

TENDENCIAS EN EL ACCESO ABIERTO A INFORMACIÓN AMBIENTAL EN LA AMAZONÍA

Espinoza-Mina, Marcos Antonio
Colina-Vargas, Alejandra Mercedes



Tendencias en el acceso abierto a información ambiental en la Amazonía.

Autor/es:

Espinoza-Mina, Marcos Antonio

Universidad Tecnológica ECOTEC

Colina-Vargas, Alejandra Mercedes

Universidad Tecnológica ECOTEC

Espinoza-Mina, M. A.
Colina-Vargas, A. M.

Tendencias en el acceso abierto a información ambiental en la Amazonía.

Editorial Grupo AEA, Ecuador, 2024

ISBN: 978-9942-651-63-1

Formato: 210 cm X 270 cm

110 págs.



Publicado por Editorial Grupo AEA

Ecuador, Santo Domingo, Vía Quinindé, Urb. Portón del Río.

Contacto: +593 983652447; +593 985244607

Email: info@editorialgrupo-aea.com

<https://www.editorialgrupo-aea.com/>

Director General:

Prof. César Casanova Villalba.

Editor en Jefe:

Prof. Giovanni Herrera Enríquez

Editora Académica:

Prof. Maybelline Jaqueline Herrera Sánchez

Supervisor de Producción:

Prof. José Luis Vera

Diseño:

Tnlgo. Oscar J. Ramírez P.

Consejo Editorial

Editorial Grupo AEA

Primera Edición, 2024

D.R. © 2024 por Autores y Editorial Grupo AEA Ecuador.

Cámara Ecuatoriana del Libro con registro editorial No 708

Disponible para su descarga gratuita en <https://www.editorialgrupo-aea.com/>

Los contenidos de este libro pueden ser descargados, reproducidos difundidos e impresos con fines de estudio, investigación y docencia o para su utilización en productos o servicios no comerciales, siempre que se reconozca adecuadamente a los autores como fuente y titulares de los derechos de propiedad intelectual, sin que ello implique en modo alguno que aprueban las opiniones, productos o servicios resultantes. En el caso de contenidos que indiquen expresamente que proceden de terceros, deberán dirigirse a la fuente original indicada para gestionar los permisos.

Título del libro:

Tendencias en el acceso abierto a información ambiental en la Amazonía.

© Espinoza Mina, Marcos Antonio; Colina Vargas, Alejandra Mercedes.

© Diciembre, 2024

Libro Digital, Primera Edición, 2024

Editado, Diseñado, Diagramado y Publicado por Comité Editorial del Grupo AEA, Santo Domingo de los Tsáchilas, Ecuador, 2024

ISBN: 978-9942-651-63-1



<https://doi.org/10.55813/egaea.l.108>

Como citar (APA 7ma Edición):

Espinoza-Mina, M. A., Colina-Vargas, A. M. (2024). *Tendencias en el acceso abierto a información ambiental en la Amazonía*. Editorial Grupo AEA. <https://doi.org/10.55813/egaea.l.108>

Cada uno de los textos de Editorial Grupo AEA han sido sometido a un proceso de evaluación por pares doble ciego externos (double-blindpaperreview) con base en la normativa del editorial.

Revisores:



Ing. Rodríguez Deza Jorge
Washington, Ph. D.

Universidad Nacional de
Huancavelica – Perú



Ing. Caicedo Aldaz Julio César,
Ph. D. (c)

Universidad Técnica Luis Vargas
Torres, Universidad Laica Eloy
Alfaro de Manabí – Ecuador



Los libros publicados por “**Editorial Grupo AEA**” cuentan con varias indexaciones y repositorios internacionales lo que respalda la calidad de las obras. Lo puede revisar en los siguientes apartados:



Editorial Grupo AEA

 <http://www.editorialgrupo-aea.com>

 Editorial Grupo AeA

 editorialgrupoea

 Editorial Grupo AEA

Aviso Legal:

La información presentada, así como el contenido, fotografías, gráficos, cuadros, tablas y referencias de este manuscrito es de exclusiva responsabilidad del/los autor/es y no necesariamente reflejan el pensamiento de la Editorial Grupo AEA.

Derechos de autor ©

Este documento se publica bajo los términos y condiciones de la licencia Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0).



El “copyright” y todos los derechos de propiedad intelectual y/o industrial sobre el contenido de esta edición son propiedad de la Editorial Grupo AEA y sus Autores. Se prohíbe rigurosamente, bajo las sanciones en las leyes, la producción o almacenamiento total y/o parcial de esta obra, ni su tratamiento informático de la presente publicación, incluyendo el diseño de la portada, así como la transmisión de la misma de ninguna forma o por cualquier medio, tanto si es electrónico, como químico, mecánico, óptico, de grabación o bien de fotocopia, sin la autorización de los titulares del copyright, salvo cuando se realice confines académicos o científicos y estrictamente no comerciales y gratuitos, debiendo citar en todo caso a la editorial. Las opiniones expresadas en los capítulos son responsabilidad de los autores.

RESEÑA DE AUTORES

AUTORES



Espinoza Mina Marcos Antonio



Universidad Tecnológica ECOTEC



mespinoza@ecotec.edu.ec



<https://orcid.org/0000-0003-1530-7243>



Ingeniero en Sistemas Computacionales, Magister en Negocios Internacionales y Comercio Exterior, Magister en Sistemas de Información, Magister en Estadística Aplicada y Doctor en Administración de Empresas. Miembro de la Red de Investigación, de Conocimiento Hardware y Software Libre. Investigador Agregado 2, Secretaría de Educación Superior, Ciencia, Tecnología e Innovación. Perito del Consejo de la Judicatura en la especialidad de Ingeniería Informática o de Sistemas. Profesional multidisciplinario, con sólida formación en investigación, docencia, y liderazgo de proyectos, tanto en los ámbitos académicos como empresariales. Posee habilidades que abarcan la gestión de proyectos, las tecnologías de la información, la gestión empresarial y el análisis de datos. Ímpetu por el aprendizaje continuo y la adaptabilidad, le ha permitido destacar en diversos roles, y aplicar la experiencia para abordar desafíos, fomentando el crecimiento, tanto de los estudiantes como de las organizaciones donde ha colaborado



Colina Vargas Alejandra Mercedes



Universidad Tecnológica ECOTEC



acolina@ecotec.edu.ec



<https://orcid.org/0000-0003-1514-8852>



Ingeniera de Sistemas, Magister en Gerencia de Tecnologías de Información y Comunicación, Magister en Sistemas de Información Mención en Inteligencia de Negocios y Analítica de Datos Masivos y Doctora en Educación. Miembro de la Red de Investigación, de Conocimiento Hardware y Software Libre. Investigador Agregado 1, Secretaría de Educación Superior, Ciencia, Tecnología e Innovación. Perito del Consejo de la Judicatura en la especialidad de Ingeniería Informática o de Sistemas. Profesional con más de 20 años de experiencia, investigadora de alto nivel con habilidades y competencias para el diseño y gestión de proyectos de apropiación, desarrollo y aplicación de las Tecnologías de Información y Comunicación, hardware y software libre, y otras tecnologías de vanguardia, que contribuyan a los procesos de transformación social y educativa del país. Destacada en la elaboración de propuestas de analítica de datos soportadas con software específicos aplicados a instituciones de administración pública y privada

Índice

Reseña de Autores.....	ix
Índice.....	xi
Índice de Tablas	xiii
Índice de Figuras.....	xiv
Introducción.....	xv
Capítulo I: El objeto de estudio	1
1.1. Definición del objeto de estudio.....	3
1.2. Criterios de acceso abierto	4
1.3. Importancia del acceso abierto en el contexto de la Amazonía.....	4
1.4. El rol de la OTCA en la promoción del acceso abierto	5
1.5. Hipótesis del estudio.....	6
1.6. Beneficios esperados del estudio	7
Capítulo II: Metodología	9
2.1. Introducción a PRISMA y su relevancia para el estudio	11
2.2. Aplicación de PRISMA.....	11
2.2.1. Identificación.....	12
2.2.2. Selección	13
2.2.3. Elegibilidad	13
2.2.4. Inclusión.....	14
2.3. Limitaciones	16
Capítulo III: Análisis bibliométrico	19
3.1. Descripción de la producción científica.....	21
3.1.1. Periodo de análisis	21
3.1.2. Mapeo de la contribución por país.....	22
3.1.3. Panorama institucional de las publicaciones	25
3.1.4. Revistas claves en la producción científica	26

3.1.5.	Distribución de los artículos según Ley de Bradford	27
3.1.6.	Análisis a través de la Ley de Lotka	29
3.2.	Análisis de redes.....	31
3.2.1.	Mapeo de redes basado en VOSviewer	31
3.2.2.	Evaluación de métricas cuantitativas de red	33
3.2.3.	Análisis por co-autoría	34
3.2.3.1.	Red de co-autoría por autor	35
3.2.3.1.1.	Mapeo de red de co-autoría por autor.....	35
3.2.3.1.2.	Medidas adicionales del análisis de co-autoría por autor	36
3.2.3.2.	Red de co-autoría por organización	38
3.2.3.2.1.	Mapeo de red de co-autoría por organización	38
3.2.3.2.2.	Medidas adicionales del análisis de co-autoría por organización.....	40
3.2.3.3.	Red de co-autoría por países	42
3.2.3.3.1.	Mapeo de red de co-autoría por países	42
3.2.3.3.2.	Medidas adicionales del análisis de co-autoría por autor	43
3.2.4.	Análisis de co-ocurrencia de palabras claves.....	46
3.2.4.1.	Mapeo de red de co-autoría por países	46
3.2.4.1.1.	Medidas adicionales del análisis de co-ocurrencia	47
3.2.5.	Análisis de co-citación	50
3.2.5.1.	Red de co-citación por autores citados	51
3.2.5.1.1.	Mapeo de red de co-citación por autores citados	51
3.2.5.1.2.	Medidas adicionales del análisis de co-citación por autores citados	52
3.2.5.2.	Red de co-citación por fuentes citadas	54
3.2.5.2.1.	Mapeo de red de co-citación por fuentes citadas.....	54
3.2.5.2.2.	Medidas adicionales del análisis de co-autoría por autor	55
3.2.5.3.	Red de co-citación por referencias citadas	59

3.2.5.3.1. Mapeo de red de co-citación por referencias citadas.....	59
3.2.5.3.2. Medidas adicionales del análisis de co-citación por referencias citadas.....	60
3.2.6. Análisis temático-basados en datos de textos.....	63
Capítulo IV: Análisis de datos avanzados	67
4.1. Técnicas de machine learning	69
4.2. Análisis Temático con LDA y K-Means	70
4.3. Interpretación de los clústeres	73
Capítulo V: Evaluación de las fuentes de acceso abierto	77
5.1. Justificación del análisis de fuentes primarias	79
5.2. Procedimiento realizado	80
5.3. Lo encontrado en la evaluación	81
Capítulo VI: Discusión, conclusiones y recomendaciones	89
6.1. Discusión de los hallazgos más relevantes	91
6.2. Conclusiones	94
Referencias Bibliográficas.....	99

Índice de Tablas

Tabla 1 Países con mayores citaciones	23
Tabla 2 Países de los autores con mayor cantidad de artículos	23
Tabla 3 Instituciones más relevantes según filiación de los autores y sus artículos	25
Tabla 4 División de revistas según las zonas de la ley de Bradford.....	27
Tabla 5 Revistas pertenecientes a la Zona 1 según ley de Bradford	28
Tabla 6 Distribución observada y teórica de la productividad científica	30
Tabla 7 Resultados de medidas adicionales del análisis de co-autor por autor.....	38
Tabla 8 Resultados de medidas adicionales del análisis de co-autoría por organización	41

Tabla 9 Resultados de medidas adicionales del análisis de co-autoría por países	45
Tabla 10 Resultados de medidas adicionales del análisis de co-ocurrencia	49
Tabla 11 Resultados de medidas adicionales del análisis de co-citación por autores citados	53
Tabla 12 Comunidades detectadas y disciplinas relacionadas	53
Tabla 13 Resultados de medidas adicionales del análisis de co-citación por fuentes citadas	57
Tabla 14 Resultados de medidas adicionales del análisis de co-citación por referencias citadas	61
Tabla 15 Resultado de medida adicional comunidad	63
Tabla 16 Temas LDA	71
Tabla 17 Palabras más frecuentes por clúster	72
Tabla 18 Interpretación de Clústeres	74
Tabla 19 Datos primarios en estudios científicos	84

Índice de Figuras

Figura 1 Producción científica reciente, último trienio	21
Figura 2 Las 10 fuentes más prolíficas en la bibliometría realizada	26
Figura 3 Frecuencia acumulada de artículos por posición de revista	29
Figura 4 Distribución de autores y productividad científica según Ley de Lotka	31
Figura 5 Visualización de red co-autoría por autor	36
Figura 6 Visualización de red co-autoría por organización	39
Figura 7 Visualización de red co-autoría por países	42
Figura 8 Visualización del análisis de co-ocurrencias en VOSviewer	46
Figura 9 Visualización de red co-citación de autores citados	51
Figura 10 Visualización de red co-citación de fuentes citadas	55
Figura 11 Visualización de red co-citación de referencias citadas	59
Figura 12 Visualización del análisis de temas basados en texto de datos	65
Figura 13 Clustering de Artículos Científicos usando LDA + K-Means	73

Introducción

La Amazonía, considerada un verdadero refugio de biodiversidad y un regulador climático esencial para el planeta, se extiende por más de 7 millones de kilómetros cuadrados y alberga el 10% de las especies del mundo. Este vasto ecosistema desempeña un papel crucial en el control del clima, el ciclo del agua y la captura de carbono, siendo indispensable para la estabilidad ambiental global (Filho et al., 2021; Yuan et al., 2022).

Más allá de su papel en la regulación climática, este ecosistema es también una fuente invaluable de compuestos naturales con propiedades beneficiosas para la salud. Un ejemplo es el cannabidiol (CBD), un compuesto derivado de plantas presentes en la región, que ha demostrado ser eficaz en la reducción de síntomas de agotamiento emocional en profesionales de la salud (Crippa et al., 2021). Además, investigaciones indican que la exposición a entornos naturales y enriquecidos, como los que ofrece la Amazonía, puede aliviar el estrés y la ansiedad, contribuyendo al bienestar mental (Figueiredo Cerqueira et al., 2023). Estos descubrimientos subrayan la necesidad de conservar este ecosistema, no solo por sus funciones ambientales, sino también por su potencial para mejorar la salud humana.

La cuenca amazónica abarca varios países de América del Sur y enfrenta múltiples desafíos ambientales, incluyendo la deforestación, la degradación de los ecosistemas y el cambio climático, factores que afectan no solo la biodiversidad sino también las comunidades locales que dependen de estos recursos (Magalhães et al., 2023; Vargas-Munoz et al., 2021). Estos desafíos resaltan la interdependencia entre la salud humana y los sistemas naturales, un concepto central en la salud planetaria, que destaca cómo la estabilidad del entorno natural es esencial para el bienestar humano (Shaw et al., 2021).

La Amazonía no solo se ve afectada por las perturbaciones naturales como incendios forestales e infestaciones de insectos, que modifican la composición de la comunidad y la diversidad funcional del ecosistema, sino también por las crecientes temperaturas globales que agravan la frecuencia y la intensidad de estos eventos (Chen et al., 2021; Viljur et al., 2022). Los incendios forestales, en particular, liberan contaminantes como el material particulado (PM2.5), asociado con un aumento en la mortalidad por enfermedades cardiovasculares y respiratorias, lo que subraya la necesidad urgente de comprender estos riesgos para la salud (Chen et al., 2021).

La sostenibilidad se ha convertido en un objetivo central para los países que buscan reducir su impacto ambiental y promover el uso eficiente de recursos naturales. Estrategias como la mejora en la productividad agrícola y la implementación de prácticas sostenibles son clave para alcanzar una gestión

más responsable del suelo y del agua, favoreciendo la conservación de la biodiversidad y la adaptación al cambio climático.

En el ámbito de la investigación científica y tecnológica, el uso de herramientas avanzadas como la inteligencia artificial ha demostrado ser una espada de doble filo. Si bien estas tecnologías tienen el potencial de acelerar la investigación y mejorar la eficiencia en el procesamiento de grandes volúmenes de datos, también existe el riesgo de que la información generada pueda ser inexacta o incluso perjudicial si no se verifica adecuadamente (Oviedo-Trespalacios et al., 2023). Esto es particularmente relevante en el contexto de la Amazonía, donde la toma de decisiones basada en datos fiables es crucial para abordar desafíos complejos relacionados con la sostenibilidad y la conservación.

El acceso abierto a la investigación es fundamental para avanzar en el conocimiento y enfrentar los desafíos de la Amazonía, permitiendo que científicos, responsables políticos y comunidades locales colaboren más eficazmente (Brack et al., 2022; Kass et al., 2023). Al eliminar las barreras económicas y legales, se promueve una difusión más equitativa de los resultados de investigación, lo que es esencial en una región donde los problemas ambientales son transfronterizos y requieren respuestas coordinadas (Pereira et al., 2021).

Además, el uso de herramientas de código abierto ha potenciado las capacidades de modelización de nichos ecológicos y distribuciones de especies, lo que ha facilitado la reproducibilidad y la colaboración dentro de la comunidad científica. Estas tecnologías permiten una mayor adaptabilidad y expansión de metodologías, promoviendo un enfoque más transparente y eficiente para abordar preguntas ecológicas complejas (Kass et al., 2023).

La investigación ambiental en la Amazonía ha avanzado significativamente gracias a la disponibilidad de grandes conjuntos de datos y tecnologías avanzadas como la teledetección y los sistemas de información geográfica (SIG) (Silva et al., 2023; Zhao et al., 2021). Estas herramientas permiten analizar la deforestación, la biodiversidad y otros procesos ecológicos a escalas sin precedentes, lo que proporciona una base sólida para la toma de decisiones informadas sobre la conservación de la región.

Técnicas como el procesamiento del lenguaje natural (PLN) y el uso de redes neuronales han permitido una extracción de información más precisa de grandes volúmenes de datos científicos, ayudando a identificar patrones y tendencias emergentes en la investigación ecológica (Camalan et al., 2022; Mansouri et al., 2021). El PLN facilita la extracción automática de información relevante de artículos científicos, lo que permite generar resúmenes y visualizaciones que hacen más accesible el conocimiento a un público más amplio.

La región amazónica también enfrenta serios problemas de salud pública debido a la prevalencia de enfermedades tropicales desatendidas, como la enfermedad

de Chagas y la leishmaniasis, que se ven exacerbadas por factores socioeconómicos y cambios ambientales (Gómez-Ochoa et al., 2022; Magalhães et al., 2023). El acceso abierto a la investigación ha sido crucial para la respuesta a emergencias de salud como la pandemia de COVID-19, facilitando la difusión de conocimientos científicos que apoyan decisiones rápidas y efectivas (Friedrich et al., 2021).

A pesar de los beneficios del acceso abierto, aún persisten importantes desafíos como la brecha digital y la falta de recursos en instituciones de la región amazónica. El fortalecimiento de las infraestructuras de datos abiertos y la capacitación de investigadores son fundamentales para superar estas barreras y fomentar una colaboración científica más efectiva (Friedrich et al., 2021; Mestanza-Ramón et al., 2023).

La Organización del Tratado de Cooperación Amazónica (OTCA) desempeña un papel fundamental en la promoción del desarrollo sostenible de la Amazonía. Como el único bloque socioambiental de Latinoamérica, la OTCA facilita la cooperación entre sus ocho países miembros para alcanzar un equilibrio entre el aprovechamiento de los recursos, su protección y conservación (OTCA, 2024). A través de la Agenda Estratégica de Cooperación Amazónica (AECA), la OTCA impulsa acciones conjuntas para abordar las prioridades de la región, incluyendo la gestión de la biodiversidad, los recursos hídricos y el cambio climático.

En este contexto, el acceso abierto a la información ambiental y ecológica se vuelve esencial para el cumplimiento de la misión de la OTCA. El acceso abierto, como componente clave de la "ciencia abierta", fortalece las capacidades regionales para la investigación científica, la toma de decisiones y la participación ciudadana en la gestión ambiental. La ciencia abierta puede ser una herramienta eficaz para fortalecer el trabajo colaborativo, mejorar la transparencia y la rendición de cuentas en la investigación (Babini & Rovelli, 2020). Además, al promover la difusión de la cultura y los saberes ancestrales de la Amazonía, el acceso abierto contribuye a la construcción de un desarrollo sostenible que respete la diversidad cultural y promueva la equidad en la región.

Dada la riqueza biológica y la importancia ecológica de la región amazónica, así como los desafíos ambientales que enfrenta, es esencial avanzar en el acceso al conocimiento científico para tomar decisiones informadas. Evaluar la evolución del acceso abierto a investigaciones ecológicas en los países amazónicos es clave para fomentar la colaboración entre científicos, responsables políticos y comunidades locales, superando barreras que limitan la difusión del conocimiento y facilitando una respuesta más efectiva a los problemas ambientales. En este contexto, el presente estudio ofrece una visión del estado actual de la difusión del conocimiento científico, destacando la necesidad de un acceso más equitativo a los datos ecológicos y la optimización de la toma de decisiones para la conservación y el desarrollo sostenible de la Amazonía.

Este análisis bibliométrico tiene como objetivo explorar las tendencias y el estado actual del acceso abierto a la información ambiental y ecológica en la región amazónica, a partir de publicaciones científicas recientes (2021-2023). Busca identificar patrones en la producción y difusión del conocimiento científico dentro de los países miembros de la Organización del Tratado de Cooperación Amazónica (OTCA), evaluando la disponibilidad y el impacto de la información accesible públicamente. Además, pretende detectar las principales barreras y oportunidades para optimizar el intercambio de datos y fortalecer la colaboración científica, apoyando así estrategias de conservación y desarrollo sostenible en la región amazónica.

CAPITULO 1

EL OBJETO DE ESTUDIO



El objeto de estudio

1.1. Definición del objeto de estudio

Este estudio se centra en el análisis del acceso abierto a la información ambiental y ecológica en la región amazónica. El análisis se enfoca específicamente en publicaciones científicas disponibles en inglés y provenientes de los países miembros de la Organización del Tratado de Cooperación Amazónica (OTCA): Bolivia, Brasil, Colombia, Ecuador, Guyana, Perú, Surinam y Venezuela.

Se entiende por "información ambiental y ecológica" el conjunto de conocimientos, datos e investigaciones científicas que documentan y analizan los ecosistemas amazónicos. Esto incluye el estudio de su biodiversidad, los recursos naturales, las amenazas ambientales que enfrentan, y las estrategias para su conservación y desarrollo sostenible. La Amazonía posee una rica biodiversidad y juega un papel crucial en la regulación de las precipitaciones en Sudamérica, pero también enfrenta amenazas como la deforestación y el cambio climático, que podrían resultar en la liberación de grandes cantidades de carbono almacenado en la vegetación (Potsdam Institute for Climate Impact Research, 2023).

A partir de una búsqueda realizada el 10 de septiembre de 2024, se identificaron 1.007 documentos que cumplen con los criterios de selección: acceso abierto, tipo artículo, fecha de publicación entre 2021 y 2023, idioma inglés y país de origen perteneciente a la OTCA. Estos documentos constituyen la base para el análisis bibliométrico que se realiza en este estudio. El análisis bibliométrico permite identificar patrones en la producción y difusión del conocimiento científico en la región, evaluar la disponibilidad y el impacto de la información de acceso abierto, y detectar las principales barreras y oportunidades para optimizar el intercambio de datos y fortalecer la colaboración científica.

1.2. Criterios de acceso abierto

Para que la información se considere de "acceso abierto", debe cumplir con criterios específicos. Debe estar disponible de forma gratuita en internet, sin restricciones de pago o suscripción. Esto permite el acceso a cualquier persona con conexión a internet, independientemente de su afiliación institucional o ubicación geográfica, democratizando el acceso al conocimiento.

La información debe ser fácilmente accesibles, interoperables y reutilizables (Redkina, 2021). Los tipos de licenciamiento deben incluir la posibilidad de realizar modificaciones y trabajos derivados, siempre que se cite la fuente original. Estas licencias, como las Creative Commons, garantizan que la información se pueda compartir y utilizar de manera amplia para fines de investigación, educación y toma de decisiones. Además, la información debe contar con metadatos completos y precisos que faciliten su descubrimiento e identificación a través de motores de búsqueda y repositorios. Los metadatos incluyen información sobre el autor, el título, la fecha de publicación, palabras clave y otros detalles relevantes que permiten la indexación y recuperación eficiente de la información.

1.3. Importancia del acceso abierto en el contexto de la Amazonía

La Amazonía, con su rica biodiversidad y vastos recursos naturales, enfrenta graves amenazas ambientales. Entre ellas, se encuentran la deforestación, la degradación de los ecosistemas, la pérdida de biodiversidad, el cambio climático y la contaminación. En este contexto, el acceso abierto a la información ambiental y ecológica se vuelve esencial para abordar estos desafíos, ya que facilita la generación y difusión del conocimiento necesario para la toma de decisiones y la acción.

El acceso abierto fortalece la investigación científica al facilitar la colaboración entre investigadores, el intercambio de conocimientos y el avance de la ciencia en la Amazonía. Permite a los investigadores acceder a una amplia gama de

datos y publicaciones, lo que facilita la replicación de estudios, la comparación de resultados y la generación de nuevo conocimiento.

Además, apoya la toma de decisiones informadas por parte de gobiernos, organizaciones y comunidades locales sobre la gestión ambiental y el desarrollo sostenible. Al proporcionar acceso a datos e investigaciones científicas, el acceso abierto contribuye a la formulación de políticas públicas más efectivas y a la implementación de estrategias de conservación basadas en evidencia. Se complementa el hecho de ayudar a reducir los costos de acceso a la información científica, tanto para las instituciones como para los investigadores individuales (Boufarss & Harviainen, 2021).

