318561. <https://doi.org/10.1371/journal.pone.0318561>
16. Maggio LA, Willinsky JM, Costello JA, et al. Integrating Wikipedia editing into health professions education: a curricular inventory and review of the literature. *Perspect Med Educ*. 2020; 9: 333–342. <https://doi.org/10.1007/s40037-020-00620-1>

17. Stine C. Crowdsourced pedagogy: Editing Wikipedia and the Framework for Information Literacy for Higher Education. *C&RL News*. 2022;83(4):166-170. <https://doi.org/10.5860/crln.83.4.166>
18. Smith D. Leveraging Wikipedia in undergraduate health sciences education: A key tool for information literacy and knowledge translation. *J Can Health Libr Assoc*. 2023;44(3):79-84. <https://doi.org/10.29173/jchla29688>
19. Evenstein Sigalov SE, Cohen A, Nachmias R. Open educational resources for literacies, diversity, equity and inclusion: The case of integrating Wikidata into higher education. In: Olivier J, Rambow A, editors. *Open educational resources in higher education: Future education and learning spaces*. Singapore: Springer; 2023. p. 279-306. https://doi.org/10.1007/978-981-19-8590-4_13
20. Zhang L, Yang X, Cota Z, et al. Which methods are the most effective in enabling novice users to participate in ontology creation? A usability study. *Database*. 2023;2023:baad076. <https://doi.org/10.1093/database/baab035>
21. Julien H, Gross M, Latham D, editors. *The information literacy framework: Case studies of successful implementation*. Lanham: Rowman & Littlefield; 2016.
22. Gallo JJ, Murray SM, Creswell JW, Deutsch C, Guetterman TC. Going virtual: Mixed methods evaluation of online versus in-person learning in the NIH mixed methods research training program retreat. *BMC Med Educ*. 2024;24(1):882. <https://doi.org/10.1186/s12909-024-05877-2>
23. Casimo K. Teaching and training with open science: From classroom teaching tool to professional development. *Neuroscience*. 2023;15(3):445-462. <https://doi.org/10.1016/j.neuroscience.2023.07.013>

24. Morata TC, Zucki F, Arrigo AJ, Cruz PC, Gong W, Matos HGC, et al. Strategies for crowdsourcing hearing health information: A comparative study of educational programs and volunteer-based campaigns on Wikimedia. *BMC Public Health*. 2024;24(1):2646. <https://doi.org/10.1186/s12889-024-20105-8>
25. Drabinski E, Tewell E. Critical Information Literacy. *The International Encyclopedia of Media Literacy* 2019; 1–4. <https://doi.org/10.1002/9781118978238.ieml0042>
26. Evenstein Sigalov S, Cohen A, Nachmias R. Transforming higher education: a decade of integrating wikipedia and wikidata for literacy enhancement and social impact. *J Comput Educ* 2024; 12: 953–995. <https://doi.org/10.1007/s40692-024-00334-x>
27. Matos HGC, Wasmann JWA, Morata TC, Alvarenga KF, Jacob LCB. Large language models readability classification: A variability analysis of sources and metrics. *medRxiv*. 2026;2026.02.20.26346638. <https://doi.org/10.64898/2026.02.20.26346638>
28. Machin-Mastromatteo JD. Literacy of the future. In: *Encyclopedia of libraries, librarianship, and information science*. Vol. 4. Amsterdam: Elsevier; 2025. p. 376-387. <https://doi.org/10.1016/B978-0-323-95689-5.00197-8>

Este preprint foi submetido sob as seguintes condições:

- Os autores declaram que os necessários Termos de Consentimento Livre e Esclarecido de participantes ou pacientes na pesquisa foram obtidos e estão descritos no manuscrito, quando aplicável.
- Os autores declaram que a elaboração do manuscrito seguiu as normas éticas de comunicação científica.
- Os autores declaram que estão cientes que são os únicos responsáveis pelo conteúdo do preprint e que o depósito no SciELO Preprints não significa nenhum compromisso de parte do SciELO, exceto sua preservação e disseminação.
- Os autores declaram que os dados, aplicativos e outros conteúdos subjacentes ao manuscrito estão referenciados.
- O manuscrito depositado está no formato PDF.
- Os autores declaram que a pesquisa que deu origem ao manuscrito seguiu as boas práticas éticas e que as necessárias aprovações de comitês de ética de pesquisa, quando aplicável, estão descritas no manuscrito.
- Os autores declaram que uma vez que um manuscrito é postado no servidor SciELO Preprints, o mesmo só poderá ser retirado mediante pedido à Secretaria Editorial do SciELO Preprints, que afixará um aviso de retratação no seu lugar.
- Os autores concordam que o manuscrito aprovado será disponibilizado sob licença [Creative Commons CC-BY](#).
- O autor submissor declara que as contribuições de todos os autores e declaração de conflito de interesses estão incluídas de maneira explícita e em seções específicas do manuscrito.
- Os autores declaram que o manuscrito não foi depositado e/ou disponibilizado previamente em outro servidor de preprints ou publicado em um periódico.
- Caso o manuscrito esteja em processo de avaliação ou sendo preparado para publicação mas ainda não publicado por um periódico, os autores declaram que receberam autorização do periódico para realizar este depósito.
- O autor submissor declara que todos os autores do manuscrito concordam com a submissão ao SciELO Preprints.