Asimismo, promueve la participación ciudadana en los debates y la toma de decisiones sobre el futuro de la Amazonía. Al facilitar el acceso a la información ambiental, el acceso abierto empodera a los ciudadanos para comprender los problemas ambientales, participar en la toma de decisiones y contribuir a la búsqueda de soluciones.

1.4. El rol de la OTCA en la promoción del acceso abierto

La Organización del Tratado de Cooperación Amazónica (OTCA) juega un papel fundamental en la promoción de la cooperación regional en la Amazonía. La OTCA busca el desarrollo armónico de la región a través de la cooperación entre sus países miembros, facilitando el intercambio de información, promoviendo la investigación científica conjunta y apoyando la toma de decisiones informadas para la conservación y el desarrollo sostenible de la Amazonía (OTCA, 2023).

El acceso abierto a la información ambiental y ecológica se alinea con la misión de la OTCA, fortaleciendo las capacidades regionales para la investigación científica, la toma de decisiones y la participación ciudadana en la gestión ambiental (OTCA, 2021). Al promover el acceso abierto, la OTCA puede contribuir a la construcción de una base de conocimiento sólida y compartida para la toma de decisiones en la región.

1.5. Hipótesis del estudio

Este estudio parte de la premisa de que el acceso abierto a la información ambiental y ecológica en la Amazonía ha experimentado un notable crecimiento en los últimos años. Factores como la implementación de políticas de acceso abierto por parte de agencias de financiamiento e instituciones, la creciente conciencia sobre los beneficios del acceso abierto dentro de la comunidad científica y el desarrollo de infraestructuras digitales que facilitan la difusión del conocimiento, han contribuido a este auge. Sin embargo, se reconoce que aún existen desafíos que impiden su total aprovechamiento para la investigación, la toma de decisiones y la participación ciudadana en temas ambientales.

En este sentido, se espera observar un aumento en la producción científica sobre temas ambientales y ecológicos de la Amazonía en acceso abierto, impulsado por las políticas de acceso abierto y el desarrollo de repositorios institucionales y temáticos. Asimismo, se anticipa una correlación positiva entre el acceso abierto y el impacto de las publicaciones científicas, medido a través de métricas tradicionales, como el número de citas, y métricas alternativas, como las menciones en redes sociales y la reutilización de datos. Se presume que el acceso abierto facilita la colaboración científica entre investigadores e instituciones de los países miembros de la OTCA, promoviendo la investigación multidisciplinaria y el intercambio de conocimientos.

No obstante, se reconoce la persistencia de barreras que dificultan el acceso abierto a la información ambiental y ecológica en la Amazonía. Estas barreras incluyen la falta de infraestructura tecnológica en algunas zonas de la región, el desconocimiento sobre las políticas y beneficios del acceso abierto, los costos de publicación en revistas de acceso abierto y la resistencia al cambio en las prácticas de publicación por parte de algunos investigadores. Además, se plantea la existencia de una brecha en el acceso a la información científica entre los países miembros de la OTCA, con una mayor concentración de la producción y el acceso en Brasil y Colombia, debido a las diferencias en el desarrollo científico y tecnológico, la inversión en investigación y las políticas de acceso abierto.

1.6. Beneficios esperados del estudio

Este estudio aspira a desempeñar un papel crucial en la visibilización del estado actual del acceso abierto, así como en la identificación de tendencias clave en la producción y difusión de información. Además, proporcionará datos valiosos que apoyarán la toma de decisiones en la formulación de políticas de acceso abierto, fomentarán la cooperación científica entre los países miembros de la OTCA y promoverán la conservación y el desarrollo sostenible de la Amazonía a través del libre acceso al conocimiento.

Los hallazgos de este estudio ofrecerán a los investigadores de la región amazónica una visión más clara y profunda del acceso abierto en su ámbito de trabajo, ayudándolos a identificar estrategias óptimas para publicar y acceder a información relevante. Asimismo, la información generada será de gran utilidad para los responsables de la toma de decisiones en gobiernos, organizaciones internacionales e instituciones de investigación, quienes podrán diseñar políticas de acceso abierto más efectivas y adaptadas a las necesidades locales. Además, el acceso abierto a la información ambiental y ecológica no solo facilita la investigación, sino que también empodera a la sociedad civil, permitiéndole participar activamente en los debates sobre el futuro de la Amazonía.

El acceso abierto a la información ambiental y ecológica es un pilar fundamental para el desarrollo sostenible de la Amazonía. Este estudio tiene como objetivo profundizar en la comprensión de este fenómeno y contribuir a su fortalecimiento en la región, impulsando el conocimiento y la colaboración en favor de la conservación y el desarrollo sostenible.

CAPITULO 2

METODOLOGÍA



Metodología

2.1. Introducción a PRISMA y su relevancia para el estudio

Este estudio bibliométrico se fundamenta en la metodología PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), una guía internacionalmente reconocida desarrollada por expertos para estandarizar la forma en que se reportan las revisiones sistemáticas y meta-análisis. PRISMA busca mejorar la transparencia, coherencia y rigurosidad de estos estudios, facilitando la comprensión del proceso de revisión y permitiendo una evaluación crítica de los resultados.

La última versión, PRISMA 2020, se compone de 27 ítems que abarcan todas las etapas de una revisión sistemática, desde la formulación de la pregunta de investigación y la definición de criterios de elegibilidad, hasta la extracción de datos, análisis e interpretación de resultados. Además, incluye un diagrama de flujo que visualiza el proceso de selección de estudios de manera clara y concisa (Page et al., 2021).

La elección de PRISMA como base metodológica para este estudio se justifica por su enfoque estandarizado y adaptable, que guía todas las etapas del proceso desde la identificación hasta la síntesis de datos. PRISMA 2020 incorpora técnicas modernas de análisis, como el aprendizaje automático y el procesamiento de lenguaje natural, que son esenciales para el análisis bibliométrico y temático sobre medio ambiente y datos abiertos.

2.2. Aplicación de PRISMA

Para realizar el análisis bibliométrico, se siguieron las fases fundamentales de PRISMA 2020: Identificación, Selección, Evaluación de la Elegibilidad e Inclusión. A continuación, se describen los detalles específicos de cada fase, adaptados a las necesidades del estudio.

2.2.1. Identificación

Durante esta fase del estudio, se llevó a cabo una búsqueda exhaustiva de investigaciones relevantes utilizando la base de datos Scopus, seleccionada como fuente principal por su amplia cobertura de publicaciones académicas en áreas clave como ciencias, tecnología, medicina y ciencias sociales, con un enfoque particular en ciencias ambientales y datos abiertos. Scopus proporciona perfiles detallados de autores e instituciones, lo que facilita el análisis de métricas fundamentales en investigaciones de gran escala, como evaluaciones de investigación, políticas científicas y clasificaciones universitarias. La precisión en el enlace de citas y la mejora continua de sus procesos de control de calidad convierten a esta base de datos en una herramienta indispensable para los estudios bibliométricos (Baas et al., 2020).

La estrategia de búsqueda se diseñó cuidadosamente para asegurar la mayor exhaustividad posible. Se utilizaron términos clave específicamente relacionados con el tema de estudio, tales como "open data," "public data," "open information," "public information," combinados con "environmental" o "ecological," para capturar la relevancia de los estudios dentro de este marco temático. Además, se apoyó este proceso con tesauros y vocabularios controlados para identificar sinónimos y variaciones terminológicas que ampliaran la cobertura de la búsqueda. Esta metodología se complementó con el uso de operadores booleanos para refinar los resultados, lo que permitió una selección más precisa de documentos pertinentes.

La búsqueda inicial arrojó un total de 111,033 documentos, los cuales fueron filtrados según criterios adicionales para aumentar su relevancia. Se priorizó la selección de estudios de acceso abierto, lo que redujo la muestra a 48,236 documentos. Posteriormente, se restringió a artículos de investigación publicados entre 2021 y 2023, lo cual ajustó el corpus a 14,020 documentos. Dado el interés en analizar investigaciones relacionadas con la región amazónica, se incluyeron solo aquellos estudios provenientes de los países miembros de la Organización del Tratado de Cooperación Amazónica, es decir, Bolivia, Brasil, Colombia, Ecuador, Guyana, Perú, Surinam y Venezuela, lo que finalmente redujo el conjunto de documentos a 1,188.

El filtro final se aplicó para limitar los estudios al idioma inglés, resultando en un total de 1,007 artículos que cumplieran con todos los criterios establecidos. Estos documentos fueron posteriormente organizados y procesados para la eliminación de duplicados mediante herramientas especializadas, asegurando así que los subconjuntos finales de datos fueran los más precisos y manejables para el análisis bibliométrico y temático, en las siguientes etapas.

2.2.2. Selección

En la fase de selección, se aplicaron criterios de inclusión y exclusión directamente a través de la estrategia de búsqueda configurada en Scopus. Dado que el número de estudios identificados superó las 1,000 referencias, se confió plenamente en la precisión de los filtros y la cadena de búsqueda utilizada, incluyendo términos y operadores booleanos cuidadosamente diseñados para capturar únicamente estudios relevantes para el análisis del uso de datos abiertos en el contexto medioambiental.

No se realizó un cribado manual de títulos y resúmenes ni una revisión a texto completo posterior. La decisión de incluir todos los estudios que cumplieran con los parámetros de la búsqueda se fundamentó en la robustez y la especificidad de los filtros aplicados en Scopus, lo que permitió asegurar que los resultados fueran directamente relevantes para los objetivos del estudio sin necesidad de una evaluación adicional.

2.2.3. Elegibilidad

En la fase de elegibilidad, se procedió a la extracción de datos relevantes de los estudios seleccionados utilizando la librería Bibliometrix de R Studio, una herramienta especializada para el análisis bibliométrico. Se extrajeron las siguientes variables para cada estudio:

- Información bibliográfica: Título, autores, año de publicación, revista y DOI.
- Afiliación: Institución y país de los autores.
- Palabras clave: Términos proporcionados por los autores para describir el contenido del estudio.

- Tipo de publicación: Clasificación del documento como artículo de investigación, revisión o capítulo de libro.
- Citas recibidas: Número total de citas que ha recibido cada artículo.
- Fuentes de datos abiertas: Identificación de las plataformas, bases de datos y repositorios de datos abiertos utilizados en los estudios, registrando la URL o identificador de cada fuente cuando fue posible.

El proceso se realizó de manera sistemática para garantizar que los datos extraídos fueran consistentes y precisos. Aunque inicialmente se consideró una revisión exhaustiva de los textos completos, en esta fase se centró exclusivamente en la extracción de información relevante para el análisis bibliométrico. Aquellos estudios que no cumplían con los criterios específicos de elegibilidad, como la falta de claridad en la aplicación de datos abiertos en el contexto ambiental, fueron excluidos y se documentaron las razones para su exclusión, siguiendo las directrices establecidas por PRISMA.

2.2.4. Inclusión

En la fase de inclusión, se llevó a cabo un análisis exhaustivo de los datos extraídos de los estudios seleccionados, utilizando herramientas avanzadas y técnicas de análisis tanto bibliométrico como de machine learning. Este proceso se centró en identificar patrones, tendencias y relaciones dentro de la investigación sobre medio ambiente y datos abiertos, así como en interpretar y contextualizar estos hallazgos en el campo de estudio.

Análisis bibliométrico: Se realizó un análisis cuantitativo detallado para caracterizar la producción científica en el área de estudio. Este análisis incluyó:

- **Análisis descriptivo:** Se evaluó la producción científica a lo largo del tiempo, identificando los autores, instituciones y países más productivos, así como las revistas más relevantes en el campo. Los resultados se presentaron mediante tablas, gráficos y mapas para facilitar la visualización de patrones y tendencias.
- **Leyes bibliométricas:** Se aplicaron las leyes de Bradford y Lotka para analizar la dispersión de la literatura y la productividad de los autores, lo

que permitió interpretar la concentración de la investigación en revistas y el desempeño de los investigadores más destacados en este ámbito.

- **Análisis de redes:** Utilizando VOSviewer, se visualizaron las redes de colaboración entre autores, países e instituciones, destacando los actores principales y sus relaciones. Este análisis abarcó diversas redes, como co-autoría, co-citación, co-ocurrencia y redes de texto. Para complementar el estudio, se empleó el paquete iGraph de R Studio, lo que permitió realizar un análisis estadístico detallado de las métricas de red, mapeando las conexiones y evaluando la influencia dentro de la comunidad científica.

Análisis de datos avanzados y machine learning: Se aplicaron técnicas de aprendizaje automático y procesamiento de lenguaje natural para profundizar en la interpretación temática de los artículos seleccionados:

- **Procesamiento del Lenguaje Natural (PLN):** Se analizó el contenido textual de los artículos para identificar conceptos clave y explorar las relaciones entre ellos mediante técnicas de PLN, como el modelado de tópicos.
- **Latent Dirichlet Allocation (LDA):** Se utilizó para descubrir temas subyacentes en la literatura, permitiendo una comprensión más profunda de las áreas de interés predominantes en la investigación sobre medio ambiente y datos abiertos.
- **k-means clustering:** Se aplicó para agrupar los estudios en clústeres temáticos basados en su similitud de contenido, clasificando las diferentes líneas de investigación presentes en el corpus analizado.
- **t-SNE (t-distributed Stochastic Neighbor Embedding):** Esta técnica de reducción dimensional se empleó para visualizar la distribución de los clústeres en un espacio bidimensional, facilitando la interpretación y representación gráfica de las relaciones entre los temas identificados.

Evaluación de las fuentes de acceso abierto: Se realizó una revisión cualitativa de un subconjunto de 50 artículos para identificar las fuentes de datos abiertas más utilizadas en los estudios seleccionados. Este análisis permitió identificar

las plataformas, bases de datos y repositorios más comunes, evaluando sus características en términos de tipo de datos, licencias y accesibilidad, así como su impacto en la disseminación del conocimiento en el ámbito medioambiental.

Finalmente, se interpretaron los resultados obtenidos del análisis bibliométrico y del machine learning en el contexto del campo de estudio, discutiendo las implicaciones de los hallazgos para futuras investigaciones, la toma de decisiones y el desarrollo de políticas relacionadas con el medio ambiente y el uso de datos abiertos.

2.3. Limitaciones

El estudio presenta ciertas limitaciones metodológicas que deben considerarse al interpretar los resultados. Una de las principales limitaciones es la dependencia exclusiva de la base de datos Scopus para la identificación de estudios, lo que pudo haber generado una subrepresentación de investigaciones relevantes que no están indexadas en esta plataforma, ya que no abarca la totalidad de las publicaciones científicas a nivel global. Además, el proceso de selección se basó únicamente en la aplicación de filtros y la cadena de búsqueda, sin una revisión manual a texto completo, lo cual podría haber afectado la especificidad y precisión en la identificación de estudios pertinentes.

Otra limitación metodológica es el sesgo de idioma, ya que la búsqueda se restringió a publicaciones en inglés y español, excluyendo potencialmente investigaciones valiosas en otros idiomas. También, el uso de términos predefinidos para la búsqueda pudo haber limitado la captación de estudios relevantes debido a las variaciones en la terminología empleada por diferentes autores.

Las técnicas de machine learning y visualización de datos, como LDA, k-means clustering y t-SNE, implican la selección de parámetros específicos que pueden influir en los resultados, introduciendo cierta subjetividad en la interpretación de las temáticas identificadas. Finalmente, aunque se evaluaron las fuentes de datos abiertas utilizadas, las conclusiones están limitadas por la calidad y la disponibilidad variable de estos datos en los repositorios seleccionados.

Estas limitaciones metodológicas subrayan la necesidad de futuros estudios que incorporen un enfoque más amplio, utilizando múltiples bases de datos, revisiones manuales y estrategias de búsqueda más flexibles para mejorar la robustez y la representatividad de los hallazgos.

CAPITULO 3

ANÁLISIS BIBLIOMÉTRICO



Análisis bibliométrico

3.1. Descripción de la producción científica

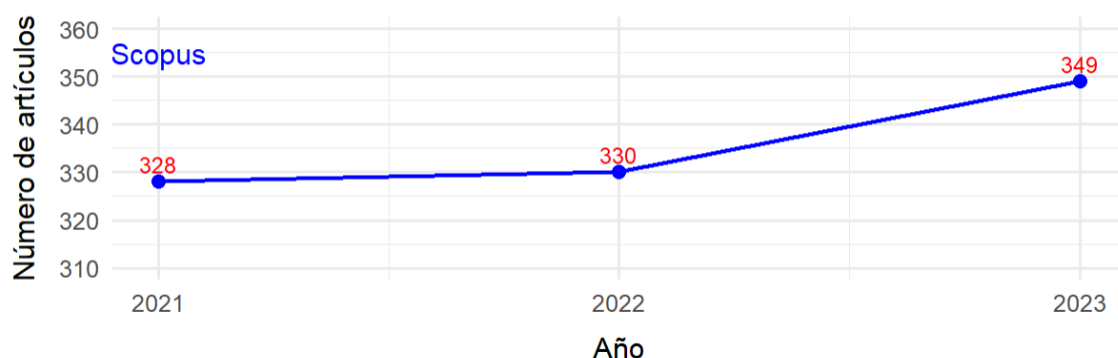
3.1.1. Periodo de análisis

La investigación ambiental actual es dinámica y requiere mantenerse al día con los avances científicos y tecnológicos. La abundancia de datos y el acceso abierto a bases de datos permiten identificar líneas emergentes, evaluar políticas y detectar brechas de conocimiento. En este sentido, centrarse en el análisis de datos recientes facilita una mejor comprensión de las tendencias y la toma de decisiones informadas para abordar los desafíos ambientales (Garrigós-Simón et al., 2021). El enfoque en los últimos tres años permite acceder a la información más relevante y actualizada, crucial para abordar los desafíos ambientales globales.

El periodo de análisis abarca desde 2021 hasta 2023, con un total de 1,000 documentos indexados en Scopus. Durante esos tres años, se observa un crecimiento anual del 3.15%, con una ligera variación en la producción de artículos: 328 en 2021, 330 en 2022 y 349 en 2023, ver Figura 1. El aumento en 2023 sugiere una tendencia moderada al alza en la producción científica, aunque el crecimiento no es particularmente acelerado, lo que podría indicar una estabilización en el ritmo de publicación en este campo.

Figura 1

Producción científica reciente, último trienio



Nota: Autores (2024).

La antigüedad promedio de los documentos examinados en este estudio es de 1.98 años. Esta métrica revela que el corpus documental analizado está constituido por publicaciones recientes, concentradas predominantemente en un intervalo temporal de tres años. Dicha característica temporal del conjunto de documentos sugiere que el análisis se sustenta en información actualizada y, por lo tanto, refleja con mayor precisión el estado actual del conocimiento en el área temática investigada.

3.1.2. Mapeo de la contribución por país

Previo a la evaluación de la contribución por país es necesario señalar que el análisis del impacto de las citas, considerando todas las publicaciones evaluadas, revela un promedio de 7,554 citas por documento y una tasa de 2,414 citas por año por documento. Si bien estos valores indican una notable repercusión de los trabajos dentro del campo de estudio, se podría señalar que los niveles de influencia son moderados, y esto puede atribuirse a la corta ventana temporal de análisis, dado que los artículos más recientes aún se encuentran en proceso de difusión y asimilación por parte de la comunidad científica. Cabe esperar que, con el transcurso del tiempo, estos trabajos incrementen su visibilidad y, consecuentemente, su número de citas.

Los resultados del procesamiento de los datos, revelan que Brasil destaca como el país con mayor número de citaciones en la región amazónica, alcanzando un total de 2765 citaciones. Este resultado refleja un impacto significativo en la producción científica de ese país, con un promedio de 4.83 citaciones por artículo. En comparación, Colombia, con 386 citaciones y un promedio de 6.23 por artículo, muestra también una contribución relevante al ámbito científico, aunque en menor escala que Brasil. En la Tabla 1, se observa que Ecuador ocupa la sexta posición en cuanto al número total de citaciones, con un valor de 264. Esto indica que los artículos científicos producidos por investigadores ecuatorianos han sido citados un número considerable de veces en comparación con otros países de la lista, lo cual sugiere una contribución científica relevante a nivel internacional.

El análisis también destaca la participación de otros países fuera de la región amazónica en esta búsqueda. Estados Unidos, por ejemplo, ha registrado 1253

citaciones, con un promedio notablemente alto de 28.48 citas por artículo, lo que resalta la influencia que sus investigaciones tienen en la comunidad científica global. Australia, con 512 citas y un promedio aún más elevado de 56.89 por artículo, sobresale como un actor clave, a pesar de estar geográficamente distante de la Amazonía. Estos resultados sugieren que, si bien los países amazónicos tienen un papel importante en las investigaciones sobre temas ambientales, la contribución de países no amazónicos en términos de citas y difusión científica es también destacada.

Tabla 1

Países con mayores citas

País	Citas	Citas promedio de artículos
BRAZIL	2765	4.83
USA	1253	28.48
AUSTRALIA	512	56.89
COLOMBIA	386	6.23
GERMANY	356	23.73
ECUADOR	264	7.76
UNITED KINGDOM	193	7.72
SPAIN	146	8.11
PERU	134	6.70
SWEDEN	80	16.00

Nota: Autores (2024).

En cuanto a la producción científica total y las colaboraciones internacionales, se puede observar en la Tabla 2 que Brasil vuelve a posicionarse como líder, con 573 artículos publicados entre 2021 y 2023. De estas publicaciones, 427 corresponden a artículos con coautores del mismo país, mientras que 146 han sido escritos en colaboración con autores internacionales, lo que representa un 25.5% de publicaciones en colaboración internacional. Esta cifra pone de relieve el dinamismo de la comunidad científica brasileña y su apertura a la cooperación con investigadores de otras partes del mundo.

Tabla 2

Países de los autores con mayor cantidad de artículos

País	Artículos	Frecuencia	Índice de colaboración intra-país	Índice de colaboración inter-país	Relación Inter-país
BRAZIL	573	0.63175	427	146	0.255

Capítulo III: Análisis bibliométrico

COLOMBIA	62	0.06836	36	26	0.419
USA	44	0.04851	0	44	1.000
ECUADOR	34	0.03749	8	26	0.765
UNITED KINGDOM	25	0.02756	0	25	1.000
PERU	20	0.02205	11	9	0.450
SPAIN	18	0.01985	1	17	0.944
GERMANY	15	0.01654	0	15	1.000
PORTUGAL	12	0.01323	1	11	0.917
AUSTRALIA	9	0.00992	0	9	1.000

Nota: Autores (2024).

Colombia, con 62 artículos publicados, presenta un patrón similar, aunque con una mayor proporción de colaboraciones internacionales. Un 41.9% de sus publicaciones ha sido producto de colaboraciones con autores de otros países, lo que refleja una estrategia de cooperación científica más pronunciada en relación con el tamaño de su producción. Ecuador, por su parte, ha publicado 34 artículos, de los cuales solo 8 fueron realizados exclusivamente por autores nacionales, mientras que 26 han sido escritos en colaboración internacional. Esto representa un alto porcentaje de cooperación internacional, concretamente un 76.5%, lo que evidencia la creciente importancia de las colaboraciones globales en la ciencia ecuatoriana.

Es importante señalar que países como Estados Unidos y el Reino Unido también tienen una presencia significativa en este corpus de artículos. En ambos casos, todas sus publicaciones durante este período han sido realizadas en colaboración internacional, con un ratio de colaboraciones múltiples (MCP Ratio) de 1.000. Esto demuestra que estos países priorizan la cooperación científica internacional cuando investigan temas relacionados con el medio ambiente y la ecología en el contexto amazónico.

Estos resultados ponen en evidencia tanto la importancia de la región amazónica en el contexto de la ciencia global, como la relevancia de las colaboraciones internacionales para potenciar la producción y el impacto de la investigación científica. Las diferencias en el número de artículos y citas entre los países reflejan no solo la variabilidad en los recursos y capacidades de investigación de cada nación, sino también su integración en redes de colaboración científica más amplias.

3.1.3. Panorama institucional de las publicaciones

El análisis de la Tabla 3, derivada de una búsqueda bibliográfica, revela un panorama de la producción científica en el ámbito institucional. La destacada presencia de la Universidad de São Paulo (USP), aportando el 2.65% de los artículos, subraya su liderazgo en la investigación relacionada con datos abiertos y medio ambiente en la Amazonía.

La participación significativa de diversas universidades federales brasileñas evidencia el compromiso de estas instituciones con la generación de conocimiento en temas ambientales clave para la región. La diversidad geográfica de estas instituciones, abarcando desde la región amazónica hasta el sur de Brasil, sugiere un interés extendido en la temática.

Es importante destacar que este análisis se basa en una selección de 1,000 artículos en inglés, publicados entre 2021 y 2023. Esta selección representa una fracción de la producción científica total sobre el tema, lo que limita la generalización de los resultados. No obstante, el análisis ofrece una visión valiosa sobre las principales instituciones que investigan datos abiertos y medio ambiente en la Amazonía, destacando la contribución de la USP y el rol crucial de las universidades federales brasileñas.

Tabla 3

Instituciones más relevantes según filiación de los autores y sus artículos

Institución	Artículos	Porcentaje
University Of São Paulo	179	2.65%
Universidade Federal De Minas Gerais	75	1.11%
Universidade De São Paulo	71	1.05%
Federal University Of Maranhão	68	1.01%
Federal University Of Rio Grande Do Norte	57	0.85%
Federal University Of Minas Gerais	50	0.74%
Universidade Federal Do Rio De Janeiro	49	0.73%
Universidade Federal Do Pará	45	0.67%
Universidade Federal Da Bahia	38	0.56%
Federal University Of Goiás	36	0.53%

Nota: Autores (2024).

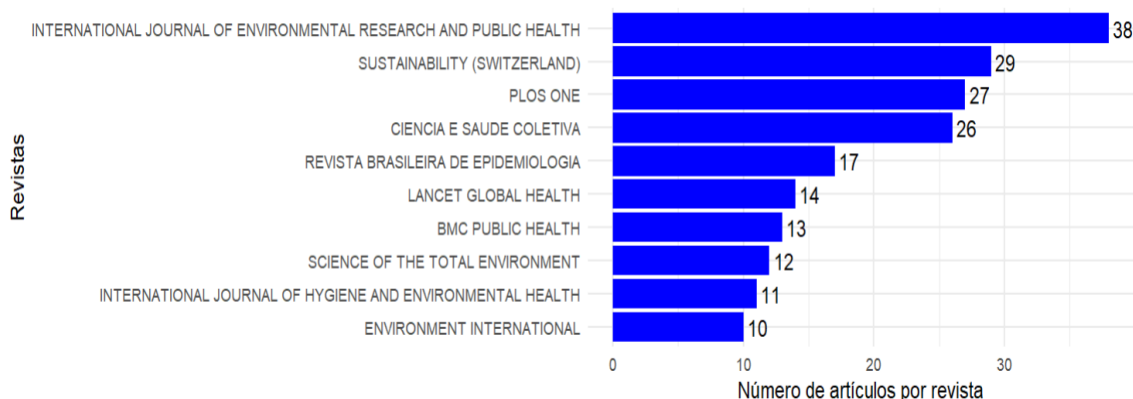
3.1.4. Revistas claves en la producción científica

El análisis de los datos bibliométricos dirigido hacia las fuentes, abarca un total de 485 revistas, lo cual revela una amplia gama de productividad entre las publicaciones consideradas. Mientras que algunas concentran un número significativo de artículos, la mayoría registra contribuciones más modestas. Esta distribución sugiere una diversidad en el ecosistema de publicaciones, donde coexisten revistas de alto impacto con otras de alcance más especializado.

El gráfico de barras horizontales de la Figura 2, representa las 10 revistas más destacadas visualmente confirma estas conclusiones. La prominencia del "International Journal of Environmental Research and Public Health" es evidente, mientras que las demás revistas, aunque relevantes, muestran una contribución decreciente en términos de número de artículos. Esta representación gráfica facilita la comprensión de la jerarquía de publicaciones y su papel en la construcción del conocimiento en el área de estudio.

Figura 2

Las 10 fuentes más prolíficas en la bibliometría realizada



Nota: Autores (2024).

El "International Journal of Environmental Research and Public Health" se destaca como la principal fuente de publicación, albergando un número considerablemente mayor de artículos que cualquier otra revista en el conjunto de datos. Le siguen "Sustainability (Switzerland)" y "PLOS ONE", consolidándose como referentes importantes en la difusión de investigaciones relacionadas con los temas de estudio.

La distribución de frecuencias calculada en base a las 485 revistas y los artículos publicados, con una media cercana a 0.0021 y un máximo de 0.0377, indica que la mayoría de las revistas tienen una presencia relativamente baja en relación con el total de artículos analizados. Esta observación refuerza la idea de un panorama editorial diverso, donde numerosas publicaciones contribuyen al conocimiento en el campo, aunque con un peso relativo menor en comparación con las revistas líderes.

3.1.5. Distribución de los artículos según Ley de Bradford

La Ley de Bradford es un principio fundamental en bibliometría y describe la distribución de artículos científicos en revistas dentro de un campo específico. Esta ley establece que un pequeño número de revistas principales ("núcleo" o "Zona 1") produce la mayor parte de los artículos, mientras que un gran número de revistas periféricas contribuye con el resto.

El análisis de los datos revela una clara manifestación de la Ley de Bradford. De un total de 1007 artículos, 334 (aproximadamente el 33%) se concentran en las 29 revistas principales que componen la zona 1, también llamada "Zona Núcleo", (Ver Tabla 4).

Entre estas revistas destacadas se encuentran el "International Journal of Environmental Research and Public Health", "Sustainability (Switzerland)" y "PLOS ONE", las cuales contribuyen significativamente a la producción científica en este campo, ver Tabla 5. En este contexto, la "posición" se refiere al lugar que ocupa una revista en una lista ordenada según su productividad de artículos, siendo la revista con más artículos la que ocupa la primera posición.

Tabla 4

División de revistas según las zonas de la ley de Bradford

Zonas	Revistas	Porcentaje revistas	Artículos	Porcentaje artículos	Promedio artículos por revista
1	29	5.9794	334	33.16782522	11.51
2	124	25.5670	341	33.86295929	2.75
3	332	68.4536	332	32.96921549	1
TOTAL	485	100	1007	100	-

Nota: Autores (2024).

Tabla 5

Revistas pertenecientes a la Zona 1 según ley de Bradford

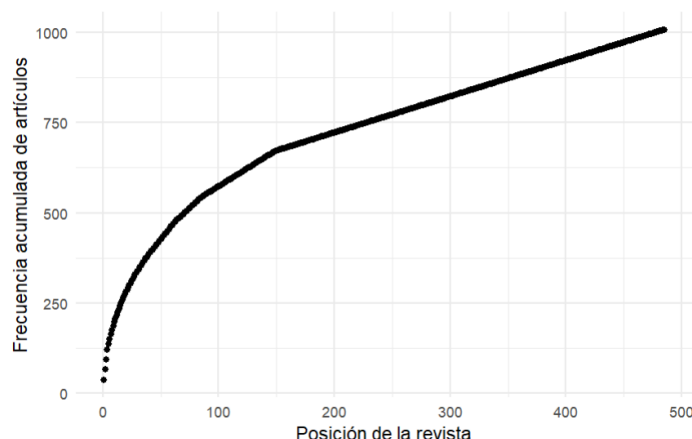
Revistas	P	F	FA	PF	PFA
INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH	1	38	38	11.38	11.38
SUSTAINABILITY (SWITZERLAND)	2	29	67	8.68	20.06
PLOS ONE	3	27	94	8.08	28.14
CIENCIA E SAUDE COLETIVA	4	26	120	7.78	35.93
REVISTA BRASILEIRA DE EPIDEMIOLOGIA	5	17	137	5.09	41.02
BMC PUBLIC HEALTH	6	14	151	4.19	45.21
PLOS NEGLECTED TROPICAL DISEASES	7	13	164	3.89	49.10
SCIENCE OF THE TOTAL ENVIRONMENT	8	12	176	3.59	52.69
BMJ OPEN	9	11	187	3.29	55.99
FRONTIERS IN PUBLIC HEALTH	10	10	197	2.99	58.98
REMOTE SENSING	11	10	207	2.99	61.98
REVISTA DE GESTAO SOCIAL E AMBIENTAL	12	10	217	2.99	64.97
DESENVOLVIMENTO E MEIO AMBIENTE	13	9	226	2.69	67.66
SCIENTIFIC REPORTS	14	9	235	2.69	70.36
AMBIENTE E SOCIEDADE	15	8	243	2.40	72.75
HELIYON	16	8	251	2.40	75.15
MUNDO DA SAUDE	17	8	259	2.40	77.54
REVISTA DA SOCIEDADE BRASILEIRA DE MEDICINA TROPICAL	18	8	267	2.40	79.94
CADERNOS DE SAUDE PUBLICA	19	7	274	2.10	82.04
JOURNAL OF HUMAN GROWTH AND DEVELOPMENT	20	7	281	2.10	84.13
BRAZILIAN ORAL RESEARCH	21	6	287	1.80	85.93
EPIDEMIOLOGIA E SERVICOS DE SAUDE	22	6	293	1.80	87.72
JOURNAL OF INFECTION IN DEVELOPING COUNTRIES	23	6	299	1.80	89.52
REVISTA BRASILEIRA DE SAUDE MATERNO INFANTIL	24	6	305	1.80	91.32
REVISTA DA ESCOLA DE ENFERMAGEM	25	6	311	1.80	93.11
REVISTA DE SAUDE PUBLICA	26	6	317	1.80	94.91
THE LANCET REGIONAL HEALTH - AMERICAS	27	6	323	1.80	96.71
TROPICAL MEDICINE AND INFECTIOUS DISEASE	28	6	329	1.80	98.50
ENVIRONMENTAL RESEARCH LETTERS	29	5	334	1.50	100.00

Nota: P: Posición, F: Frecuencia, FA: Frecuencia acumulada, PF: Porcentaje de frecuencia, PFA: Porcentaje de frecuencia acumulada, Autores (2024)

La gráfica de dispersión en la Figura 3 ilustra visualmente la distribución de Bradford. La curva ascendente pronunciada al principio refleja la rápida acumulación de artículos en las revistas que ocupan las primeras posiciones. A medida que avanzamos hacia las revistas de menor posición, la curva se aplana, lo que indica una disminución en la productividad de artículos por revista. La división en tres zonas (núcleo, zona 2 y zona 3) se basa en puntos de inflexión observados en esta curva, reflejando cambios en la tasa de acumulación de artículos.

Figura 3

Frecuencia acumulada de artículos por posición de revista



Nota: Autores (2024).

El análisis de Bradford puede revelar patrones y tendencias en la comunicación académica. Por ejemplo, un núcleo muy pequeño en comparación con una zona extendida grande podría sugerir una concentración de la investigación en unas pocas instituciones o revistas de élite. Por el contrario, un núcleo más grande podría indicar una mayor diversidad en la producción de conocimiento. En el contexto de los datos estudiados, el análisis de Bradford revela un núcleo relativamente pequeño en comparación con la zona extendida. Aproximadamente el 33% de los artículos se concentran en solo el 6% de las revistas, lo que sugiere una moderada concentración de la producción de conocimiento en un grupo selecto de publicaciones. Si bien esto no indica una situación extrema de dominio por unas pocas revistas de élite, sí apunta hacia una cierta centralización en la comunicación académica dentro del campo estudiado.

3.1.6. Análisis a través de la Ley de Lotka

La Ley de Lotka es una herramienta fundamental en bibliometría y ofrece una perspectiva valiosa sobre la distribución de la productividad científica entre los investigadores (Kawamura et al., 2000). Esta ley establece que el número de autores con un determinado número de publicaciones disminuye a medida que aumenta el número de publicaciones, reflejando una realidad común en la investigación: un pequeño grupo de investigadores altamente productivos genera una parte sustancial de la producción científica total, mientras que la

mayoría de los autores contribuyen con un número relativamente bajo de publicaciones.

El análisis de los datos, basado en la Ley de Lotka, revela una distribución de la productividad científica que se ajusta a este patrón. La Tabla 6 presenta la distribución observada de la productividad en comparación con la distribución teórica derivada de la ley, utilizando un coeficiente beta de 2.4012 y una constante C de 0.8873. El coeficiente beta, al ser mayor que 2, indica una concentración de la productividad en un número reducido de autores, lo que se refleja en la rápida disminución del número de autores a medida que aumenta su número de publicaciones.

Tabla 6

Distribución observada y teórica de la productividad científica

Número de artículos	Número de autores	Frecuencia (Distribución Observada)	Distribución Teórica	Apariciones de autor
1	7202	0.927495	0.8873	7202
2	467	0.0601417	0.168	934
3	68	0.00875724	0.0634	204
4	13	0.00167418	0.0318	52
5	8	0.00103026	0.0186	40
6	4	0.000515132	0.012	24
7	1	0.000128783	0.0083	7
11	1	0.000128783	0.0028	11
16	1	0.000128783	0.0011	16
TOTALES	7765	1	-	8490

Nota: Autores (2024).

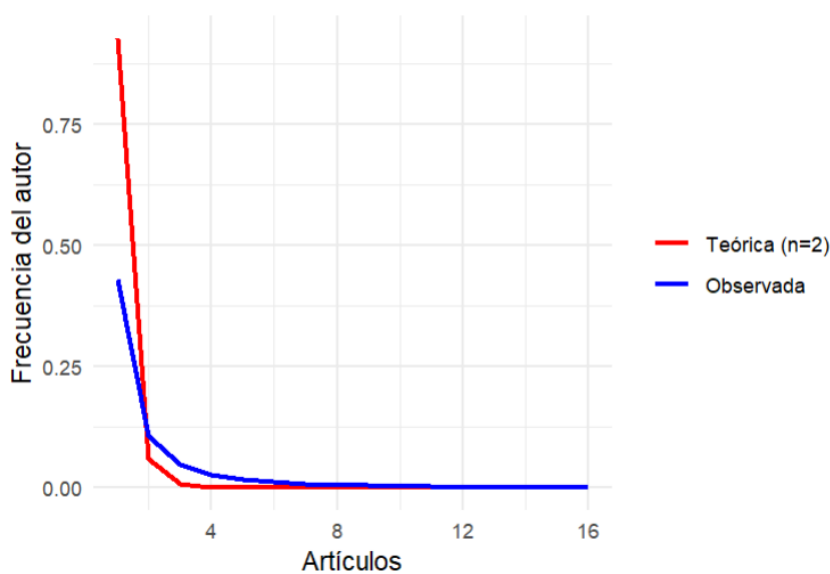
La excelente bondad de ajuste del modelo, evidenciada por un R^2 de 0.9821 y un valor p de 0.9545 en la prueba de Kolmogorov-Smirnov, confirma que la Ley de Lotka describe con notable precisión la distribución de la productividad científica observada en los datos.

La Figura 4 ilustra la distribución observada de la frecuencia de autores en relación con el número de artículos publicados, comparándola con una distribución teórica basada en la Ley de Lotka con un exponente $n=2$. Se evidencia una clara concordancia entre ambas distribuciones, especialmente en el rango de menor número de artículos, donde la mayoría de los autores se concentran. Esto ratifica que la Ley de Lotka, con el exponente ajustado, modela

adecuadamente el patrón de productividad científica observado en este conjunto de datos, confirmando la tendencia a que un pequeño número de autores publique la mayoría de los artículos, mientras que la gran mayoría contribuye con un número reducido de publicaciones.

Figura 4

Distribución de autores y productividad científica según Ley de Lotka



Nota: Autores (2024).

3.2. Análisis de redes

3.2.1. Mapeo de redes basado en VOSviewer

En esta investigación bibliométrica sobre datos abiertos relacionados con el medio ambiente, se ha utilizado el software VOSviewer para llevar a cabo un análisis detallado de redes. Este análisis se centra en la exploración y visualización de las interconexiones entre diversos componentes clave en este campo. VOSviewer aplica un enfoque basado en la distancia para generar mapas de redes que representan gráficamente las relaciones entre entidades como autores, instituciones, países y palabras clave, todo en el contexto de la aplicación de los datos abiertos al ámbito medioambiental (Anas et al., 2024).

Cada uno de estos elementos se visualiza como un círculo dentro de la red, mientras que las líneas que los conectan representan diferentes tipos de

relaciones: coautoría (número de publicaciones compartidas entre dos investigadores), co-citación (frecuencia con la que dos documentos o autores son citados conjuntamente en otros trabajos) y co-ocurrencia (número de publicaciones en las que dos términos aparecen juntos).

Estas conexiones son esenciales para comprender cómo interactúan y se relacionan los diferentes actores y conceptos dentro del campo, proporcionando una visión clara de las dinámicas que configuran este ámbito de estudio (van Eck & Waltman, 2020).

La evaluación de la solidez de las conexiones en el análisis de redes se lleva a cabo a través de la asignación de pesos o valores numéricos positivos, los cuales indican la intensidad de las relaciones entre los nodos. Los valores más elevados denotan vínculos más fuertes, lo que refleja una mayor frecuencia de interacciones, como la coautoría o citas compartidas entre elementos relacionados, por ejemplo, en el contexto de publicaciones sobre datos abiertos ambientales (Van Eck & Waltman, 2014).

Evaluar la intensidad de las conexiones es esencial para identificar patrones y analizar la formación de clústeres, lo que permite una comprensión más profunda de las dinámicas subyacentes en el análisis bibliométrico. En este contexto, las visualizaciones producidas por VOSviewer emplean una combinación de colores y tamaños para destacar la relevancia de los elementos en el análisis bibliométrico. Los círculos de mayor tamaño indican una mayor importancia o un mayor número de citas recibidas, mientras que la proximidad entre los nodos representa la solidez de sus conexiones (van Eck & Waltman, 2020).

El algoritmo de clustering implementado por VOSviewer optimiza la organización de los elementos en grupos, permitiendo la identificación de subconjuntos más específicos y homogéneos en el contexto de los datos abiertos medioambientales. Este proceso agrupa elementos con características similares, facilitando un análisis más detallado de las tendencias y conexiones presentes en el conjunto de datos. Si bien alcanzar una homogeneidad total en los clústeres puede ser desafiante debido a la diversidad inherente de la información, el objetivo principal es maximizar la consistencia dentro de cada grupo, empleando

un algoritmo de movimiento local diseñado específicamente para esta tarea (Van Eck & Waltman, 2014).

En consecuencia, el análisis de redes mediante VOSviewer en esta investigación bibliométrica sobre datos abiertos medioambientales ofrece una comprensión profunda de la estructura y dinámica de este campo. Al visualizar las interrelaciones y agrupamientos de actores clave, instituciones y conceptos, se facilita la identificación de tendencias emergentes y áreas de investigación relevantes, proporcionando una base sólida para futuros estudios y avances en la aplicación de datos abiertos en la ciencia medioambiental (Li et al., 2021).

3.2.2. Evaluación de métricas cuantitativas de red

Adicionalmente, para complementar el análisis de redes, se llevó a cabo un análisis estadístico utilizando métricas cuantitativas que describen tanto las características globales de la red como las de sus nodos individuales. Estas métricas proporcionan una visión general de la estructura de la red y de sus propiedades fundamentales. La elección de la métrica adecuada depende del contexto específico y de los objetivos del estudio (Fonseca et al., 2016).

El análisis estadístico se llevó a cabo utilizando el software R Studio, empleando el paquete iGraph para calcular los parámetros de la red. Se obtuvieron principalmente, medidas de centralidad, como el grado, la intermediación y la cercanía, lo que permitió una comprensión más detallada de la red de relaciones en el ámbito de los datos abiertos medioambientales (Saetnan & Kipling, 2016).

Según Zhang (2010) la centralidad es una medida clave en el análisis de redes sociales, utilizada para identificar a los actores más prominentes dentro de una red, aquellos con mayor participación o relevancia debido a sus relaciones con otros miembros. En su estudio, Zhang propone varios tipos principales de centralidad:

- **Centralidad de grado:** Se refiere al número de conexiones directas que tiene un nodo (actor). Un mayor número de conexiones indica que el actor es más activo o popular dentro de la red. Esta métrica se calcula contando el número de relaciones de cada nodo y refleja tanto las conexiones entrantes como salientes (Osareh et al., 2022).

- Centralidad de intermediación: Mide la frecuencia con la que un nodo actúa como intermediario o puente entre otros nodos que no están directamente conectados. Un nodo con alta intermediación tiene la capacidad de controlar el flujo de información o recursos, lo que subraya su importancia en términos de su posición estratégica para la transmisión de información en la red (Osareh et al., 2022).
- Centralidad de cercanía: Esta medida se basa en la proximidad de un nodo a todos los demás en la red. Un nodo con alta cercanía tiene acceso eficiente a todos los demás actores, lo que le permite una mayor independencia y rapidez en la transmisión de información (Zhang, 2010).
- Centralidad de autovector: Es una medida que no solo toma en cuenta el número de conexiones directas de un nodo, como ocurre con la centralidad de grado, sino también la relevancia de los nodos a los que está conectado. Un nodo tiene un puntaje alto de centralidad de autovector si está conectado a otros nodos igualmente influyentes dentro de la red. Esta métrica valora tanto la cantidad como la calidad de las conexiones, asignando mayor importancia a las interacciones con nodos prominentes (Zhang, 2010).

3.2.3. Análisis por co-autoría

El análisis de co-autoría se basa en la creación de representaciones visuales que ilustran las relaciones y colaboraciones entre países, organizaciones y autores en la investigación académica. Este enfoque permite comprender y evaluar los patrones de colaboración científica, facilitando el análisis de las dinámicas de cooperación dentro de la comunidad investigadora (Fonseca et al., 2016) y las posiciones individuales de los investigadores (Osareh et al., 2022). Para el estudio de cada red, se ha utilizado el método de "conteo completo" en VOSviewer (San-Juan-Heras et al., 2024).

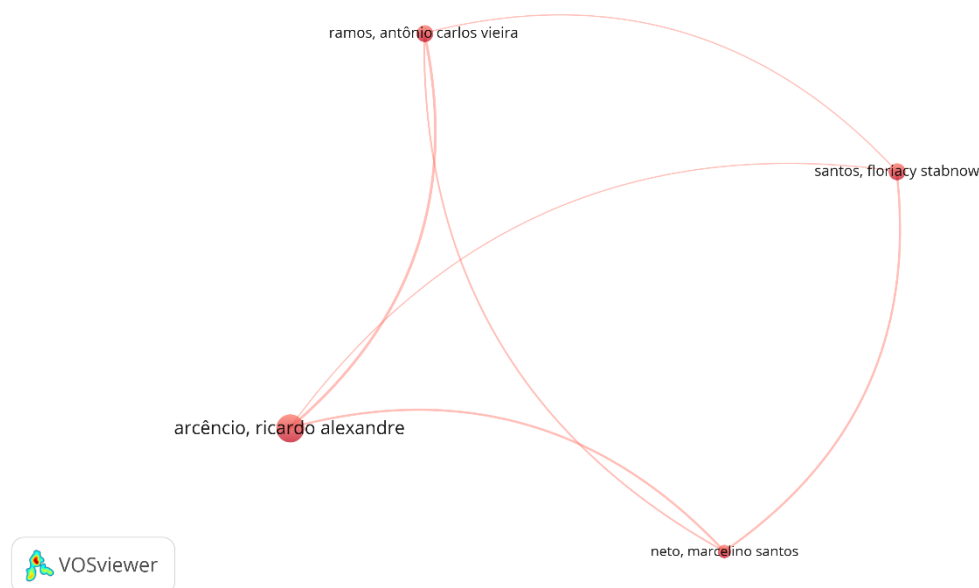
3.2.3.1. Red de co-autoría por autor

3.2.3.1.1. Mapeo de red de co-autoría por autor

La primera red corresponde a la unidad de análisis basada en los autores. El software VOSviewer facilita el análisis de co-autoría evaluando las interacciones entre los autores, mostrando las redes de colaboración presentes en la producción científica (Omar & Kamaruzaman, 2024). Se utilizó el tesauro propio del programa para optimizar la identificación y clasificación de términos. Los criterios de selección establecidos incluyeron un mínimo de cinco publicaciones por autor y al menos una cita por autor.

La aplicación de estos criterios permitió identificar a los autores más influyentes en la investigación sobre datos abiertos medioambientales a partir de documentos extraídos de Scopus, involucrando un total de 5,617 autores. La Figura 5 facilita el reconocimiento de los grupos de colaboración y de los nodos clave que desempeñan un papel central en la producción y difusión del conocimiento en el campo estudiado. Se identificaron 9 autores, de los cuales se analizaron 4, lo que representa un 0.07% del total, debido a su alto grado de relación.

Los resultados del análisis de coautoría por autor revelan patrones de colaboración y conexiones entre los diferentes investigadores. Esta visualización facilita la identificación clara de las relaciones y el grado de interacción entre los autores, destacando a aquellos con mayor influencia y visibilidad en la red de coautoría. Es importante resaltar que, el autor Arcêncio Ricardo Alexander, seguido por Neto Marcelino Santos, se destacan como los colaboradores más prominentes, ambos con un valor de 11 en la fuerza total de sus conexiones.

Figura 5*Visualización de red co-autoría por autor**Nota:* Autores (2024).

3.2.3.1.2. Medidas adicionales del análisis de co-autoría por autor

El análisis de coautoría se complementa con la evaluación de medidas adicionales, presentadas en la Tabla 7, como el grado de centralidad, que mide el número de conexiones directas que tiene cada nodo en la red. La centralidad de una red es una medida de la cohesión de su estructura. En el contexto de redes de coautoría, esto se refleja en la variabilidad del grado de centralidad de los nodos, es decir, en las diferencias en la cantidad de conexiones de coautoría que poseen los distintos autores dentro de la red.

En este sentido, la columna GC de la Tabla 7 muestra que todos los autores (Arcêncio, Neto, Ramos y Santos) presentan un grado de centralidad de 3, lo que significa que cada uno está conectado a tres otros autores. Esto sugiere una interconexión sólida entre estos investigadores, lo que podría potenciar la colaboración y el intercambio de ideas dentro de esta red de coautoría. Según Borgatti & Everett (2006), un alto grado de centralidad puede ser indicativo de una posición privilegiada en redes de colaboración que promueven la innovación y el desarrollo de nuevos conocimientos.

De manera similar, la columna GI de la Tabla 7 muestra que Arcêncio, Neto y Ramos tienen un grado de intermediación de 0, lo que significa que no actúan como intermediarios en las conexiones entre otros autores. En contraste, Santos presenta un grado de intermediación de 1, lo que indica que podría desempeñar un papel en la conexión entre otros dos autores. Este fenómeno se alinea con las observaciones de Freeman (1977), que sugieren que los nodos con un alto grado de intermediación son cruciales para la comunicación dentro de redes académicas.

Otra medida importante en el análisis de coautoría es la centralidad de intermediación. La columna CI de la Tabla 7 revela que Arcêncio, Neto y Ramos tienen una centralidad de intermediación de 0, mientras que Santos presenta un valor de 1. Esto refuerza la idea de que Santos desempeña un papel más relevante en la facilitación de conexiones entre otros autores, lo que podría ser beneficioso para fortalecer su red de investigación. Este hallazgo está respaldado por la investigación de Burt (2004), que señala que los nodos con alta centralidad de intermediación tienden a tener un impacto mayor en el flujo de información y recursos.

Por otro lado, la centralidad de cercanía mide cuán próximo está un nodo a los demás nodos de la red. Un valor más alto indica que un autor puede conectarse con otros nodos de manera más rápida y eficiente. En la columna CC de la Tabla 7, Santos presenta la mayor centralidad de cercanía (0.125), seguido por Ramos (0.111), lo que sugiere que estos autores son más eficientes en términos de comunicación o intercambio de información. En contraste, Arcêncio y Neto tienen valores más bajos, lo que indica que podrían estar más alejados en cuanto a la interacción dentro de la red. Esta idea se refuerza con el trabajo de Valente (2011), que subraya la importancia de la cercanía en redes sociales y su influencia en la difusión de información.

La centralidad de autovector mide la influencia de un nodo en la red, considerando la relevancia de las conexiones que tiene. En la columna CA de la Tabla 7, se observa que Arcêncio tiene la mayor centralidad de autovector (1.0000000), lo que sugiere que, aunque no actúe como intermediario, sus conexiones son muy influyentes dentro de la red. Esto podría indicar que sus

publicaciones o colaboraciones son altamente valoradas, a pesar de su baja interconexión. Neto y Ramos también presentan valores altos, mientras que Santos tiene un valor más bajo, lo que sugiere que su influencia es menor en comparación con los otros autores, a pesar de su alto grado de interconexión. (Bonacich, 1987) sostiene que la centralidad de autovector es fundamental para entender el poder y la influencia dentro de redes complejas.

Tabla 7

Resultados de medidas adicionales del análisis de co-autor por autor

Autor	GC	GI	CI	CC	CA	# Doc.
Arcêncio, Ricardo Alexandre	3	0	0	0.10000000	1.00000000	11
Neto, Marcelino Santos	3	0	0	0.09090909	0.9763480	6
Ramos, Antônio Carlos Vieira	3	0	0	0.11111111	0.9354242	6
Santos, Floriacy Stabnow	3	1	1	0.12500000	0.7684481	5

Nota: GC: Grado de centralidad, GI: Grado de intermediación, CI: Centralidad de intermediación, CC: Centralidad de cercanía, CA: Centralidad de autovector, # Doc.: número de documentos.

En resumen, los resultados indican la presencia de una red de coautoría bien interconectada entre los autores analizados, con un equilibrio entre cercanía e influencia. Se observa que los grupos de coautores mantienen vínculos directos entre sí, sin conexiones con autores externos al grupo. Por lo tanto, la red consta de un solo componente, es decir, un único clúster (Saetnan & Kipling, 2016).

Santos desempeña un papel clave como intermediario, mientras que Arcêncio sobresale en términos de influencia, a pesar de su limitada capacidad para conectar a otros autores. La colaboración entre ambos podría ser fundamental para el avance de sus investigaciones conjuntas. Esta dinámica se ve respaldada por estudios que demuestran que la cooperación efectiva en redes académicas fomenta una mayor productividad y mejora la calidad de la investigación (Katz & Martin, 1997).

3.2.3.2. Red de co-autoría por organización

3.2.3.2.1. Mapeo de red de co-autoría por organización

En la segunda red, cuya unidad de análisis es la organización, el análisis de coautoría permite identificar un mapa detallado de las colaboraciones más relevantes a nivel nacional e internacional en el ámbito de los datos abiertos medioambientales. Este enfoque destaca las redes de cooperación entre

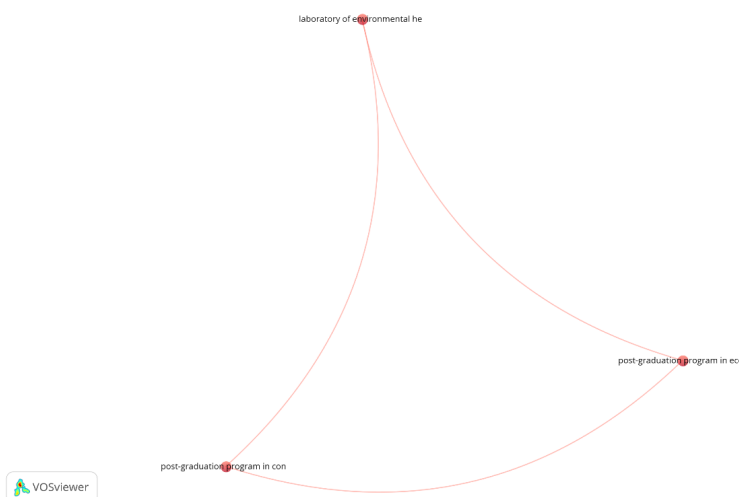
universidades, laboratorios, institutos de investigación y otras entidades, ofreciendo una visión clara sobre los procesos de generación y difusión del conocimiento en este campo.

Para el análisis realizado con VOSviewer, se estableció un umbral mínimo de tres documentos por organización y al menos una cita por organización. Como resultado, se identificaron 3,957 organizaciones, de las cuales solo 9 cumplieron con estos criterios, representando un 0.22% del total. Entre estas, únicamente 3 organizaciones evidenciaron una interrelación significativa en la red de coautoría.

Tal como se muestra en la Figura 6, el "Laboratorio de salud ambiental y ecotoxicología, Departamento de ciencias ambientales de la Universidad de Jahangirnagar (Laboratory of environmental health and ecotoxicology, department of environmental sciences jahangirnagar university)", el "Programa de posgrado en conservación de los recursos naturales del cerrado, Instituto Federal de Goiano (Post-graduation program in conservation of cerrado natural resources, goiano federal institute)", y el "Programa de posgrado en ecología, conservación y biodiversidad, Universidad Federal de Uberlândia (Post-graduation program in ecology, conservation, and biodiversity, federal university of uberlândia)" presentan el mismo número de citaciones (14) y una fuerza total de enlace (FTE) idéntica (6) en la red de co-autoría.

Figura 6

Visualización de red co-autoría por organización



Nota: Autores (2024).

3.2.3.2.2. Medidas adicionales del análisis de co-autoría por organización

En la Tabla 8, la columna correspondiente al grado de centralidad (GC) muestra que tres organizaciones destacan: el Laboratorio de Salud Ambiental y Ecotoxicología de la Universidad de Jahangirnagar, el Programa de Posgrado en Conservación de los Recursos Naturales del Cerrado del Instituto Federal de Goiano, y el Programa de Posgrado en Ecología, Conservación y Biodiversidad de la Universidad Federal de Uberlândia. Estas tres instituciones tienen un grado de centralidad de 2 y están asociadas con tres documentos cada una.

Este grado de centralidad indica que cada organización mantiene conexiones directas con otros nodos, ya sean otras organizaciones o autores, dentro de la red de coautoría. Aunque el grado de centralidad es moderado, sugiere que estas instituciones participan activamente en colaboraciones relacionadas con investigaciones sobre el medio ambiente y la biodiversidad.

De manera similar, en la Tabla 8, la columna correspondiente al grado de intermediación (GI) muestra que todas las organizaciones tienen un valor de 0. Esto indica que ninguna de estas organizaciones actúa como puente en las conexiones entre otros nodos dentro de la red de coautoría. La ausencia de grado de intermediación sugiere que, aunque estas instituciones están colaborando, no están facilitando conexiones entre otros autores o instituciones, lo que limita su papel en la mediación del flujo de información dentro de la red. Este resultado puede sugerir que hay poca interacción entre los diferentes grupos de investigación o que las colaboraciones son más internas que externas.

De igual manera, la columna correspondiente a la centralidad de intermediación (CI) en la Tabla 8 muestra valores de 0 para todas las organizaciones, al igual que el grado de intermediación. Esto refuerza la idea de que ninguna de estas instituciones desempeña un papel clave en la conectividad de la red, lo que podría indicar que están más enfocadas en sus propias investigaciones sin actuar como intermediarias entre otros grupos de investigación. Este hallazgo es importante porque la capacidad de ser un intermediario en redes académicas a menudo está asociada con una mayor influencia y capacidad para acceder a recursos y oportunidades de colaboración (Freeman, 1977).

En la columna correspondiente a la centralidad de cercanía (CC) de la Tabla 8, se observa que todas las organizaciones presentan un valor de 0.1666667. Este valor sugiere que, aunque no actúan como intermediarias, están relativamente cerca de otros nodos dentro de la red, lo que podría facilitar la comunicación y la colaboración directa. Este nivel de cercanía indica que, si bien no desempeñan un rol intermediario, tienen la capacidad de acceder rápidamente a otros nodos, lo cual es beneficioso para la difusión de información y el establecimiento de nuevas colaboraciones.

Finalmente, en la columna correspondiente a la Centralidad de Autovector (CA) de la Tabla 8, se observa que todas las organizaciones tienen un valor de 1. Esto indica que todas están altamente conectadas con otros nodos influyentes dentro de la red. Este resultado es significativo, ya que sugiere que, aunque no actúan como intermediarias, las colaboraciones que mantienen son con nodos que ejercen una gran influencia en el campo. Esto implica que, a pesar de no mediar en las conexiones, estas organizaciones están vinculadas a actores clave en la investigación, lo que puede mejorar su reputación y visibilidad en el ámbito académico.

Tabla 8

Resultados de medidas adicionales del análisis de co-autoría por organización

Organización	GC	GI	CI	CC	CA	# Doc.
Laboratorio de salud ambiental y ecotoxicología, Departamento de ciencias ambientales de la Universidad de Jahangirnagar, Bangladesh	2	0	0	0.1666667	1	3
Programa de posgrado en conservación de los recursos naturales del cerrado, Instituto Federal de Goiano, Brasil	2	0	0	0.1666667	1	3
Programa de posgrado en ecología, conservación y biodiversidad, Universidad Federal de Uberlândia, Brasil	2	0	0	0.1666667	1	3

Nota: GC: Grado de centralidad, GI: Grado de intermediación, CI: Centralidad de intermediación, CC: Centralidad de cercanía, CA: Centralidad de autovector, # Doc.: número de documentos.

En síntesis, los resultados revelan que las organizaciones analizadas presentan un grado de centralidad y de cercanía moderados, lo que sugiere una participación en la coautoría, aunque sin desempeñar un rol intermediario dentro

de la red. Esto podría reflejar un enfoque en la colaboración interna y una limitada conexión con otros grupos de investigación. No obstante, su alta centralidad de autovector indica que están colaborando con nodos influyentes, lo cual puede ser ventajoso para su posicionamiento en el campo de estudio.

3.2.3.3. Red de co-autoría por países

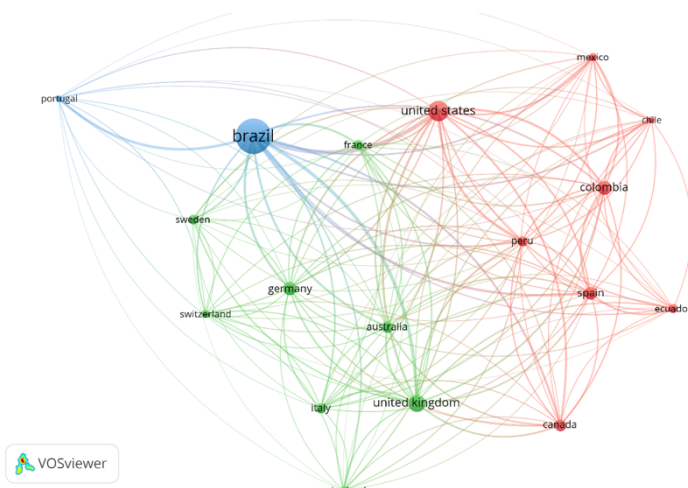
3.2.3.3.1. Mapeo de red de co-autoría por países

En la tercera red, cuyo análisis se centra en los países, el estudio de coautoría revela la influencia que cada nación ejerce dentro de la comunidad científica, medida a través del número de citas recibidas. Para este análisis, se establecieron como criterios un mínimo de 20 documentos por país y al menos una cita por país. Estos parámetros permiten identificar los países con mayor impacto y relevancia en el campo de estudio, basándose en sus contribuciones académicas.

Al aplicar el análisis de coautoría con VOSviewer a los datos extraídos de Scopus, se identificaron un total de 110 países, de los cuales solo 18 conforman una red con conexiones sólidas, lo que representa el 16.36% del total. Tal como se muestra en la Figura 7, Brasil se destaca como el país con el mayor número de citas en Scopus, lo que refleja su considerable influencia e impacto en las investigaciones dentro del área de estudio, así como su papel clave en la promoción de la colaboración internacional.

Figura 7

Visualización de red co-autoría por países



Nota: Autores (2024).

Se identificaron tres clústeres principales que representan distintas regiones geográficas, con países como Canadá, Australia y Brasil liderando, por grupos, la productividad en este campo. El primer grupo ($n=8$), representado en color rojo, tiene a Colombia como el nodo más grande, con 130 publicaciones y una fuerza total del enlace (FTE) de 197, seguido por España (73 publicaciones y 187 FTE), Perú (73 publicaciones y 180 FTE), Ecuador (63 publicaciones y 103 FTE), Canadá (48 publicaciones y 165 FTE), México (31 publicaciones y 124 FTE) y Chile (28 publicaciones y 112 FTE).

El segundo grupo, representado en color verde ($n=8$), está liderado por el Reino Unido, con 99 publicaciones y una fuerza total del enlace (FTE) de 287, seguido por Alemania (46 publicaciones y 167 FTE), Francia (37 publicaciones y 130 FTE), Australia (36 publicaciones y 166 FTE), Italia (28 publicaciones y 121 FTE), Países Bajos (23 publicaciones y 85 FTE), Suiza (22 publicaciones y 101 FTE) y Suecia (22 publicaciones y 85 FTE). El tercer grupo, en color celeste ($n=2$), tiene a Brasil como el nodo más destacado, con 788 documentos y 502 FTE, seguido por Portugal con 45 documentos y 83 FTE.

3.2.3.3.2. Medidas adicionales del análisis de co-autoría por autor

En otro orden de ideas, los resultados obtenidos del análisis de coautoría por países en el contexto de los datos abiertos medioambientales pueden complementarse con medidas bibliométricas adicionales. Estos resultados se alinean con estudios recientes sobre redes de colaboración científica, proporcionando una interpretación más integral de las dinámicas de cooperación internacional en este ámbito.

Una de las medidas utilizadas es el grado de centralidad, que refleja la importancia de un país dentro de la red de coautoría en función de sus conexiones directas con otros países. En la columna GC de la Tabla 9, se observan países como Estados Unidos, Colombia y Brasil con un alto grado de centralidad, lo que indica una participación activa y directa en redes internacionales de investigación. Este patrón es consistente con estudios previos que destacan cómo los países con alta productividad científica tienden a ocupar

nodos centrales en las redes de colaboración, lo que incrementa su visibilidad e influencia dentro del campo de investigación (Glänzel & Schubert, 2005).

Por otro lado, las medidas de grado de intermediación y centralidad de intermediación evalúan la capacidad de un país para actuar como “puente” o intermediario entre otras naciones dentro de la red. En la columna GI y CI de la Tabla 9, se destaca que países como Portugal, Ecuador, Suecia, Países Bajos y Chile sobresalen en estas métricas, lo que sugiere que facilitan la colaboración entre otros nodos (países). Este hallazgo resalta su rol crucial en la promoción de conexiones internacionales. Esto es consistente con investigaciones que muestran que las redes de co-autoría no solo se fortalecen por la cantidad de publicaciones, sino también por la capacidad de ciertos actores de conectar a otros en la red (Biscaro & Giupponi, 2014).

En cuanto a, la medida de centralidad de cercanía indica la rapidez con la que un país puede acceder a otros nodos dentro de la red. Un país como Brasil, que presenta una cifra baja según la columna CC de la Tabla 9, podría, en principio, sugerir un error en la medición. Sin embargo, estudios bien establecidos han demostrado que un bajo valor de CC puede indicar una rápida difusión de información científica dentro de la red, lo que maximiza el impacto de las publicaciones (Glänzel & Schubert, 2005).

Finalmente, al analizar la centralidad de autovector, se observa en la columna CA de la Tabla 9 que países como Brasil (CA = 1.0000000), Estados Unidos (CA = 0.8709249), Reino Unido (CA = 0.6512679) y Colombia (CA = 0.4058418) muestran altos valores, lo que refleja su considerable influencia en la red. Estos países no solo participan en numerosas colaboraciones, sino que también lo hacen con otros nodos clave, lo que incrementa su capacidad de influir en la agenda global de investigación. Investigaciones recientes subrayan que la centralidad de autovector es un buen indicador de la capacidad de un país para influir en las dinámicas globales de producción científica (Haunschild et al., 2016).

Tabla 9

Resultados de medidas adicionales del análisis de co-autoría por países

País	GC	GI	CI	CC	CA	# Doc.
Alemania	17	0.0000000	0.0000000	0.010101010	0.3745379	46
Australia	17	0.0000000	0.0000000	0.009615385	0.3765792	36
Brasil	17	0.0000000	0.0000000	0.004901961	1.0000000	788
Canadá	16	0.0000000	0.0000000	0.009259259	0.3850945	48
Chile	17	1.0000000	1.0000000	0.015625000	0.2454846	28
Colombia	17	0.0000000	0.0000000	0.010989011	0.4058418	130
Ecuador	17	32.4000000	32.4000000	0.016666667	0.2145696	63
España	17	0.0000000	0.0000000	0.010309278	0.4014205	73
Estados Unidos	17	0.0000000	0.0000000	0.006756757	0.8709249	146
Francia	17	0.0000000	0.0000000	0.012500000	0.3006952	37
Italia	17	0.0000000	0.0000000	0.012345679	0.2580986	28
México	17	0.2000000	0.2000000	0.015625000	0.2743332	31
Países bajos	17	9.6666667	9.6666667	0.016666667	0.1926442	23
Perú	17	0.0000000	0.0000000	0.010526316	0.3840434	73
Portugal	16	79.9000.000	79.9000000	0.019230769	0.2513589	45
Reino Unido	17	0.0000000	0.0000000	0.010000000	0.6512679	99
Suecia	17	19.8666667	19.8666667	0.015873016	0.2071458	22
Suiza	17	0.6666667	0.6666667	0.013513514	0.2156397	22

Nota: GC: Grado de centralidad, GI: Grado de intermediación, CI: Centralidad de intermediación, CC: Centralidad de cercanía, CA: Centralidad de autovector, # Doc.: número de documentos

En resumen, países como Estados Unidos y Colombia se destacan por su alta influencia y participación directa en la red, mientras que Chile y Ecuador se posicionan como intermediarios clave que conectan a otros países. Por su parte, Brasil presenta un caso particular, con un número significativamente mayor de documentos (788), lo que lo convierte en uno de los países más productivos. Sin embargo, su centralidad de intermediación (CI = 0.004901961) es relativamente baja.

Estos hallazgos están en consonancia con la literatura científica actual, que resalta la importancia de las redes de coautoría y la colaboración internacional en la producción de conocimiento sobre temas críticos como el cambio climático y la sostenibilidad (Glänzel & Schubert, 2005; Haunschild et al., 2016).

3.2.4. Análisis de co-ocurrencia de palabras claves

3.2.4.1. Mapeo de red de co-autoría por países

El análisis de co-ocurrencia es una visualización que refleja la frecuencia con la que términos, palabras clave o conceptos específicos aparecen juntos en un conjunto de publicaciones académicas, permitiendo identificar las palabras clave más influyentes (San-Juan-Heras et al., 2024). Se seleccionó la unidad de análisis Palabras clave del autor y como método de conteo el “conteo completo”.

Como umbral para la visualización, se estableció un número mínimo de 10 ocurrencias para las palabras clave. El análisis resultó en la identificación de 3,293 palabras clave, de las cuales 29 cumplen con este criterio, representando el 0.88% del total.

Los resultados generados por el software VOSviewer presentan las visualizaciones de las palabras clave extraídas de los datos de Scopus, como se muestra en la Figura 8. Destacan principalmente los cinco primeros términos: “public health”, “covid-19”, “brazil”, “spatial analysis” y “mortality”.

Figura 8

Visualización del análisis de co-ocurrencias en VOSviewer



Nota: Autores (2024).

3.2.4.1.1. Medidas adicionales del análisis de co-ocurrencia

El análisis de co-ocurrencia de términos es una herramienta esencial para comprender la interrelación entre diferentes conceptos o palabras clave en un corpus de documentos o estudios. En este estudio, se han incorporado métricas adicionales de co-ocurrencia entre términos clave, utilizando el software VOSviewer, en el contexto de los datos abiertos medioambientales. Estas medidas proporcionan una visión más profunda de las conexiones conceptuales presentes en la literatura, facilitando el análisis estructural del campo de estudio.

Se llevaron a cabo cálculos de medidas adicionales utilizando R Studio y el paquete iGraph, centrados en el grado de centralidad (GC) y la centralidad de intermediación (CI). Estas métricas, fundamentales en los análisis de redes, permiten evaluar tanto la relevancia de las palabras clave como su papel como intermediarias dentro de la red. Su aplicación es crucial para identificar los nodos clave que estructuran las interrelaciones en el ámbito de los datos abiertos medioambientales.

El Grado de Centralidad (GC) es una métrica que refleja el número de conexiones directas que una palabra clave tiene con otras palabras dentro de una red, lo que indica su relevancia como nodo central. De acuerdo con Freeman (1977), un nodo con un mayor grado de centralidad presenta un mayor número de conexiones directas, lo que lo posiciona como un elemento central dentro de la red de interacciones, sugiriendo su importancia en el flujo de información.

Por otro lado, la centralidad de intermediación (CI) mide la capacidad de una palabra clave para actuar como un puente entre otras palabras, convirtiéndose así en un intermediario crucial dentro de la red. Esta métrica, introducida por Freeman (1977), destaca que los nodos con alta intermediación controlan el flujo de información entre segmentos desconectados de la red. Burt (2004) refuerza este concepto, al señalar que un nodo con elevada intermediación posee una ventaja estructural al conectar diferentes grupos, lo que le otorga un papel estratégico en la dinámica de la red.

La Tabla 10 presenta los resultados del análisis complementario de co-ocurrencia, destacando los siguientes términos clave:

- COVID-19 y Salud Pública:

Estos dos términos se destacan significativamente en cuanto a su grado de centralidad (GC) y centralidad de intermediación (CI). COVID-19 posee el GC más elevado (22) y una CI notablemente alta (67.37), lo que indica su posición central dentro del análisis, siendo el término más co-ocurrente con otros, lo que refleja su papel predominante en el discurso científico contemporáneo. Además, su alta CI sugiere que actúa como un conector crucial entre múltiples áreas temáticas, confirmando su rol como intermediario entre diversas disciplinas. Por su parte, Salud Pública tiene un GC de 21 y una CI de 18.23, lo que también la posiciona como un término altamente central en la red, aunque con un menor grado de intermediación en comparación con COVID-19.

- Brasil, Mortalidad y Análisis Espaciales:

Estos términos presentan altos valores de centralidad. Brasil tiene un GC de 14 y una CI de 59.47, lo que indica su fuerte conexión con otros términos clave, probablemente debido a la pandemia de COVID-19 y su impacto en esta región. Su alta CI refuerza la idea de que Brasil es un nodo importante que conecta temas relacionados con la epidemiología, la salud pública y la sostenibilidad. Por otro lado, Mortalidad (GC=14, CI=5.53) y Análisis Espaciales (GC=18, CI=18.03) juegan roles destacados, aunque con características diferentes. Mientras que "Mortalidad" parece estar asociada a un conjunto más específico de temas, "Análisis Espaciales" tiene un papel más amplio como intermediario entre estudios diversos, lo que sugiere aplicaciones en campos como la epidemiología, la teledetección y el análisis de datos geoespaciales.

- Coronavirus, Epidemiología y SARS-CoV-2:

Estos términos también presentan altos niveles de coocurrencia, especialmente Epidemiología (GC=16, CI=5.43) y SARS-CoV-2 (GC=9, CI=13.46). La prominencia de estos términos está claramente vinculada con el contexto de la pandemia y los estudios relacionados con la propagación del virus y las respuestas a la crisis. Coronavirus (GC=19, CI=17.31) está estrechamente relacionado con "COVID-19" y otros términos de salud pública, lo que sugiere un enfoque integral en la investigación sobre virus y su impacto en la salud global.

- Cambio Climático y Sostenibilidad:

Cambio Climático presenta un GC de 5 y una CI relativamente alta (24.42), lo que indica que, aunque no es un nodo central en la red, actúa como un conector importante en varios estudios. Esto podría reflejar un enfoque que conecta la salud pública con los efectos del cambio climático en la epidemiología y la sostenibilidad. Sostenibilidad, por su parte, tiene un GC moderado (7) y una CI de 8.69, lo que sugiere una conexión relevante en la red, aunque no desempeña un papel principal en este análisis específico.

Por otro lado, algunos términos, como Una Salud, Políticas Públicas, Desarrollo Sostenible y Aprendizaje Automático, presentan bajos valores de Grado de Centralidad y Centralidad de Intermediación. Esto sugiere que, en este análisis específico, estos temas no son tan prominentes ni están tan interconectados como otros. No obstante, su presencia en la red indica que ocupan un espacio dentro de la investigación, aunque su impacto o su relación con otras áreas podría estar menos desarrollada en este corpus de datos.

Tabla 10

Resultados de medidas adicionales del análisis de co-ocurrencia

Palabra clave	GC	CI	Ocurrencias
Amazonas	5	6.3435341	13
Brasil	14	59.4763633	34
cambio climático	5	24.4226190	19
Colombia	5	5.0830622	11
Coronavirus	19	17.3074176	15
COVID-19	22	67.3667155	66
estudios ecológicos	9	5.7379482	13
Ambiente	9	24.6763485	16
salud ambiental	10	19.8750934	13
epidemiología	16	5.4373950	65
sistemas de información sanitaria	9	7.0741814	12
hospitalización	9	12.6093784	11
Incidencia	12	12.0804335	12
aprendizaje automático	4	0.8261905	14
Mortalidad	14	5.5324.017	35
una salud	3	0.0000000	11
atención primaria de salud	9	10.8329365	12
salud publica	21	18.2372.209	69
vigilancia de salud pública	11	22.6205.331	17
políticas públicas	2	0.0000000	11
política pública	7	14.1694444	15
teledetección	4	3.0039683	17

SARS-CoV-2	9	13.4640183	21
determinantes sociales de la salud	8	28.7678988	11
análisis espaciales	18	18.0315221	44
Sostenibilidad	7	8.6978058	22
desarrollo sostenible	3	0.0000000	12
estudios de series de tiempo	10	0.0000000	19
Tuberculosis	08	5.0788822	12

Nota: GC: Grado de centralidad, CI: Centralidad de intermediación

En resumen, este análisis revela una marcada orientación hacia estudios relacionados con la pandemia de COVID-19, la epidemiología y su impacto en la salud pública, destacando notablemente el enfoque en países como Brasil. Los términos vinculados a la mortalidad, la vigilancia de la salud pública y el análisis espacial también desempeñan un papel relevante en la red, lo que sugiere que estos temas han sido fundamentales en la investigación reciente. Asimismo, se observa una interconexión significativa entre el cambio climático, la sostenibilidad y la salud, lo que indica un creciente interés por entender cómo los problemas ambientales interactúan con la salud pública.

Por otro lado, los términos con baja centralidad sugieren áreas de investigación emergentes o menos desarrolladas en este contexto, que podrían adquirir mayor relevancia a medida que los estudios sobre la interrelación entre salud, políticas públicas y desarrollo sostenible se expandan.

3.2.5. Análisis de co-citación

El análisis de co-citación se basa en la construcción de representaciones visuales que mapean las relaciones entre artículos académicos, autores o revistas, en función de la frecuencia con la que son citados conjuntamente en otros trabajos. Este enfoque permite identificar a los autores, fuentes o revistas, así como a los documentos que son citados con mayor frecuencia en conjunto, proporcionando una visión clara de los núcleos de conocimiento más influyentes dentro de un campo de estudio (San-Juan-Heras et al., 2024).

Para realizar este análisis, se utilizó la herramienta VOSviewer con el objetivo de identificar y visualizar las relaciones entre los autores citados, las citas y las fuentes referenciadas. Esta herramienta permitió mapear de manera efectiva las conexiones y patrones de co-citación, facilitando una comprensión más profunda de las dinámicas entre los principales actores dentro del campo de estudio.

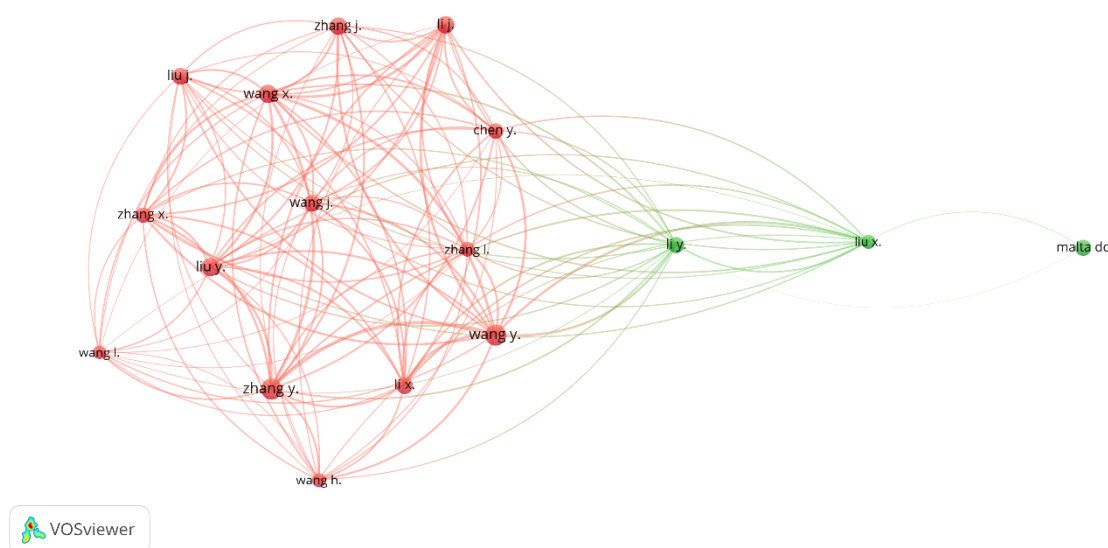
3.2.5.1. Red de co-citación por autores citados

3.2.5.1.1. Mapeo de red de co-citación por autores citados

La primera red se centró en los autores citados. Para ello, se seleccionó la unidad de análisis "Autores citados" y se aplicó el método de "conteo completo". El umbral mínimo para el número de citas de un autor se estableció en 50 citas en la base de datos Scopus. Los resultados destacan la influencia de ciertos autores dentro del grupo de autores activos; de un total de 130,721 autores que componen la red, solo 17 ítems forman parte de la red principal y están fuertemente interconectados. Los autores más citados están representados por nodos de mayor tamaño en el gráfico, mientras que las relaciones de co-citación más sólidas se visualizan mediante líneas más gruesas (López-Fraile et al., 2023). Destacan en la red los autores "Li J." y "Chen Y." por acumular el mayor número de citas, con un total de 114 citas en la base de datos Scopus. Estos autores ocupan una posición central en la red, como se muestra en la Figura 9.

Figura 9

Visualización de red co-citación de autores citados



Nota: Autores (2024).

3.2.5.1.2. Medidas adicionales del análisis de co-citación por autores citados

El análisis de co-citación por autor se amplió mediante el uso del paquete iGraph en R Studio, lo que permitió calcular métricas clave como el grado de centralidad, la centralidad de intermediación y la centralidad de cercanía para los autores más frecuentemente citados en los datos extraídos de Scopus. Además, se empleó el algoritmo de Louvain para la detección de comunidades, facilitando la identificación de patrones de agrupamiento entre los autores citados, lo que proporciona una visión más detallada de las relaciones y estructuras dentro de la red de co-citación.

En lo que respecta al grado de centralidad (GC), representado en la columna correspondiente de la Tabla 11, esta métrica mide el número de conexiones directas que un autor tiene con otros autores dentro de la red de co-citación. La mayoría de los autores presentan valores similares en esta métrica, con la excepción de Malta Dc, quien registra un GC de 2, lo que indica que este autor posee menos co-citaciones en comparación con el resto de los autores analizados.

Por otro lado, la columna correspondiente a la Centralidad de Intermediación (CI) refleja la capacidad de un autor para actuar como puente dentro de la red. En este caso, Wang L. presenta el valor más alto (CI = 53.0), lo que sugiere que este autor desempeña un papel crucial al conectar diferentes grupos en la red, un posible indicativo de su influencia en diversos temas. En contraste, la mayoría de los autores tienen valores de intermediación nulos o bajos, lo que sugiere una estructura de red menos jerárquica y más distribuida.

En cuanto a la centralidad de cercanía (CC), la Tabla 11, en su columna correspondiente, muestra qué tan próximo está un autor al resto de los autores en términos de distancias dentro de la red. Wang L. también sobresale en esta métrica, con un valor de 0.0040, lo que refuerza su relevancia estructural dentro de la red de co-citación, indicando que está bien posicionado para acceder a la información a través de múltiples rutas cortas en la red.

Tabla 11

Resultados de medidas adicionales del análisis de co-citación por autores citados

Autores citados	GC	CI	CC	Citaciones
Chen Y.	15	0.0	0.002710027	69
Li J.	15	0.0	0.002652520	81
Li X.	15	0.0	0.002053388	78
Li Y.	16	12.5	0.003134796	74
Liu J.	15	8.0	0.002785515	75
Liu X.	16	2.0	0.003344482	56
Liu Y.	15	0.0	0.001872659	88
Malta Dc	2	1.5	0.003030303	70
Wang H.	15	1.0	0.003048780	56
Wang J.	15	0.0	0.002512563	77
Wang L.	15	53.0	0.004016064	51
Wang X.	15	0.0	0.001754386	89
Wang Y.	15	0.0	0.001692047	114
Zhang J.	15	0.0	0.002544529	81
Zhang L.	15	1.0	0.002808989	60
Zhang X.	15	0.0	0.002028398	68
Zhang Y.	15	0.0	0.001689189	114

Nota: GC: Grado de centralidad, CI: Centralidad de intermediación, CC: Centralidad de cercanía

Las comunidades dentro de una red de co-citación por autor fueron identificadas utilizando algoritmos como el de Louvain, que permite agrupar a los autores que son citados de manera similar, sugiriendo que están trabajando en áreas temáticas o disciplinas afines. La Tabla 12 muestra las comunidades detectadas y sus respectivas disciplinas, basadas en las revistas o fuentes en las que los autores publican sus trabajos, así como en los temas principales abordados en sus artículos.

Tabla 12

Comunidades detectadas y disciplinas relacionadas

Comunidad	Autores citados	Disciplinas o áreas de conocimientos
1	Chen Y.	Salud pública
	Li X.	Epidemiología
	Li Y.	Sostenibilidad y medio ambiente
	Liu X.	Ciencias de la salud (especialmente en temas
	Malta Dc	colectivos o de salud pública en países en
	Wang H.	desarrollo)
	Wang Y.	

2	Zhang L.	
	Zhang Y.	
	Li J.	Investigación biomédica
	Liu J.	Ciencias biológicas
	Liu Y.	Enfermedades tropicales y salud global (con la
	Wang J.	inclusión de PLoS Neglected Tropical
	Wang L.	Diseases)
	Wang X.	
	Zhang J.	
	Zhang X.	

Nota: Autores (2024).

Finalmente, el valor de la modularidad (0.0040) sugiere una débil división en la red, lo que indica que las comunidades identificadas presentan una estructura modular poco pronunciada. Esto puede interpretarse como un indicio de que los temas de investigación abordados por los autores citados están altamente interconectados entre las distintas comunidades, lo que implica que las áreas de investigación no están claramente delimitadas y comparten un grado significativo de superposición temática.

Estos resultados son consistentes con estudios recientes sobre redes de co-citación, los cuales subrayan la relevancia de las redes altamente conectadas para el intercambio de conocimiento y la difusión de ideas dentro de las disciplinas científicas. Este tipo de redes facilita la circulación de información y promueve la colaboración entre autores, lo que contribuye al avance y desarrollo de los campos de estudio (Li et al., 2021; Small, 1973).

La co-citación de autores destacados refuerza su rol central en la consolidación del conocimiento dentro de sus respectivas áreas de investigación. Su alta frecuencia de citación conjunta sugiere que sus trabajos son considerados fundamentales y ejercen una influencia significativa en el desarrollo de sus campos de estudio.

3.2.5.2. Red de co-citación por fuentes citadas

3.2.5.2.1. Mapeo de red de co-citación por fuentes citadas

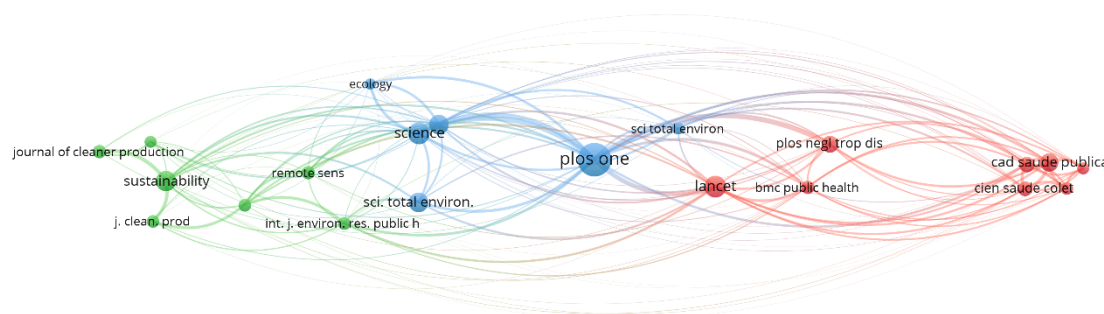
La segunda red se centró en las revistas citadas. Se seleccionó la unidad de análisis "Fuentes citadas" y se utilizó el método de "conteo completo" para el análisis. El umbral mínimo de citas para una fuente se estableció en 100,

resultando en un total de 21 fuentes que cumplen con este criterio, de un total de 21,873 fuentes identificadas. Estas 21 fuentes están fuertemente relacionadas entre sí dentro de la red.

Entre las principales fuentes con mayor número de citas y fuerza de enlace se destacan: en la Figura 10: "PLOS ONE" (128 citas, 2588 de fuerza de enlace), "Science" (110 citas, 1442 de fuerza de enlace), "The Lancet" (164 citas, 1,380 de fuerza de enlace), "Nature" (152 citas, 1,288 de fuerza de enlace) y "Cadernos de Saúde Pública" (306 citas, 1,130 de fuerza de enlace).

Figura 10

Visualización de red co-citación de fuentes citadas



Nota: Autores (2024).

3.2.5.2.2. Medidas adicionales del análisis de co-autoría por autor

Complementando los resultados, se llevó a cabo un análisis estadístico basado en las métricas adicionales del análisis de co-citación de las fuentes citadas. La Tabla 12 presenta los resultados sobre diversas fuentes citadas en el contexto del presente estudio científico.

La métrica de grado de centralidad mide el número de conexiones directas que posee cada fuente en la red de co-citación. En la Tabla 12, la columna correspondiente al grado de centralidad (GC) presenta las publicaciones con mayor grado de centralidad, destacando aquellas que actúan como los nodos más interconectados dentro de la red.

- Nature, PLOS ONE, Science y The Lancet son las fuentes que presentan el mayor grado de centralidad (GC) con valores de 20, 20, 20 y 19, respectivamente.

Estas revistas son altamente referenciadas y desempeñan un papel central en la red de co-citación, lo que sugiere que son ampliamente citadas en investigaciones relacionadas. Se posicionan como nodos centrales debido a su alta frecuencia de citación, actuando como pilares de la red científica global, similar a lo que se ha demostrado en estudios sobre la evolución de redes científicas en áreas multidisciplinarias (Lafia et al., 2022).

- BMC Public Health y Sustainability también se destacan con un grado de centralidad (GC) de 19, lo que indica su relevancia como nodos clave en la red de co-citación de este estudio.

La centralidad de intermediación refleja la capacidad de una fuente para conectar a otras fuentes en la red, actuando como un "puente" que facilita el flujo de información entre diferentes nodos. Estas fuentes no solo vinculan a autores, sino que también promueven la integración de ideas de distintos campos, lo que se alinea con investigaciones que destacan la importancia de la intermediación en la promoción de la innovación y la difusión de ideas entre dominios diversos (Rossini, 2024). La columna etiquetada como Centralidad de Intermediación (CI) en la Tabla 12 muestra los resultados que indican:

- Ecology (100.0) y Remote Sens (50.67) presentan valores de intermediación significativamente altos, indicando que estas publicaciones desempeñan un papel crucial al conectar diferentes grupos de citación. Esto puede sugerir que sus contenidos abarcan múltiples áreas temáticas o interdisciplinarias.
- En contraste, muchas de las fuentes, como Lancet, Plos One y BMC Public Health, tienen valores de CI bajos o nulos, lo que sugiere que, aunque son frecuentemente citadas, no juegan un rol crítico como nodos intermediarios entre otras fuentes.

Por otro lado, la centralidad de cercanía mide la proximidad de una fuente a todas las demás en la red. Esta métrica está asociada a la capacidad de una fuente

para funcionar como un centro de intercambio intelectual dentro de su campo. Las fuentes con valores más altos de centralidad de cercanía se encuentran más cerca de las demás en términos de distancia dentro de la red y están representadas en la columna CC de la Tabla 13.

Las fuentes con los valores más altos de centralidad de cercanía incluyen BMC Public Health (0.01), Ecology (0.01) y Science (0.006), lo que sugiere que estas revistas facilitan un acceso eficiente a otras publicaciones dentro de la red.

Por otro lado, publicaciones como PLOS ONE (0.002) presentan una menor centralidad de cercanía, lo que podría indicar que, aunque son ampliamente citadas, no están tan interconectadas con otras fuentes en la red de co-citación en términos de proximidad.

Tabla 13

Resultados de medidas adicionales del análisis de co-citación por fuentes citadas

Fuentes citadas	GC	CI	CC	Citaciones
Bmc Public Health	19	43.833333	0.01000000	127
Cad Saude Publica	12	0.000000	0.005952381	219
Cien Saude Colet	16	29.666667	0.009708738	152
Ecology	15	100.000000	0.010416667	102
Epidemiol Serv Saude	13	2.833333	0.009174312	119
Int. J. Environ. Res. Public Health	12	0.000000	0.005813953	122
J. Clean. Prod	9	22.166667	0.007936508	114
Journal Of Cleaner Production	5	1.000.000	0.004310345	128
Lancet	19	0.000000	0.005882353	298
Nature	20	1.333333	0.007462687	245
Plos Negl Trop Dis	14	0.000000	0.004807692	189
Plos One	20	0.000000	0.002341920	634
Remote Sens	14	50.666667	0.009708738	133
Rev Bras Epidemiol	13	21.500000	0.009174312	111
Rev Saude Publica	13	22.000000	0.009009009	164
Sci Total Environ	15	32.666667	0.007575758	110
Sci. Total Environ	12	0.000000	0.005586592	116
Sci. Total Environ.	9	0.000000	0.003300330	259
Science	20	18.000000	0.006097561	306
Science Of The Total Environment	11	1.000000	0.006711409	106
Sustainability	19	19.500000	0.009009009	264

Nota: GC: Grado de centralidad, CI: Centralidad de intermediación, CC: Centralidad de cercanía

En relación con el resultado de la modularidad, el análisis presenta un valor de 0.27, lo que indica una estructura comunitaria moderada en la red. Esto sugiere que, aunque existen grupos distintivos de fuentes que tienden a citarse mutuamente, la red no está fuertemente fragmentada.

La identificación de comunidades dentro de la red es consistente con estudios que indican que las redes científicas tienden a formar "clusters" o grupos temáticos. La modularidad en redes de co-citación refleja la especialización temática, como se observa en las comunidades centradas en salud pública, ecología y sostenibilidad, lo cual ha sido evidenciado en análisis recientes sobre las estructuras de investigación en áreas multidisciplinarias.

Se identificaron tres comunidades principales dentro de la red de co-citación:

- Comunidad 1: Incluye fuentes centradas principalmente en la salud pública y la epidemiología, como BMC Public Health, Cadernos de Saúde Pública, The Lancet y Revista de Saúde Pública. Estas fuentes conforman una comunidad coherente relacionada con temas de salud pública y epidemiología.
- Comunidad 2: Agrupa revistas de impacto interdisciplinar y global, tales como Ecology, PLOS ONE y Science, lo que sugiere que estas fuentes abordan una amplia gama de disciplinas científicas.
- Comunidad 3: Está compuesta por revistas vinculadas a la investigación ambiental y la sostenibilidad, como International Journal of Environmental Research and Public Health, Journal of Cleaner Production y Sustainability. Estas fuentes se citan entre sí en estudios relacionados con sostenibilidad y medio ambiente.

El análisis de co-citación revela una red en la que ciertas fuentes, como Nature, PLOS ONE y Science, desempeñan roles centrales debido a su alta frecuencia de citación. Por otro lado, fuentes como Ecology y Remote Sensing son cruciales por su capacidad para conectar diferentes áreas temáticas. La modularidad relativamente baja sugiere que, aunque existen comunidades temáticas diferenciadas, la red está interconectada, lo que permite la circulación de ideas entre disciplinas como la salud pública, la ecología y la sostenibilidad.

En términos generales, este análisis refleja cómo ciertas revistas no solo actúan como nodos de alta citación, sino también como puntos de conexión clave entre diferentes áreas de investigación. Esta interconexión es esencial para el avance interdisciplinario en campos como la salud y el medio ambiente.

3.2.5.3. Red de co-citación por referencias citadas

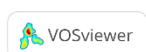
3.2.5.3.1. Mapeo de red de co-citación por referencias citadas

La tercera red se centró en los documentos. Se seleccionó la unidad de análisis "Referencias citadas" y se utilizó el método de "conteo completo" para el análisis. El umbral mínimo de citas para una referencia se estableció en 5, resultando en un total de 52,340 citas referenciadas. De estas, 16 referencias cumplen con el criterio establecido, y solo 4 de ellas están fuertemente relacionadas entre sí.

La Figura 11 presenta la referencia más citada, que corresponde a "R: A Language and Environment for Statistical Computing" (2021), con un total de 15 citaciones. Esta referencia es seguida por "R: A Language and Environment for Statistical Computing" (2020), que cuenta con 12 citaciones.

Figura 11

Visualización de red co-citación de referencias citadas



Nota: Autores (2024).

3.2.5.3.2. Medidas adicionales del análisis de co-citación por referencias citadas

El análisis de co-citación facilita la identificación de las relaciones y conexiones entre las referencias citadas en un grupo de estudios. En este contexto, los resultados presentados en la Tabla 14 reflejan el análisis de diversas medidas adicionales, tales como el grado de centralidad (GC), la centralidad de intermediación (CI) y la centralidad de cercanía (CC), las cuales permiten evaluar el rol y la influencia de las referencias clave dentro de la red de co-citación.

1. Kim Hj, Fay Mp, Feuer Ej, Midthune Dn., "Permutation Tests For Joinpoint Regression With Applications To Cancer Rates", Stat Med, 2000:

Esta referencia presenta un bajo grado de centralidad ($GC = 1$), lo que sugiere que su papel en la red es modesto en términos de conexiones directas con otros artículos. La centralidad de intermediación ($CI = 0$) indica que no actúa como un intermediario importante dentro de la red de co-citación, es decir, no conecta grupos de referencias entre sí. Sin embargo, su centralidad de cercanía ($CC = 0.1667$) refleja que se encuentra relativamente cerca de otras referencias, lo que implica que su acceso a otros nodos dentro de la red no es extremadamente limitado. Aunque ha sido citada seis veces, su impacto en la red general de co-citación es bajo.

2. R: A Language And Environment For Statistical Computing, 2020:

Esta referencia muestra una centralidad de intermediación significativa ($CI = 2$), lo que sugiere que desempeña un papel clave en la conexión entre diferentes conjuntos de estudios dentro de la red de co-citación. Su centralidad de cercanía ($CC = 0.2500$) indica que tiene un acceso relativamente directo a otros nodos en la red. Con 12 citaciones, esta referencia es influyente y juega un rol relevante en el campo de la investigación relacionada con el uso de software estadístico en análisis medioambientales y otros contextos.

3. R: A Language And Environment For Statistical Computing, 2021:

Similar a su versión del 2020, esta referencia es aún más prominente en términos de grado de centralidad ($GC = 3$), lo que refleja su fuerte conexión con múltiples estudios. También presenta una importante centralidad de intermediación ($CI =$

2), indicando que conecta distintos grupos de estudios, lo que es crucial para la difusión del conocimiento a través de la red. Su centralidad de cercanía ($CC = 0.2500$) sugiere que está bien posicionada dentro de la red, facilitando un acceso eficiente a otras referencias. Con 15 citaciones, es una referencia clave en la literatura actual sobre herramientas estadísticas.

4. Wickham H., "Ggplot2: Elegant Graphics For Data Analysis", 2016:

Esta referencia, centrada en la visualización de datos, presenta un grado de centralidad de 2, lo que indica que está conectada con otros trabajos relevantes. Sin embargo, su centralidad de intermediación ($CI = 0$) revela que no actúa como un conector significativo entre distintos conjuntos de estudios. Su centralidad de cercanía ($CC = 0.1667$) sugiere que, aunque está algo próxima a otros nodos, su influencia en términos de acceso a otras referencias no es tan alta como otros estudios en la red. Con 11 citaciones, sigue siendo una referencia relevante, particularmente en el ámbito de la visualización de datos en la investigación.

Tabla 14

Resultados de medidas adicionales del análisis de co-citación por referencias citadas

Referencias citadas	GC	CI	CC	Citaciones
Kim Hj, Fay Mp, Feuer Ej, Midthune Dn., Permutation Tests For Joinpoint Regression With Applications To Cancer Rates, Stat Med, 19, 3, Pp. 335-351, (2000)	1	0	0.1666667	6
R: A Language And Environment For Statistical Computing, (2020)	2	2	0.2500000	12
R: A Language And Environment For Statistical Computing, (2021)	3	2	0.2500000	15
Wickham H., Ggplot2: Elegant Graphics For Data Analysis, (2016)	2	0	0.1666667	11

Nota: GC: Grado de centralidad, CI: Centralidad de intermediación, CC: Centralidad de cercanía

En síntesis, las referencias que destacan en el análisis son las relacionadas con el software estadístico R. Estas publicaciones no solo son las más citadas, sino que también presentan altos valores en la centralidad de intermediación y cercanía, lo que sugiere que actúan como puntos clave en la red de co-citación. Esto se debe a su relevancia en el análisis de datos, tanto para el contexto medioambiental como para otros campos científicos. La referencia relacionada con la visualización de datos mediante ggplot2 también resulta ser notablemente

citada, aunque su papel dentro de la red parece más periférico en comparación con las publicaciones sobre R.

Por otro lado, el artículo de Kim Hj et al. presenta valores bajos en todas las medidas, lo que refleja que, aunque ha sido citado, no tiene un papel central en la red de co-citación, lo que podría deberse a la especificidad de su aplicación en análisis de tasas de cáncer.

Este análisis resalta la importancia de herramientas estadísticas como R y la visualización de datos con ggplot2 en el campo de los datos abiertos medioambientales. La alta centralidad de intermediación y cercanía de estas referencias sugiere que actúan como pilares fundamentales para la investigación actual y futura en este campo.

La identificación de comunidades es otra de las medidas adicionales calculadas. Basado en el cálculo del modularidad de la co-citación, se ha obtenido una Comunidad. Esta comunidad revela un grupo de referencias citadas que comparten una temática común, lo que indica una estrecha relación entre ellas en cuanto a su uso conjunto en investigaciones. La Tabla 14 contiene el resultado de estas referencias, las cuales se agrupan en torno a dos áreas principales:

- Métodos estadísticos y herramientas de análisis de datos: Incluye referencias como R: A Language and Environment for Statistical Computing (2020, 2021) y ggplot2 (Wickham, 2016), que son fundamentales para la manipulación y visualización de datos en investigaciones científicas. Estas herramientas se emplean ampliamente en diversas disciplinas para realizar análisis estadísticos complejos y generar representaciones gráficas efectivas.
- Análisis de tendencias y puntos de cambio: La referencia de Kim Hj et al. (2000), que aborda la regresión de puntos de cambio (joinpoint regression), es representativa de estudios centrados en el análisis de tendencias, especialmente en contextos donde es crucial detectar variaciones significativas a lo largo del tiempo, como en estudios epidemiológicos o de tasas de enfermedades.

Estas dos áreas temáticas reflejan la interrelación de las referencias citadas dentro de esta comunidad, que combina el uso de herramientas estadísticas avanzadas con técnicas de análisis de tendencias.

Tabla 15

Resultado de medida adicional comunidad

Comunidad	Referencias citadas
1	Kim Hj, Fay Mp, Feuer Ej, Midthune Dn., Permutation Tests For Joinpoint Regression With Applications To Cancer Rates, Stat Med, 19, 3, Pp. 335-351, (2000) R: A Language And Environment For Statistical Computing, (2020) R: A Language And Environment For Statistical Computing, (2021) Wickham H., Ggplot2: Elegant Graphics For Data Analysis, (2016)

Nota: Autores (2024).

En conjunto, esta comunidad parece estar compuesta por estudios que se apoyan en técnicas estadísticas avanzadas y herramientas computacionales robustas, como R y ggplot2, para abordar problemas complejos, como la identificación de puntos de cambio en datos temporales. Las conexiones entre estas referencias sugieren que los investigadores dentro de esta comunidad recurren a estos métodos para obtener resultados precisos y visualizaciones efectivas en sus análisis científicos.

3.2.6. Análisis temático-basados en datos de textos

Para llevar a cabo el análisis y la detección de tendencias en el tema de estudio, se utilizó la función de generación de mapas basada en datos textuales de VOSviewer, extrayendo información de los campos de título y resumen de los artículos seleccionados de Scopus. Se estableció como criterio de selección un mínimo de 150 ocurrencias de los datos extraídos por VOSviewer. Se empleó el método de conteo completo o "Full counting", complementado con la incorporación del thesaurus predeterminado de temas de VOSviewer para refinar el análisis.

Como resultado, la Figura 12 presenta un mapa de densidad que identifica un total de 27,882 términos, de los cuales 33 se destacan como los más relevantes, según la evaluación del modelo. El mapa de calor, generado por VOSviewer a partir de un análisis de redes textuales de artículos extraídos de Scopus, muestra

las relaciones entre los términos clave obtenidos de los títulos y resúmenes de artículos relacionados con datos abiertos medioambientales.

El mapa refleja la frecuencia de co-ocurrencia de los términos, en el cual los colores indican la densidad o relevancia de los mismos dentro del corpus de artículos analizados. El tamaño de los círculos representa el peso o número de veces que un término aparece, mientras que el degradado de color refleja las puntuaciones medias de citas de los artículos.

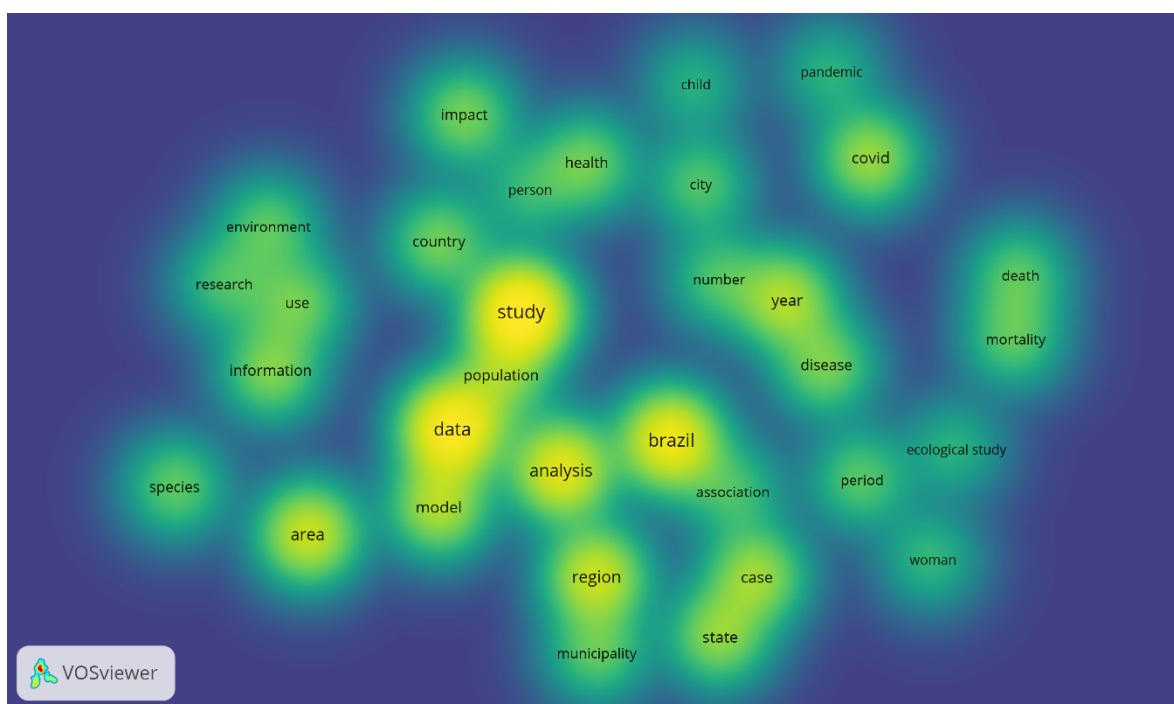
Los términos más destacados, representados en color amarillo y ubicados cerca unos de otros, señalan conceptos estrechamente relacionados que aparecen con mayor frecuencia en el análisis. Un ejemplo de esto son términos como *study*, *data*, *analysis*, *Brazil* y *region*, los cuales destacan como temas clave en el conjunto de artículos analizados.

- Estudios en Brasil y regiones específicas: El término “Brazil” se destaca como uno de los puntos centrales en el mapa, lo que indica una fuerte representación de investigaciones medioambientales centradas en este país. Términos como “región” y “state” sugieren que muchos de estos estudios están orientados hacia niveles geográficos específicos, lo que refleja un enfoque territorial en la investigación.
- Análisis de datos: La frecuencia de términos como “data”, “analysis” y “model” señala que una parte considerable de los estudios utiliza enfoques cuantitativos y modelos estadísticos para analizar datos medioambientales abiertos. Esto subraya el rol fundamental de las herramientas de análisis en la comprensión de fenómenos ambientales.
- Pandemia y salud pública: Términos como “covid”, “health” e “impact” ponen de relieve un notable interés en la interrelación entre el medio ambiente y la salud pública, probablemente relacionado con los efectos de la pandemia de COVID-19 en el ámbito ambiental y social.
- Estudios ecológicos y mortalidad: Términos como “disease”, “mortality” y “ecological study” indican la existencia de investigaciones que examinan cómo el entorno natural impacta la salud de las poblaciones, incluyendo la mortalidad y la incidencia de enfermedades.

- Conceptos medioambientales: Palabras como “environment”, “species” y “research” reflejan la relevancia de estudios relacionados con la ecología y el medio ambiente, probablemente vinculados a temas como la biodiversidad y los efectos del cambio climático en los ecosistemas.

Figura 12

Visualización del análisis de temas basados en texto de datos



Nota: Autores (2024).

En síntesis, el análisis de textos sobre datos abiertos medioambientales destaca varias áreas clave en la literatura científica. Los estudios se enfocan en la aplicación de métodos cuantitativos y en el uso de herramientas estadísticas para analizar grandes conjuntos de datos, frecuentemente con un enfoque geográfico y regional, como es el caso de Brasil. Asimismo, la interacción entre el medio ambiente y la salud pública, exacerbada por la pandemia de COVID-19, se identifica como un tema relevante.

Finalmente, el análisis muestra un interés considerable en la conservación de la biodiversidad y el monitoreo de ecosistemas, campos en los que los datos abiertos desempeñan un papel fundamental. Estos hallazgos subrayan la relevancia de los datos abiertos para el avance de la investigación y la resolución de los desafíos medioambientales actuales y futuros.

CAPITULO 4

ANÁLISIS DE DATOS AVANZADOS



Análisis de datos avanzados

4.1. Técnicas de machine learning

El análisis realizado en este estudio se apoyó en el uso de técnicas de machine learning como un complemento poderoso a las metodologías tradicionales de bibliometría. Mientras que la bibliometría clásica se enfoca en el análisis cuantitativo de la literatura científica a través de métricas como el recuento de publicaciones y citas, la incorporación de machine learning permite una exploración más profunda de los datos textuales. Estas técnicas hacen posible identificar patrones ocultos, descubrir relaciones complejas entre conceptos y organizar grandes volúmenes de información de manera sistemática, ampliando significativamente el alcance del análisis. El uso de machine learning facilita la clasificación automática de los contenidos temáticos, ayudando a desentrañar tendencias emergentes y proporcionando una base sólida para la toma de decisiones informadas en la investigación científica.

Para implementar estas técnicas avanzadas de análisis, se eligió Python como el lenguaje de programación principal, dada su versatilidad y amplia adopción en la comunidad de machine learning. Python cuenta con un ecosistema robusto de bibliotecas especializadas como scikit-learn, pandas y numpy, que permiten realizar desde tareas de procesamiento de datos hasta modelado y visualización de resultados con gran eficiencia. Su facilidad de uso y la extensa documentación que lo respalda lo convierten en una opción ideal para manejar el gran volumen de datos textuales involucrados en este estudio, optimizando tanto el preprocesamiento como el análisis detallado de los resúmenes científicos.

En este estudio se llevó a cabo un análisis exhaustivo de un conjunto de datos bibliográficos utilizando técnicas de procesamiento de lenguaje natural (PLN) y algoritmos de machine learning. El objetivo principal fue identificar y organizar los temas predominantes en la investigación sobre medio ambiente y el uso de datos abiertos, proporcionando una visión clara y estructurada del panorama actual en este campo. Este enfoque buscó facilitar la interpretación y el análisis

de grandes volúmenes de datos textuales, permitiendo una mejor comprensión de las tendencias y líneas de investigación relevantes.

Se utilizaron algoritmos como Latent Dirichlet Allocation (LDA) para identificar las temáticas subyacentes en los resúmenes de los artículos científicos, clasificando los contenidos en categorías temáticas coherentes. Posteriormente, se aplicó el algoritmo de K-Means para agrupar los estudios en clústeres según sus similitudes temáticas, lo cual ayudó a visualizar las conexiones entre las distintas áreas de investigación. Estos métodos fueron elegidos por su capacidad para manejar datos complejos y ofrecer una organización clara del conocimiento disponible.

Adicionalmente, se empleó la técnica de t-SNE para proyectar los resultados en un espacio de dos dimensiones, proporcionando una representación gráfica que facilita la interpretación de cómo se agrupan los artículos según sus similitudes. Esta visualización permitió observar patrones temáticos y relaciones entre los clústeres identificados, ayudando a resaltar las áreas de mayor concentración de estudios y las posibles conexiones entre ellas.

El propósito general del análisis fue ofrecer una herramienta eficaz para explorar y entender mejor las dinámicas de investigación en el campo de la ciencia ambiental y los datos abiertos. Manteniendo un enfoque riguroso, pero no conclusivo, se buscó proporcionar una base sólida para que futuros estudios puedan profundizar en las áreas identificadas, fomentando así un entendimiento más organizado y sistemático del conocimiento científico en esta área.

4.2. Análisis Temático con LDA y K-Means

El código en Python fue diseñado para analizar el conjunto de datos bibliográficos extraído, específicamente los resúmenes de artículos científicos, utilizando técnicas de procesamiento de lenguaje natural (PLN) y algoritmos de machine learning. La primera acción que se realizó fue la carga un archivo de que contenía los datos bibliográficos. A continuación, el código seleccionó la columna de los resúmenes (columna "AB") y los preparó para su análisis, eliminando los valores nulos, para continuar con el preprocesamiento del texto.

Posteriormente, se aplicó la técnica de vectorización utilizando el CountVectorizer. Este método permitió convertir los resúmenes textuales en una matriz de frecuencias, donde cada palabra del texto se representó numéricamente en función de su frecuencia de aparición en los resúmenes. Se establecieron parámetros para evitar palabras demasiado comunes o demasiado raras mediante las opciones `max_df=0.95` y `min_df=2`, respectivamente.

Una vez vectorizados los resúmenes, se implementó el modelo de Latent Dirichlet Allocation (LDA), que es una técnica de modelado de temas. Este modelo identificó cinco temas subyacentes en los resúmenes, basándose en la probabilidad de que cada documento perteneciera a un tema específico. Las palabras clave más representativas de cada tema fueron extraídas y se generó una tabla que las listaba, destacando las principales palabras que definían cada uno de los cinco temas. En la Tabla 15 se puede observar cómo palabras como *species*, *environmental* y *conservation* se destacaron en uno de los temas, indicando un enfoque hacia la investigación ambiental y la conservación de especies.

Tabla 16

Temas LDA

Tema	Palabras
0	species, environmental, data, forest, areas, conservation, ecological, use, open, public
1	health, study, mortality, Brazil, data, cases, 19, public, years, COVID
2	environmental, data, public, social, research, results, analysis, management, economic, study
3	study, data, health, environmental, risk, using, models, air, results, exposure
4	environmental, water, data, public, areas, use, study, soil, information, results

Nota: Autores (2024).

El siguiente paso fue aplicar el algoritmo de K-Means para agrupar los resúmenes en cinco clústeres, basándose en los temas identificados previamente por el modelo LDA. Los clústeres asignados a cada artículo se añadieron como una nueva columna en el DataFrame original. Con esto, el código permitió identificar la pertenencia de cada artículo a un clúster específico, lo que facilitó una mejor organización y análisis del contenido temático.

Luego, se desarrolló una función para generar resúmenes temáticos para cada clúster, en la que se utilizaron los resúmenes y títulos de los artículos. Esta función permitió vectorizar el texto combinado de los artículos en cada clúster y extraer las palabras más frecuentes. Los resultados fueron organizados en un DataFrame con las palabras más comunes y sus frecuencias, lo que ofreció un resumen de los términos más representativos de cada clúster, ver Tabla 17. En uno de los clústeres, las palabras más frecuentes fueron health, study y mortality, indicando que los artículos dentro de ese grupo tenían un enfoque en temas de salud pública y mortalidad. Esto proporcionó una visión clara de los temas predominantes dentro de cada conjunto de artículos.

Tabla 17

Palabras más frecuentes por clúster

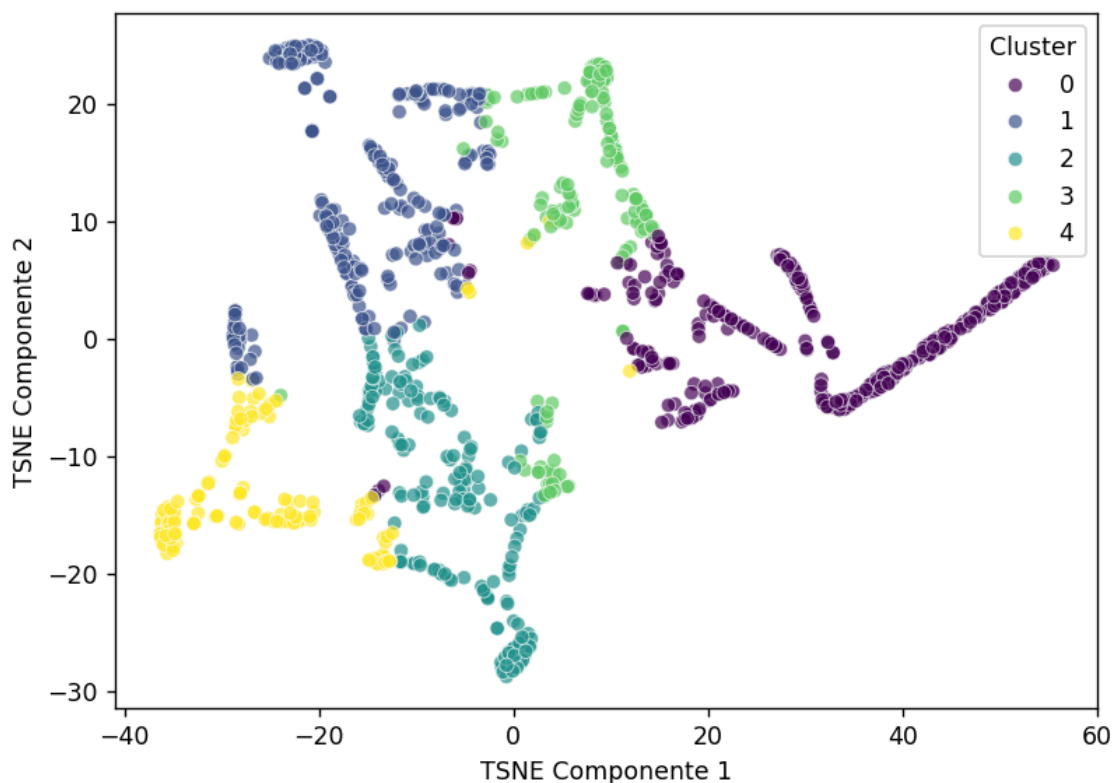
Clus ter	Palabras representativas	Frecuencia
0	health, study, mortality, brazil, data, 19, cases, covid, public, spatial	602, 510, 489, 481, 365, 332, 312, 298, 290, 286
1	environmental, data, public, health, social, study, analysis, results, research, based	455, 324, 284, 273, 239, 219, 203, 191, 183, 159
2	environmental, data, water, public, study, health, use, results, areas, information	305, 282, 211, 190, 187, 149, 144, 139, 137, 136
3	study, data, health, environmental, using, risk, public, results, brazil, dengue	223, 198, 177, 164, 139, 135, 122, 121, 97, 93
4	species, environmental, data, forest, areas, study, public, ecological, open, conservation	316, 146, 145, 105, 104, 96, 95, 93, 92, 85

Nota: Autores (2024).

Para visualizar los resultados, se utilizó la técnica de reducción de dimensionalidad t-SNE (Stochastic Neighbor Embedding), que permitió proyectar los resúmenes agrupados en un espacio de dos dimensiones. Esto facilitó la visualización gráfica de los clústeres, mostrando cómo los artículos estaban agrupados en función de su similitud temática. El gráfico generado evidenció la distribución de los artículos en función de los clústeres de K-Means, ver Figura 13.

Figura 13

Clustering de Artículos Científicos usando LDA + K-Means



Nota: Autores (2024).

4.3. Interpretación de los clústeres

La interpretación de los clústeres se realizó con el objetivo de obtener una visión más profunda y detallada de los patrones temáticos presentes en la literatura científica analizada. Esta fase del análisis es fundamental para ir más allá de la simple agrupación de artículos y comprender las conexiones y relaciones entre los conceptos identificados. El propósito de esta interpretación es revelar las dinámicas internas y la estructura temática de la investigación en campos como el medio ambiente y el uso de datos abiertos, proporcionando una base sólida para el análisis de las tendencias científicas en estas áreas.

La justificación para llevar a cabo esta interpretación detallada radica en la necesidad de transformar datos textuales y numéricos en conocimiento significativo. Al analizar las palabras clave y los títulos dentro de cada clúster, se busca asegurar que los temas identificados reflejen con precisión las áreas de

investigación predominantes y que la agrupación realizada por los algoritmos de machine learning sea coherente y válida. Este enfoque permite validar los resultados del análisis y facilita la identificación de posibles tendencias emergentes, orientando así futuros estudios y esfuerzos en el campo de la investigación científica.

Para realizar la interpretación, primero se revisaron las palabras clave predominantes, las cuales reflejaron los conceptos principales abordados por los artículos dentro de cada clúster. Estas palabras clave permitieron identificar de forma rápida los temas recurrentes, proporcionando una visión general sobre el enfoque temático de los artículos agrupados.

Luego, se observó la frecuencia de las palabras clave. Las palabras con mayor frecuencia indicaron los conceptos más discutidos en los artículos y fueron representativas del núcleo temático del clúster. Cuanto más alta fue la frecuencia de una palabra clave, mayor fue su relevancia dentro del clúster.

Se procedió a combinar las palabras clave principales y secundarias para obtener una comprensión más completa del contexto. Se analizaron las relaciones entre las palabras clave para formar una visión coherente del tema tratado por los artículos, evitando considerarlas de forma aislada.

Se revisaron los títulos de los artículos asignados a cada clúster, lo que brindó mayor claridad y contexto a las palabras clave. Esta comparación entre títulos y palabras clave permitió afinar la interpretación del tema y confirmar la coherencia entre los artículos y el enfoque temático del clúster, ver Tabla 18.

Tabla 18

Interpretación de Clústeres

Título del cluster	Palabras clave	Palabras con frecuencia alta	Interpretación
0 Salud y mortalidad, con un enfoque en Brasil y COVID-19	health, study, mortality, brazil, data, cases, 19, public, years, covid	health, mortality, covid, brazil	Estudios relacionados con salud pública, específicamente mortalidad, COVID-19 y estudios en Brasil.
1 Estudios ambientales,	environmental, data, public, health, social, study,	environmen tal, public,	Explora salud pública en el contexto de factores ambientales

sociales y de salud pública	research, management, economic	results, health	social, water, areas, soil, health	y sociales, con enfoque en gestión y economía.
2 Estudios sobre el agua, salud y uso de recursos ambientales	environmental, data, water, public, study, health, areas, soil, information, results	water, areas, soil, health		Aborda la gestión del agua, salud pública y el uso de suelos en el contexto ambiental.
3 Estudios de riesgos ambientales y modelos de salud	study, data, health, environmental, using, risk, models, air, results, exposure	risk, models, exposure, air		Estudios de riesgo ambiental, modelos predictivos y exposición al aire, relacionados con la salud.
4 Conservación, biodiversidad y estudios ecológicos	species, environmental, data, forest, areas, conservation, ecological, use, open, public	species, forest, conservatio, n, ecological		Estudios sobre conservación, biodiversidad y ecosistemas forestales, con énfasis en especies y ecología.

Nota: Autores (2024).

La interpretación de los clústeres reveló algunos hallazgos significativos sobre las temáticas abordadas en los artículos analizados. El primer clúster se centró en temas de salud pública, destacando específicamente la mortalidad y el impacto de la COVID-19 en Brasil. Las palabras clave predominantes como "health," "mortality," "covid" y "brazil" subrayan la relevancia de los estudios que investigan la relación entre la pandemia y sus efectos en la población brasileña, indicando un fuerte enfoque en la salud y las políticas públicas en respuesta a esta crisis.

El segundo y tercer clústeres abordaron interacciones entre factores ambientales y sociales con temas de salud pública y gestión de recursos. El clúster dos se enfocó en la gestión del agua y el uso de suelos en el contexto de la salud pública, con palabras clave como "water," "areas," y "soil" que sugieren investigaciones sobre la sostenibilidad de los recursos naturales y su impacto en la calidad de vida. Mientras tanto, el tercer clúster se concentró en estudios sobre riesgos ambientales y modelos predictivos, con un énfasis claro en la exposición al aire y otros factores de riesgo para la salud.

El último clúster se destacó por su enfoque en la conservación, biodiversidad y estudios ecológicos, resaltando la importancia de las especies y los ecosistemas forestales. Las palabras clave "species," "forest," "conservation" y "ecological" indican un interés en la protección de la biodiversidad y la preservación de los

ecosistemas naturales. Este clúster refleja la investigación centrada en la ecología y la necesidad de estrategias para conservar la flora y fauna en sus entornos naturales, destacando la conexión entre los estudios ambientales y las políticas de conservación.

CAPITULO 5

ÉVALUACIÓN DE LAS FUENTES DE ACCESO ABIERTO



Evaluación de las fuentes de acceso abierto

5.1. Justificación del análisis de fuentes primarias

En el marco del análisis bibliométrico realizado, se consideró fundamental complementar el estudio con una evaluación detallada del uso de fuentes primarias en las investigaciones más influyentes. Este enfoque surge de la necesidad de comprender no solo las tendencias temáticas y colaborativas en la literatura científica, sino también de examinar la calidad y accesibilidad de los datos utilizados en los estudios que más impactan el campo. Evaluar las fuentes primarias es esencial para entender cómo se recopila y utiliza la información en la investigación científica y para identificar las prácticas que promueven la transparencia y la reproducibilidad en el uso de datos abiertos.

La decisión de analizar una muestra de 50 artículos se basó en su relevancia dentro del corpus bibliográfico estudiado, seleccionados mediante una métrica de citación normalizada por la edad del artículo. Esta métrica se diseñó para garantizar que los estudios más influyentes, aquellos que han mantenido su relevancia a lo largo del tiempo, fueran incluidos en el análisis. Al aplicar este enfoque, se buscó no solo identificar a los estudios con mayor visibilidad, sino también resaltar la manera en que estos trabajos hacen uso de las fuentes primarias de datos abiertos, destacando su papel en la generación y validación del conocimiento científico.

El análisis de estas fuentes primarias se justifica, además, por el compromiso con la reproducibilidad y la integridad de la investigación científica. Entender cómo se accede y utiliza la información primaria en los estudios más citados ofrece una visión crítica sobre las prácticas actuales en la ciencia abierta. Este enfoque permite detectar posibles limitaciones en la disponibilidad y calidad de los datos reportados, así como en las estrategias de documentación y accesibilidad empleadas por los autores. Evaluar estos aspectos es crucial para mejorar las normativas de transparencia en la publicación de resultados científicos.

La metodología utilizada para llevar a cabo esta evaluación combina un análisis crítico del contenido de los artículos con la verificación directa de la accesibilidad de los datos en repositorios abiertos. Esta estrategia se diseñó para asegurar que los recursos utilizados en los estudios sean efectivamente accesibles y estén descritos de manera clara, permitiendo a otros investigadores replicar y validar los hallazgos de manera eficiente. La revisión detallada de resúmenes, métodos y secciones adicionales de cada artículo proporcionó una base sólida para identificar y evaluar el uso de datos abiertos en el contexto de la investigación científica.

Esta evaluación busca contribuir a una comprensión más profunda del impacto que las prácticas de datos abiertos tienen en la calidad y la transparencia de la investigación. Al identificar las fortalezas y debilidades en la manera en que se reportan las fuentes de datos primarias, el estudio proporciona una guía para futuras mejoras en la documentación y accesibilidad de la información en publicaciones científicas, fomentando así una cultura de investigación más abierta y colaborativa.

5.2. Procedimiento realizado

Para obtener una muestra representativa de las prácticas en el uso de fuentes primarias en los estudios evaluados, se escogieron 50 trabajos de investigación. La selección se basó en una tasa de citación normalizada por la edad del artículo, calculada al dividir el total de citaciones para la diferencia entre 2024 y el año de publicación. Este método permitió identificar los estudios con mayor influencia, considerando tanto su popularidad como su relevancia temporal, para luego analizar las fuentes primarias empleadas en cada uno.

Si bien esta métrica facilitó la selección inicial de los artículos, es importante reconocer sus limitaciones. No se consideraron factores cualitativos como la calidad de las citaciones o el impacto real de la investigación, y el cálculo pudo verse afectado por sesgos como la tendencia a que los artículos más citados reciban aún más atención, independientemente de su calidad intrínseca.

Ya con esa selección, para determinar las fuentes de datos primarias de acceso abierto en los estudios elegidos, se empleó un enfoque metodológico que combina la lectura crítica del contenido con la verificación de la disponibilidad de los datos.

En primer lugar, se realizó una lectura exhaustiva del resumen y la sección de métodos de cada artículo, buscando frases clave que indicaran la naturaleza de los datos primarios, como "datos recopilados", "participantes reclutados" o "muestras ambientales". Además, se identificaron las bases de datos o repositorios mencionados en el texto, verificando su acceso abierto y la disponibilidad de los datos primarios en cuestión.

En segundo lugar, se examinó el documento completo para recopilar información adicional sobre las fuentes de datos, ya que esta información puede estar presente en diferentes secciones del artículo. Finalmente, se corroboró la accesibilidad de los datos primarios en los sitios web referenciados, confirmando así su disponibilidad pública.

Este proceso, aunque riguroso, está sujeto a ciertas limitaciones. La interpretación de información compleja o ambigua puede ser un desafío, y la falta de claridad en la descripción de las fuentes de datos en algunos artículos puede dificultar la identificación precisa. No obstante, el objetivo principal es proporcionar una descripción de las fuentes de datos primarias de acceso abierto utilizadas en cada estudio, reconociendo las posibles limitaciones como los sesgos inconscientes que podrían influir en la interpretación de la información, y el conocimiento previo sobre el tema que condiciona la identificación de las fuentes.

5.3. Lo encontrado en la evaluación

El análisis detallado de 50 artículos científicos reveló una notable diversidad en los enfoques metodológicos y en las fuentes de datos utilizadas, destacando el predominio de estudios empíricos que emplean datos primarios para sustentar sus hallazgos. Esta tendencia evidencia el compromiso de la comunidad científica con la obtención de datos directos y cuantificables, que son

fundamentales para respaldar conclusiones robustas y para avanzar en el conocimiento en diversas áreas de investigación. Desde estudios observacionales hasta ensayos clínicos y revisiones sistemáticas, los artículos analizados reflejan una amplia gama de metodologías que enriquecen el campo científico con datos precisos y específicos.

Una de las observaciones más relevantes es el uso extendido de fuentes de datos primarias provenientes de tecnologías avanzadas como imágenes satelitales (Landsat-8 y Sentinel-2), utilizadas para monitorear la deforestación, la contaminación del aire y la calidad del agua. Estos recursos permiten un análisis ambiental detallado, crucial para evaluar el impacto de las actividades humanas y del cambio climático en diferentes ecosistemas. Además, el empleo de muestras ambientales tomadas de aguas residuales, ríos y otras fuentes naturales ha sido clave para la detección de contaminantes y patógenos, lo que resalta el papel fundamental de la investigación empírica en la protección de la salud pública y el entorno natural.

El estudio también identificó el uso de modelos predictivos y experimentales como una herramienta valiosa para analizar el impacto del cambio climático y para estudiar respuestas biológicas en condiciones controladas. Modelos como los experimentales murinos y las simulaciones climáticas se utilizaron para investigar desde la respuesta inmune hasta los efectos del cambio climático sobre la biodiversidad. Estos métodos proporcionan una base sólida para desarrollar estrategias de mitigación y adaptación frente a las crisis ambientales y de salud emergentes, destacando la importancia de la predicción y la planificación en la ciencia moderna.

En paralelo, se encontraron estudios que utilizan datos secundarios provenientes de bases de datos globales bien establecidas, como el Global Burden of Disease Study (GBD) y la base de datos InvaCost. Estas fuentes ofrecen un contexto agregado y comparativo que es fundamental para los análisis epidemiológicos y económicos a gran escala. Aunque no se recogen directamente por los autores, estos datos secundarios son esenciales para examinar tendencias globales y evaluar la carga de enfermedades, así como para estudiar los impactos económicos de las invasiones biológicas.

La multidisciplinariedad de las fuentes de datos es otro aspecto que resalta en el análisis. Los estudios abordan una amplia variedad de enfoques, desde la ciencia de la sostenibilidad y la ecología hasta la salud pública y la demografía, utilizando datos gubernamentales, encuestas en línea y especímenes de museo para explorar la distribución geográfica de especies. Esta diversidad de fuentes y metodologías permite a los investigadores abordar problemas complejos desde múltiples ángulos, enriqueciendo la calidad y la profundidad de sus conclusiones.

Además, es significativo el compromiso con la ciencia abierta y la transparencia en la investigación. Un alto porcentaje de estudios utiliza datos de acceso abierto, lo que no solo facilita la reproducibilidad de los resultados, sino que también promueve la colaboración interdisciplinaria y la verificación independiente de los hallazgos. Este enfoque hacia la apertura de datos es crucial para la evolución científica, ya que permite a otros investigadores construir sobre el trabajo existente y avanzar en la resolución de desafíos globales.

Sin embargo, se identificaron áreas importantes de mejora en cuanto a la claridad y precisión con la que se reportan las fuentes de datos en las publicaciones científicas. En varios casos, la falta de detalles específicos sobre las fuentes primarias utilizadas dificultó la evaluación completa de la calidad y la fiabilidad de los datos empleados. Esta deficiencia subraya la necesidad de una mayor rigurosidad en la documentación de los métodos de recolección y en la descripción de los recursos utilizados en los estudios.

Los estudios que se apoyan en revisiones de la literatura y enunciados de consenso, aunque no aportan datos primarios directos, cumplen un papel esencial al sintetizar y analizar información existente. Estos enfoques permiten consolidar el conocimiento previo y proporcionan una base sólida para el desarrollo de nuevas hipótesis y teorías, demostrando la importancia de las revisiones sistemáticas para identificar vacíos en la investigación y guiar futuros estudios.

El análisis global de estos artículos muestra una clara tendencia hacia enfoques metodológicos diversificados y la integración de tecnologías avanzadas en la investigación científica. El predominio de estudios empíricos, junto con el uso de

modelos predictivos y datos secundarios, destaca un enfoque equilibrado que combina observación directa y análisis teórico. Esta combinación es fundamental para abordar de manera integral los problemas complejos que enfrenta la sociedad moderna, como el cambio climático, la salud pública y la sostenibilidad.

El estudio de estas 50 publicaciones científicas revela un panorama dinámico y en constante evolución en la investigación, donde la diversidad metodológica, el uso de fuentes de datos primarias y la apuesta por la ciencia abierta son elementos clave. Estas prácticas no solo enriquecen el conocimiento científico, sino que también fomentan un enfoque colaborativo para enfrentar los desafíos globales, promoviendo un futuro más sostenible y basado en la evidencia, ver Tabla 19.

Tabla 19

Datos primarios en estudios científicos

ID de archivo	Tipo de estudio	Fuente de datos primaria	Observaciones
(Ong et al., 2023)	Estudio observacional	N/A	Utiliza datos secundarios del Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) para analizar la carga mundial de enfermedades.
(Zhao et al., 2021)	Estudio observacional	Multi-Country Multi-City (MCC) Collaborative Research Network	Datos sobre la contaminación del aire en múltiples ciudades
(Shaw et al., 2021)	Enunciado de consenso	N/A	No utiliza datos primarios, es una revisión de estudios existentes.
(Canabarro et al., 2023)	Estudio ecológico	Datos del gobierno (no especifica cuál) y de la Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO)	Datos sobre la producción y el consumo de alimentos
(Chen et al., 2021)	Estudio observacional	Multi-Country Multi-City (MCC) Collaborative Research Network	Datos sobre la contaminación del aire en múltiples ciudades
(Filho et al., 2021)	Estudio descriptivo	Encuesta en línea (no especifica nombre)	Datos sobre la percepción pública de la ciencia
(Kass et al., 2023)	Nota de software	N/A	No utiliza datos primarios, describe una nueva aplicación de software.

Capítulo VI: Discusión, conclusiones y recomendaciones

(Oviedo-Trespalcios et al., 2023)	Revisión de la literatura	N/A	No utiliza datos primarios, utiliza ChatGPT para analizar información de otras fuentes.
(Urban & Nakada, 2021)	Estudio de caso	Sistema Nacional de Información Sanitaria de Brasil (SNIS) y Asociación Empresarial para el Reciclaje (CEMPRE)	Datos sobre la gestión de residuos en Brasil
(Viljur et al., 2022)	Estudio de caso	Serie de estudios de casos (no especifica cuáles)	Datos sobre la ciencia de la sostenibilidad en diferentes contextos
(Vargas-Munoz et al., 2021)	Artículo de revista	N/A	No utiliza datos primarios, es una revisión de estudios existentes sobre la gestión de residuos sólidos urbanos.
(Local Burden of Disease HIV Collaborators, 2021)	Estudio observacional	N/A	Utiliza datos secundarios del Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) para analizar la carga mundial de enfermedades relacionadas con la contaminación del aire.
(Yuan et al., 2022)	Artículo de revista	N/A	No utiliza datos primarios, discute el papel de la inteligencia artificial en la salud global.
(Mansouri et al., 2021)	Estudio de modelado	Modelos predictivos (no especifica cuáles)	Datos sobre el impacto del cambio climático en la biodiversidad.
(Parra et al., 2021)	Artículo de revista	N/A	No utiliza datos primarios, analiza la ética en la investigación en salud pública.
(Heinrich et al., 2022)	Estudio cualitativo	Entrevistas a personas con enfermedades crónicas (no especifica el número)	Datos sobre las experiencias de las personas con enfermedades crónicas.
(Friedrich et al., 2021)	Estudio observacional	Departamento de Informática del Sistema de Salud Pública de Brasil (DATASUS)	Datos sobre la mortalidad infantil en Brasil.
(Pereira et al., 2021)	Estudio observacional	Imágenes de satélite Landsat-8	Datos sobre la deforestación en la Amazonía.
(Brack et al., 2022)	Artículo de política	N/A	No utiliza datos primarios, describe una nueva política de salud pública.
(Mestanza-Ramón et al., 2023)	Estudio experimental	Muestras de agua de ríos y arroyos (no especifica la ubicación)	Datos sobre la calidad del agua y la presencia de contaminantes.

Capítulo VI: Discusión, conclusiones y recomendaciones

(Stapleton et al., 2023)	Revisión sistemática	N/A	No utiliza datos primarios, es una revisión de estudios existentes sobre la contaminación del aire y la salud.
(Gómez-Ochoa et al., 2022)	Estudio observacional	N/A	Utiliza datos secundarios del Global Burden of Disease Study para analizar la mortalidad por enfermedades no transmisibles.
(Crippa et al., 2021)	Ensayo clínico aleatorizado	Datos del ensayo clínico (no especifica nombre ni participantes)	Datos sobre la eficacia de un nuevo tratamiento para la diabetes.
(Magalhães et al., 2023)	Estudio ecológico	Ministerio de Salud de Brasil y proyecto MapBiomias	Datos sobre la deforestación y la salud en Brasil.
(Augusto et al., 2022)	Estudio experimental	Muestras de aguas residuales (no especifica la ubicación)	Datos sobre la presencia de contaminantes emergentes en el agua.
(Silva et al., 2023)	Estudio observacional	Datos de los satélites TRMM y MODIS	Datos sobre la precipitación y la temperatura en la Amazonía.
(Lense et al., 2023)	Estudio descriptivo	Mapas de suelos (no especifica la ubicación ni la fuente)	Datos sobre las propiedades del suelo en una región determinada.
(Wallace et al., 2023)	Revisión de la literatura	N/A	No utiliza datos primarios, es una revisión de estudios existentes sobre el impacto del cambio climático en la agricultura.
(Camalan et al., 2022)	Estudio observacional	Imágenes de satélite Sentinel-2	Datos sobre la calidad del agua en embalses.
(Figueiredo Cerqueira et al., 2023)	Estudio experimental	Datos de un modelo experimental murino	Datos sobre la eficacia de una vacuna contra la leishmaniasis.
(Zaneti et al., 2022)	Estudio de caso	Datos de paneles fotovoltaicos instalados y de la operación de un autobús eléctrico en las instalaciones de Unicamp	Datos sobre la eficiencia energética de un sistema de generación y almacenamiento de energía solar.
(Nascimento et al., 2021)	Revisión taxonómica	Especímenes de museo (no especifica cuáles) y modelos de nicho ecológico (no especifica cuáles)	Datos sobre la distribución geográfica de una especie de lagartija.
(Nathan et al., 2021)	Estudio de cohorte	Datos de células T de memoria de 259 individuos en una cohorte de progresión de tuberculosis peruana	Datos sobre la respuesta inmune a la tuberculosis.

Capítulo VI: Discusión, conclusiones y recomendaciones

(Abrahão et al., 2021)	Estudio descriptivo	Muestras ambientales recolectadas en un área urbana de Brasil (no especifica la ubicación exacta)	Datos sobre la presencia de SARS-CoV-2 en superficies públicas.
(Marques et al., 2021)	Estudio transversal	Cuestionario en línea a residentes de Río de Janeiro, Brasil	Datos sobre la importancia de los jardines domésticos para el bienestar mental durante la pandemia de COVID-19.
(Boing et al., 2022)	Estudio ecológico	Centro Nacional de Estadísticas de Salud (NCHS) de Estados Unidos, Encuesta de la Comunidad Estadounidense y modelos de aprendizaje automático de química atmosférica (no especifica cuáles)	Datos sobre la contaminación del aire y la mortalidad en Estados Unidos.
(Martins et al., 2021)	Estudio de caso	Imágenes aéreas RGB de un área urbana en el bioma del Cerrado	Datos sobre la segmentación semántica de la copa de los árboles.
(Stoler et al., 2021)	Estudio transversal	Datos de 8,297 hogares en 29 sitios en 23 países de ingresos bajos y medianos	Datos sobre la inseguridad hídrica en hogares.
(Pachito et al., 2021)	Revisión sistemática y meta-análisis	Datos de 14 estudios de cohorte (no especifica cuáles)	Datos sobre el efecto de la exposición a largas jornadas de trabajo en el consumo de alcohol.
(Shah et al., 2021)	Enunciado de consenso	N/A	No utiliza datos primarios, es una revisión de estudios existentes sobre el manejo de la sepsis.
(Claro et al., 2021)	Estudio observacional	Muestras de aguas residuales (no especifica la ubicación)	Datos sobre la presencia de SARS-CoV-2 en aguas residuales para la vigilancia epidemiológica.
(Arruda et al., 2021)	Estudio de caso	Imágenes de satélite Landsat-8	Datos sobre la cartografía de cicatrices de quemaduras en el Pantanal brasileño.
(Hacket-Pain et al., 2022)	Estudio observacional	Datos de una serie de estudios de observación (no especifica cuáles)	Datos sobre el esfuerzo reproductivo de las plantas en respuesta a la herbivoría.
(Ahmed et al., 2023)	Revisión de la literatura	Base de datos InvaCost	Datos sobre los costos económicos de las invasiones biológicas. Aunque la base de datos es una recopilación, los

			datos de costos en sí mismos pueden considerarse primarios.
(Schneider et al., 2021)	Estudio ecológico	Sistema Nacional de Información Sanitaria de Brasil (SNIS)	Datos sobre mordeduras de serpientes en Brasil.
(Rodríguez-Villamizar et al., 2021)	Estudio ecológico	Modelo Copernicus Atmospheric Monitoring Service Reanalysis (CAMSRA)	Datos sobre la contaminación del aire y la mortalidad por COVID-19. El modelo en sí no es una fuente primaria, pero los datos generados por el modelo para este estudio sí lo son.
(Del-Aguila-Arcatales et al., 2022)	Estudio observacional	Índice de los Objetivos de Desarrollo Sostenible (ODS), Informe de Competitividad Global e Informe del Monitor de Emprendimiento Global	Datos sobre la influencia de los ODS en la continuidad del emprendimiento. Estos informes contienen datos agregados de diversas fuentes, por lo que se consideran secundarios.
(Shao et al., 2022)	Estudio observacional	Datos de la calidad del aire y las condiciones meteorológicas en Wuhan, China (no especifica la fuente de los datos)	Datos sobre la correlación entre la mortalidad por COVID-19 y la exposición al aire libre de PM2.5.
(Colón-González et al., 2021)	Estudio de modelado	Datos de casos de dengue en Vietnam (no especifica la fuente de los datos)	Datos para la predicción estacional probabilística del dengue.
(Ruel et al., 2021)	Ensayo clínico	Datos de un ensayo clínico (no especifica nombre ni participantes)	Datos sobre la farmacocinética, seguridad, tolerabilidad y actividad antiviral de las tabletas dispersables de dolutegravir en bebés y niños con VIH-1.

Nota: Autores (2024).

CAPITULO 6

DISCUSIÓN, CONCLUSIONES Y RECOMENDACIONES



Discusión, conclusiones y recomendaciones

6.1. Discusión de los hallazgos más relevantes

El análisis bibliométrico de las publicaciones científicas sobre acceso abierto a la información ambiental y ecológica en la región amazónica durante el período 2021-2023 revela un crecimiento sostenido, aunque moderado, en la producción científica. Este aumento refleja un interés creciente y sostenido por la investigación ambiental en la Amazonía, impulsado en gran parte por la disponibilidad de datos abiertos. Brasil se destaca como el principal productor de conocimiento en la región, liderando en número de citaciones y publicaciones, lo que evidencia su papel crucial en la comunidad científica global. No obstante, la participación significativa de países fuera de la región amazónica, como Estados Unidos y Australia, subraya que la investigación sobre la Amazonía tiene un alcance verdaderamente global, facilitada por colaboraciones internacionales.

A nivel institucional, la Universidad de São Paulo (USP) y otras universidades brasileñas emergen como actores clave en la producción científica de acceso abierto. Revistas de alto impacto, como el *International Journal of Environmental Research and Public Health*, juegan un papel fundamental en la difusión de estas investigaciones, lo que señala la importancia de algunos canales específicos para la diseminación del conocimiento. La concentración de publicaciones en un número reducido de revistas y la alta colaboración internacional indican la necesidad de una mayor integración de esfuerzos globales para mejorar el acceso abierto a datos ambientales y avanzar en la gestión sostenible de los recursos naturales en la Amazonía.

El análisis de redes realizado mediante VOSviewer ofrece una visualización detallada de las dinámicas de colaboración entre autores, instituciones y países en el campo de los datos abiertos sobre el medio ambiente. Brasil ocupa una posición central en la red de coautoría y co-citación, consolidándose como un nodo principal. Sin embargo, la presencia destacada de países no amazónicos, como Estados Unidos y Australia, refuerza la naturaleza internacional de la investigación ecológica, potenciada por redes globales de colaboración. A nivel

institucional, las universidades brasileñas, lideradas por la USP, juegan un rol destacado, lo que reafirma la relevancia de este país en la investigación ambiental.

El análisis temático y de co-ocurrencia de palabras clave subraya que temas como la salud pública, el impacto de la pandemia de COVID-19 y el análisis espacial dominan las investigaciones recientes. Palabras clave como "mortalidad" y "cambio climático" sugieren una fuerte conexión entre la investigación ambiental y la salud humana, lo que refleja una creciente preocupación por los efectos del cambio climático en la población. La visualización de redes de co-citación destaca la influencia de revistas de alto perfil como *PLOS ONE* y *The Lancet*, que han jugado un papel crucial en la difusión de investigaciones ambientales de acceso abierto.

El análisis estadístico utilizando el paquete iGraph complementa el análisis de redes realizado en VOSviewer, proporcionando una perspectiva cuantitativa de la estructura de la red. Las métricas de centralidad, intermediación y cercanía revelan que Brasil, seguido de cerca por Estados Unidos y Australia, ocupa posiciones estratégicas dentro de las redes de colaboración internacional. Estos países no solo son productores clave de conocimiento, sino que también actúan como nodos cruciales en la comunicación y difusión de investigaciones ambientales. Las posiciones centrales de estos países en la red refuerzan su rol en facilitar el flujo de información científica.

A nivel individual, investigadores como Arcêncio y Neto destacan por su centralidad en la red de coautoría, reafirmando el papel prominente de Brasil en la investigación ecológica. Sin embargo, autores como Santos sobresalen en términos de intermediación, actuando como "puentes" que facilitan la conexión entre distintos grupos de investigación, lo que es esencial para impulsar la colaboración interdisciplinaria y transnacional en la ciencia abierta.

El uso de técnicas avanzadas de *machine learning*, como Latent Dirichlet Allocation (LDA), K-Means y t-SNE, permitió una comprensión más profunda de los patrones temáticos en la literatura sobre datos abiertos medioambientales. Estas herramientas han complementado los enfoques bibliométricos tradicionales, permitiendo la identificación de clústeres temáticos ocultos y la

organización eficiente de grandes volúmenes de datos textuales. El análisis resultante identificó cinco clústeres temáticos principales: (1) salud pública y mortalidad, especialmente en relación con la pandemia de COVID-19; (2) la intersección entre factores ambientales y sociales; (3) la gestión de recursos naturales, como el agua y el suelo; (4) los riesgos ambientales y la exposición a contaminantes; y (5) la conservación de la biodiversidad. Estos clústeres proporcionan una visión clara de las áreas prioritarias en la investigación científica actual sobre la Amazonía.

El análisis de una selección de 50 artículos altamente influyentes revela una diversidad considerable en los enfoques metodológicos y el uso de fuentes de datos. Predominan los estudios observacionales, ecológicos y experimentales, con un uso extenso de tecnologías avanzadas, como imágenes satelitales (Landsat-8 y Sentinel-2) y modelos predictivos, que se emplean para monitorear fenómenos ambientales y gestionar recursos naturales y salud pública. El uso generalizado de datos abiertos refuerza el compromiso de la comunidad científica con la transparencia y la reproducibilidad, lo que promueve una mayor colaboración interdisciplinaria.

No obstante, a pesar del notable progreso en la adopción de datos abiertos, se han identificado carencias en la claridad y precisión con las que se reportan las fuentes de datos en algunos estudios. Esta falta de especificidad podría dificultar la verificación y replicabilidad de los estudios, lo que resalta la necesidad de mejorar las directrices sobre la descripción y accesibilidad de las fuentes de datos primarias. Además, los estudios que se basan en datos secundarios de grandes bases de datos globales, como el *Global Burden of Disease Study* (GBD) y la base *InvaCost*, destacan la importancia de estos recursos para investigaciones a gran escala, aunque no siempre se recojan directamente por los investigadores.

En síntesis, los hallazgos de este estudio destacan la relevancia internacional y la naturaleza colaborativa de la investigación sobre el acceso abierto a la información ambiental en la Amazonía. La ciencia abierta, impulsada por la colaboración global y el uso de datos abiertos, se perfila como una herramienta fundamental para mejorar nuestra comprensión y gestión del medio ambiente

amazónico. Sin embargo, persisten desafíos, como la necesidad de mejorar la transparencia y la accesibilidad de los datos, lo que será clave para avanzar en la conservación y gestión sostenible de esta región crucial.

6.2. Conclusiones

El presente estudio bibliométrico revela que, aunque el campo de la investigación sobre acceso abierto a la información ambiental y ecológica en la región amazónica ha experimentado un crecimiento sostenido entre 2021 y 2023, este incremento es moderado. Brasil se consolida como el principal motor de producción científica en la región, destacándose tanto en cantidad de publicaciones como en impacto, con una notable influencia global. A pesar de esto, países externos a la región amazónica, como Estados Unidos y Australia, también han tenido una participación significativa, reforzando el carácter internacional de la investigación ambiental sobre la Amazonía y la importancia de las colaboraciones transnacionales.

El estudio muestra que la producción científica está concentrada en un núcleo reducido de revistas de alto impacto y un grupo limitado de autores e instituciones, lo que refleja una alta especialización y colaboración en torno a nodos clave. Universidades brasileñas, lideradas por la Universidad de São Paulo, emergen como centros fundamentales de generación de conocimiento, mientras que la fuerte colaboración internacional —especialmente con países como Ecuador y Colombia dentro de la región, y con Estados Unidos y Australia fuera de ella— ha sido crucial para expandir el impacto global de estas investigaciones.

Las redes de coautoría y co-citación han evidenciado que un grupo selecto de investigadores y organizaciones desempeñan un papel desproporcionado en conectar a diferentes actores, actuando como facilitadores clave de la colaboración y la difusión de conocimiento. Estas redes también confirman que Brasil, Estados Unidos y Australia no solo lideran en la producción científica, sino que también son esenciales para la disseminación del conocimiento en torno al acceso abierto a datos medioambientales.

El análisis temático sugiere que los temas emergentes más relevantes incluyen la relación entre salud pública y medio ambiente, el impacto de la pandemia de COVID-19, y el uso de análisis espaciales avanzados. Estos temas, junto con la biodiversidad y el cambio climático, destacan como áreas prioritarias en la investigación actual, subrayando cómo los datos abiertos están siendo fundamentales para abordar los desafíos contemporáneos en sostenibilidad y salud pública.

Las técnicas estadísticas basadas en redes y machine learning han sido fundamentales para estructurar y analizar grandes volúmenes de datos bibliográficos, lo que permitió identificar cinco áreas temáticas principales: la salud pública, especialmente en relación con la pandemia; la gestión de recursos naturales; la exposición a riesgos ambientales y contaminantes; la intersección entre factores sociales y ambientales; y la conservación de la biodiversidad. Estas áreas temáticas reflejan la diversidad y complejidad de los problemas investigados en el contexto amazónico, y señalan cómo los datos abiertos están facilitando nuevas formas de abordar estos desafíos.

Un aspecto destacado del estudio es la evolución en el uso de tecnologías avanzadas, como imágenes satelitales y modelos predictivos, que han permitido a los investigadores abordar problemas ambientales y de salud pública con mayor precisión. Estas herramientas han sido especialmente útiles en estudios sobre biodiversidad y el impacto del cambio climático, proporcionando datos cuantificables que fortalecen la calidad y reproducibilidad de los hallazgos.

Asimismo, la creciente adopción de la ciencia abierta en este campo es un avance significativo. Un alto porcentaje de los estudios analizados emplea fuentes de datos accesibles al público, lo que no solo fomenta la transparencia, sino que también promueve la colaboración interdisciplinaria. Sin embargo, se han identificado áreas de mejora, especialmente en la especificidad y precisión con que se reportan las fuentes de datos. Esta deficiencia podría comprometer la replicabilidad de los estudios y subraya la necesidad de desarrollar normativas más rigurosas para mejorar la documentación de las fuentes de datos.

Finalmente, el uso de bases de datos secundarias globales, como el *Global Burden of Disease Study* (GBD), ha sido fundamental para permitir análisis a

gran escala que abarcan cuestiones relacionadas con la salud pública y los impactos ambientales globales. Aunque estos datos no son generados directamente por los investigadores, su empleo permite realizar estudios comparativos globales esenciales para abordar problemas ambientales y de sostenibilidad.

En conclusión, el panorama actual de la investigación sobre acceso abierto a la información ambiental en la Amazonía es dinámico y está en evolución. La creciente integración de datos abiertos, el uso de tecnologías avanzadas y el compromiso con la ciencia abierta están configurando un enfoque más colaborativo y global para abordar los complejos problemas ambientales y ecológicos de la región. Sin embargo, para avanzar de manera efectiva, es crucial mejorar las normativas sobre la descripción y accesibilidad de las fuentes de datos y seguir fomentando la colaboración internacional en este campo.

Recomendaciones

- Fortalecer la infraestructura para datos abiertos: Es necesario aumentar la inversión en plataformas que faciliten el intercambio de datos medioambientales de calidad, tanto a nivel regional como global, para potenciar la capacidad de investigación y el impacto de las publicaciones.
- Impulsar la adopción de la ciencia abierta: A pesar del uso creciente de datos abiertos, es necesario fomentar aún más esta práctica, exigiendo que los datos utilizados en las investigaciones sean accesibles y organizados de manera clara.
- Mejorar la documentación de las fuentes de datos: Es crucial que los investigadores proporcionen descripciones detalladas de las fuentes de datos y metodologías utilizadas para asegurar la calidad, replicabilidad y transparencia en los estudios científicos.
- Fortalecer la capacidad analítica: Dado el creciente uso de análisis espaciales y de datos geográficos, es importante desarrollar y adoptar herramientas avanzadas para manejar grandes volúmenes de datos y mejorar la comprensión de los fenómenos ecológicos.

- Fomentar la colaboración internacional: Se debe fortalecer la cooperación entre países amazónicos y no amazónicos, con Brasil liderando iniciativas globales que amplíen la participación en la investigación medioambiental.
- Priorizar la difusión en revistas de alto impacto: Publicar en revistas influyentes como PLOS ONE y The Lancet es crucial para aumentar la visibilidad y el impacto de las investigaciones sobre datos abiertos medioambientales.
- Intensificar la investigación en la intersección entre salud y medio ambiente: Se debe profundizar en los estudios que exploren la relación entre los factores medioambientales y la salud pública, especialmente en contextos de crisis sanitarias, como lo evidenció la pandemia de COVID-19.
- Promover la gestión sostenible de recursos naturales: Continuar investigando la gestión del agua, el suelo y otros recursos naturales es fundamental para desarrollar políticas de sostenibilidad que mejoren el bienestar humano y ambiental.
- Fomentar la conservación de la biodiversidad: Es esencial destinar más recursos a la investigación y conservación de la biodiversidad en áreas críticas de la Amazonía, utilizando datos abiertos para diseñar estrategias de protección efectivas.
- Ampliar el uso de machine learning en estudios medioambientales: La integración de técnicas avanzadas de machine learning como LDA y K-Means permite identificar patrones complejos en grandes conjuntos de datos, mejorando la toma de decisiones en políticas medioambientales.
- Fomentar el uso de tecnologías avanzadas: Tecnologías como imágenes satelitales y modelos predictivos han demostrado su valor en el análisis de problemas ambientales y de salud, por lo que es recomendable seguir invirtiendo en su adopción y desarrollo para aumentar la precisión de los estudios.

REFERENCIAS BIBLIOGRÁFICAS



Referencias Bibliográficas

- Abrahão, J. S., Sacchetto, L., Rezende, I. M., Rodrigues, R. A. L., Crispim, A. P. C., Moura, C., Mendonça, D. C., Reis, E., Souza, F., Oliveira, G. F. G., Domingos, I., De Miranda Boratto, P. V., Silva, P. H. B., Queiroz, V. F., Machado, T. B., Andrade, L. A. F., Lourenço, K. L., Silva, T., Oliveira, G. P., ... Drumond, B. P. (2021). Detection of SARS-CoV-2 RNA on public surfaces in a densely populated urban area of Brazil: A potential tool for monitoring the circulation of infected patients. *Science of The Total Environment*, 766, 142645. <https://doi.org/10.1016/j.scitotenv.2020.142645>
- Ahmed, D. A., Haubrock, P. J., Cuthbert, R. N., Bang, A., Soto, I., Balzani, P., Tarkan, A. S., Macêdo, R. L., Carneiro, L., Bodey, T. W., Oficialdegui, F. J., Courtois, P., Kourantidou, M., Angulo, E., Heringer, G., Renault, D., Turbelin, A. J., Hudgins, E. J., Liu, C., ... Courchamp, F. (2023). Recent advances in availability and synthesis of the economic costs of biological invasions. *BioScience*, 73(8), 560-574. <https://doi.org/10.1093/biosci/biad060>
- Arruda, V. L. S., Piontekowski, V. J., Alencar, A., Pereira, R. S., & Matricardi, E. A. T. (2021). An alternative approach for mapping burn scars using Landsat imagery, Google Earth Engine, and Deep Learning in the Brazilian Savanna. *Remote Sensing Applications: Society and Environment*, 22, 100472. <https://doi.org/10.1016/j.rsase.2021.100472>
- Augusto, M. R., Claro, I. C. M., Siqueira, A. K., Sousa, G. S., Caldereiro, C. R., Duran, A. F. A., De Miranda, T. B., Bomediano Camillo, L. D. M., Cabral, A. D., & De Freitas Bueno, R. (2022). Sampling strategies for wastewater surveillance: Evaluating the variability of SARS-COV-2 RNA concentration in composite and grab samples. *Journal of Environmental Chemical Engineering*, 10(3), 107478. <https://doi.org/10.1016/j.jece.2022.107478>
- Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, 1(1), 377-386. https://doi.org/10.1162/qss_a_00019
- Babini, D., & Rovelli, L. (2020). *Tendencias recientes en las políticas científicas de ciencia abierta y acceso abierto en Iberoamérica*. Consejo Latinoamericano de Ciencias Sociales. CLACSO. <https://doi.org/10.2307/j.ctv1gm02tq>
- Biscaro, C., & Giupponi, C. (2014). Co-Authorship and Bibliographic Coupling Network Effects on Citations. *PLoS ONE*, 9(6), e99502. <https://doi.org/10.1371/journal.pone.0099502>
- Boing, A. F., deSouza, P., Boing, A. C., Kim, R., & Subramanian, S. V. (2022). Air Pollution, Socioeconomic Status, and Age-Specific Mortality Risk in the

- United States. *JAMA Network Open*, 5(5), e2213540. <https://doi.org/10.1001/jamanetworkopen.2022.13540>
- Bonacich, P. (1987). Power and Centrality: A Family of Measures. *American Journal of Sociology*, 92(5), 1170-1182.
- Borgatti, S. P., & Everett, M. G. (2006). A Graph-theoretic perspective on centrality. *Social Networks*, 28(4), 466-484. <https://doi.org/10.1016/j.socnet.2005.11.005>
- Boufarss, M., & Harviainen, J. T. (2021). Librarians as gate-openers in open access publishing: A case study in the United Arab Emirates. *The Journal of Academic Librarianship*, 47(5), 102425. <https://doi.org/10.1016/j.acalib.2021.102425>
- Brack, W., Barcelo Culleres, D., Boxall, A. B. A., Budzinski, H., Castiglioni, S., Covaci, A., Dulio, V., Escher, B. I., Fantke, P., Kandie, F., Fatta-Kassinos, D., Hernández, F. J., Hilscherová, K., Hollender, J., Hollert, H., Jahnke, A., Kasprzyk-Hordern, B., Khan, S. J., Kortenkamp, A., ... Zuccato, E. (2022). One planet: One health. A call to support the initiative on a global science-policy body on chemicals and waste. *Environmental Sciences Europe*, 34(1), 21. <https://doi.org/10.1186/s12302-022-00602-6>
- Burt, R. S. (2004). Structural Holes and Good Ideas. *American Journal of Sociology*, 110(2), 349-399. <https://doi.org/10.1086/421787>
- Camalan, S., Cui, K., Pauca, V. P., Alqahtani, S., Silman, M., Chan, R., Plemmons, R. J., Dethier, E. N., Fernandez, L. E., & Lutz, D. A. (2022). Change Detection of Amazonian Alluvial Gold Mining Using Deep Learning and Sentinel-2 Imagery. *Remote Sensing*, 14(7), 1746. <https://doi.org/10.3390/rs14071746>
- Canabarro, N. I., Silva-Ortiz, P., Nogueira, L. A. H., Cantarella, H., Maciel-Filho, R., & Souza, G. M. (2023). Sustainability assessment of ethanol and biodiesel production in Argentina, Brazil, Colombia, and Guatemala. *Renewable and Sustainable Energy Reviews*, 171, 113019. <https://doi.org/10.1016/j.rser.2022.113019>
- Chen, G., Guo, Y., Yue, X., Tong, S., Gasparrini, A., Bell, M. L., Armstrong, B., Schwartz, J., Jaakkola, J. J. K., Zanobetti, A., Lavigne, E., Nascimento Saldiva, P. H., Kan, H., Royé, D., Milojevic, A., Overcenco, A., Urban, A., Schneider, A., Entezari, A., ... Li, S. (2021). Mortality risk attributable to wildfire-related PM2.5 pollution: A global time series study in 749 locations. *The Lancet Planetary Health*, 5(9), e579-e587. [https://doi.org/10.1016/S2542-5196\(21\)00200-X](https://doi.org/10.1016/S2542-5196(21)00200-X)
- Claro, I. C. M., Cabral, A. D., Augusto, M. R., Duran, A. F. A., Graciosa, M. C. P., Fonseca, F. L. A., Speranca, M. A., & Bueno, R. D. F. (2021). Long-term monitoring of SARS-COV-2 RNA in wastewater in Brazil: A more responsive and economical approach. *Water Research*, 203, 117534. <https://doi.org/10.1016/j.watres.2021.117534>
- Colón-González, F. J., Soares Bastos, L., Hofmann, B., Hopkin, A., Harpham, Q., Crocker, T., Amato, R., Ferrario, I., Moschini, F., James, S., Malde, S.,

- Ainscoe, E., Sinh Nam, V., Quang Tan, D., Duc Khoa, N., Harrison, M., Tsarouchi, G., Lumbroso, D., Brady, O. J., & Lowe, R. (2021). Probabilistic seasonal dengue forecasting in Vietnam: A modelling study using superensembles. *PLOS Medicine*, 18(3), e1003542. <https://doi.org/10.1371/journal.pmed.1003542>
- Crippa, J. A. S., Zuardi, A. W., Guimarães, F. S., Campos, A. C., De Lima Osório, F., Loureiro, S. R., Dos Santos, R. G., Souza, J. D. S., Ushirohira, J. M., Pacheco, J. C., Ferreira, R. R., Mancini Costa, K. C., Scomparin, D. S., Scarante, F. F., Pires-Dos-Santos, I., Mechoulam, R., Kapczinski, F., Fonseca, B. A. L., Esposito, D. L. A., ... Coutinho, B. M. (2021). Efficacy and Safety of Cannabidiol Plus Standard Care vs Standard Care Alone for the Treatment of Emotional Exhaustion and Burnout Among Frontline Health Care Workers During the COVID-19 Pandemic: A Randomized Clinical Trial. *JAMA Network Open*, 4(8), e2120603. <https://doi.org/10.1001/jamanetworkopen.2021.20603>
- Del-Aguila-Arcentales, S., Alvarez-Risco, A., Jaramillo-Arévalo, M., De-la-Cruz-Díaz, M., & Anderson-Seminario, M. D. L. M. (2022). Influence of Social, Environmental and Economic Sustainable Development Goals (SDGs) over Continuation of Entrepreneurship and Competitiveness. *Journal of Open Innovation: Technology, Market, and Complexity*, 8(2), 73. <https://doi.org/10.3390/joitmc8020073>
- Figueiredo Cerqueira, M. M. D., Castro, M. M. L., Vieira, A. A., Kurosawa, J. A. A., Amaral Junior, F. L. D., Siqueira Mendes, F. D. C. C. D., & Sosthenes, M. C. K. (2023). Comparative analysis between Open Field and Elevated Plus Maze tests as a method for evaluating anxiety-like behavior in mice. *Heliyon*, 9(4), e14522. <https://doi.org/10.1016/j.heliyon.2023.e14522>
- Filho, W. L., Salvia, A. L., Bonoli, A., Saari, U. A., Voronova, V., Klõga, M., Kumbhar, S. S., Olszewski, K., De Quevedo, D. M., & Barbir, J. (2021). An assessment of attitudes towards plastics and bioplastics in Europe. *Science of The Total Environment*, 755, 142732. <https://doi.org/10.1016/j.scitotenv.2020.142732>
- Fonseca, B. D. P. F. E., Sampaio, R. B., Fonseca, M. V. D. A., & Zicker, F. (2016). Co-authorship network analysis in health research: Method and potential use. *Health Research Policy and Systems*, 14(1), 34. <https://doi.org/10.1186/s12961-016-0104-5>
- Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, 40(1), 35. <https://doi.org/10.2307/3033543>
- Friedrich, F., Ongaratto, R., Scotta, M. C., Veras, T. N., Stein, R. T., Lumertz, M. S., Jones, M. H., Comaru, T., & Pinto, L. A. (2021). Early Impact of Social Distancing in Response to Coronavirus Disease 2019 on Hospitalizations for Acute Bronchiolitis in Infants in Brazil. *Clinical Infectious Diseases*, 72(12), 2071-2075. <https://doi.org/10.1093/cid/ciaa1458>
- Garrigós-Simón, F., Sanz-Blas, S., Narangajavana, Y., & Buzova, D. (2021). The Nexus between Big Data and Sustainability: An Analysis of Current Trends and Developments. *Sustainability*, 13(12), 6632. <https://doi.org/10.3390/su13126632>

- Glänzel, W., & Schubert, A. (2005). Analysing Scientific Networks Through Co-Authorship. En H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research* (pp. 257-276). Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-2755-9_12
- Gómez-Ochoa, S. A., Rojas, L. Z., Echeverría, L. E., Muka, T., & Franco, O. H. (2022). Global, Regional, and National Trends of Chagas Disease from 1990 to 2019: Comprehensive Analysis of the Global Burden of Disease Study. *Global Heart*, 17(1), 59. <https://doi.org/10.5334/gh.1150>
- Hacket-Pain, A., Foest, J. J., Pearse, I. S., LaMontagne, J. M., Koenig, W. D., Vacchiano, G., Bogdziewicz, M., Caignard, T., Celebias, P., Van Dormolen, J., Fernández-Martínez, M., Moris, J. V., Palaghianu, C., Pesendorfer, M., Satake, A., Schermer, E., Tanentzap, A. J., Thomas, P. A., Vecchio, D., ... Ascoli, D. (2022). MASTREE+: Time-series of plant reproductive effort from six continents. *Global Change Biology*, 28(9), 3066-3082. <https://doi.org/10.1111/gcb.16130>
- Haunschild, R., Bornmann, L., & Marx, W. (2016). Climate Change Research in View of Bibliometrics. *PLOS ONE*, 11(7), e0160393. <https://doi.org/10.1371/journal.pone.0160393>
- Heinrich, C. J., Camacho, S., Henderson, S. C., Hernández, M., & Joshi, E. (2022). Consequences of Administrative Burden for Social Safety Nets that Support the Healthy Development of Children. *Journal of Policy Analysis and Management*, 41(1), 11-44. <https://doi.org/10.1002/pam.22324>
- Kass, J. M., Pinilla-Buitrago, G. E., Paz, A., Johnson, B. A., Grisales-Betancur, V., Meenan, S. I., Attali, D., Broennimann, O., Galante, P. J., Maitner, B. S., Owens, H. L., Varela, S., Aiello-Lammens, M. E., Merow, C., Blair, M. E., & Anderson, R. P. (2023). wallace 2: A shiny app for modeling species niches and distributions redesigned to facilitate expansion via module contributions. *Ecography*, 2023(3), e06547. <https://doi.org/10.1111/ecog.06547>
- Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26(1), 1-18. [https://doi.org/10.1016/S0048-7333\(96\)00917-1](https://doi.org/10.1016/S0048-7333(96)00917-1)
- Kawamura, M., Thomas, C. D. L., Tsurumoto, A., Sasahara, H., & Kawaguchi, Y. (2000). Lotka's Law and Productivity Index of Authors in a Scientific Journal. *Journal of Oral Science*, 42(2), 75-78. <https://doi.org/10.2334/josnusd.42.75>
- Lense, G. H. E., Lämmle, L., Ayer, J. E. B., Lama, G. F. C., Rubira, F. G., & Mincato, R. L. (2023). Modeling of Soil Loss by Water Erosion and Its Impacts on the Cantareira System, Brazil. *Water*, 15(8), 1490. <https://doi.org/10.3390/w15081490>
- Li, J., Goerlandt, F., & Reniers, G. (2021). An overview of scientometric mapping for the safety science community: Methods, tools, and framework. *Safety Science*, 134, 105093. <https://doi.org/10.1016/j.ssci.2020.105093>

- Local Burden of Disease HIV Collaborators. (2021). Mapping subnational HIV mortality in six Latin American countries with incomplete vital registration systems. *BMC Medicine*, 19(1), 4. <https://doi.org/10.1186/s12916-020-01876-4>
- Magalhães, A. R., Codeço, C. T., Svenning, J.-C., Escobar, L. E., Van De Vuurst, P., & Gonçalves-Souza, T. (2023). Neglected tropical diseases risk correlates with poverty and early ecosystem destruction. *Infectious Diseases of Poverty*, 12(1), 32. <https://doi.org/10.1186/s40249-023-01084-1>
- Mansouri, K., Karmaus, A. L., Fitzpatrick, J., Patlewicz, G., Pradeep, P., Alberga, D., Alepee, N., Allen, T. E. H., Allen, D., Alves, V. M., Andrade, C. H., Auernhammer, T. R., Ballabio, D., Bell, S., Benfenati, E., Bhattacharya, S., Bastos, J. V., Boyd, S., Brown, J. B., ... Kleinstreuer, N. C. (2021). CATMoS: Collaborative Acute Toxicity Modeling Suite. *Environmental Health Perspectives*, 129(4), 047013. <https://doi.org/10.1289/EHP8495>
- Marques, P., Silva, A. S., Quaresma, Y., Manna, L. R., De Magalhães Neto, N., & Mazzoni, R. (2021). Home gardens can be more important than other urban green infrastructure for mental well-being during COVID-19 pandemics. *Urban Forestry & Urban Greening*, 64, 127268. <https://doi.org/10.1016/j.ufug.2021.127268>
- Martins, J. A. C., Nogueira, K., Osco, L. P., Gomes, F. D. G., Furuya, D. E. G., Gonçalves, W. N., Sant'Ana, D. A., Ramos, A. P. M., Liesenberg, V., Dos Santos, J. A., De Oliveira, P. T. S., & Junior, J. M. (2021). Semantic Segmentation of Tree-Canopy in Urban Environment with Pixel-Wise Deep Learning. *Remote Sensing*, 13(16), 3054. <https://doi.org/10.3390/rs13163054>
- Mestanza-Ramón, C., Jiménez-Oyola, S., Montoya, A. V. G., Vizuete, D. D. C., D'Orio, G., Cedeño-Laje, J., & Straface, S. (2023). Assessment of Hg pollution in stream waters and human health risk in areas impacted by mining activities in the Ecuadorian Amazon. *Environmental Geochemistry and Health*, 45(10), 7183-7197. <https://doi.org/10.1007/s10653-023-01597-6>
- Nascimento, F. O. D., Cheng, J., & Feijó, A. (2021). Taxonomic revision of the pampas cat *Leopardus colocola* complex (Carnivora: Felidae): an integrative approach. *Zoological Journal of the Linnean Society*, 191(2), 575-611. <https://doi.org/10.1093/zoolinnean/zlaa043>
- Nathan, A., Beynor, J. I., Baglaenko, Y., Suliman, S., Ishigaki, K., Asgari, S., Huang, C.-C., Luo, Y., Zhang, Z., Lopez, K., Lindestam Arlehamn, C. S., Ernst, J. D., Jimenez, J., Calderón, R. I., Lecca, L., Van Rhijn, I., Moody, D. B., Murray, M. B., & Raychaudhuri, S. (2021). Multimodally profiling memory T cells from a tuberculosis cohort identifies cell state associations with demographics, environment and disease. *Nature Immunology*, 22(6), 781-793. <https://doi.org/10.1038/s41590-021-00933-1>
- Omar, M., & Kamaruzaman, F. M. (2024). Technical and vocational education training and industry collaboration: A bibliometric review. *Journal of*

- Education and Learning (EduLearn)*, 18(4), 1582-1592.
<https://doi.org/10.11591/edulearn.v18i4.21120>
- Ong, K. L., Stafford, L. K., McLaughlin, S. A., Boyko, E. J., Vollset, S. E., Smith, A. E., Dalton, B. E., Duprey, J., Cruz, J. A., Hagins, H., Lindstedt, P. A., Aali, A., Abate, Y. H., Abate, M. D., Abbasian, M., Abbasi-Kangevari, Z., Abbasi-Kangevari, M., Abd ElHafeez, S., Abd-Rabu, R., ... Vos, T. (2023). Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: A systematic analysis for the Global Burden of Disease Study 2021. *The Lancet*, 402(10397), 203-234.
[https://doi.org/10.1016/S0140-6736\(23\)01301-6](https://doi.org/10.1016/S0140-6736(23)01301-6)
- Osareh, F., Salehi Zahabi, S., & Akbarzadeh, F. (2022). Co-authorship Network Analysis of Medical Images Researchers with Emphasis on Micro and Macro Metrics. *Journal of Clinical Research in Paramedical Sciences*, 11(2). <https://doi.org/10.5812/jcrps-131621>
- OTCA. (2021). *Programa de acciones estratégicas, estrategia regional para la gestión integrada de los recursos hídricos de la cuenca amazónica*. Otca.
- OTCA. (2024). *¿Quiénes Somos?* - OTCA [Página informativa]. *¿Quiénes Somos?* <https://otca.org/quienes-somos/>
- Oviedo-Trespalacios, O., Peden, A. E., Cole-Hunter, T., Costantini, A., Haghani, M., Rod, J. E., Kelly, S., Torkamaan, H., Tariq, A., David Albert Newton, J., Gallagher, T., Steinert, S., Filtness, A. J., & Reniers, G. (2023). The risks of using ChatGPT to obtain common safety-related information and advice. *Safety Science*, 167, 106244.
<https://doi.org/10.1016/j.ssci.2023.106244>
- Pachito, D. V., Pega, F., Bakusic, J., Boonen, E., Clays, E., Descatha, A., Delvaux, E., De Bacquer, D., Koskenvuo, K., Kröger, H., Lambrechts, M.-C., Latorraca, C. O. C., Li, J., Cabrera Martimbianco, A. L., Riera, R., Rugulies, R., Sembajwe, G., Siegrist, J., Sillanmäki, L., ... Godderis, L. (2021). The effect of exposure to long working hours on alcohol consumption, risky drinking and alcohol use disorder: A systematic review and meta-analysis from the WHO/ILO Joint Estimates of the Work-related Burden of Disease and Injury. *Environment International*, 146, 106205.
<https://doi.org/10.1016/j.envint.2020.106205>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Alonso-Fernández, S. (2021). Declaración PRISMA 2020: Una guía actualizada para la publicación de revisiones sistemáticas. *Revista Española de Cardiología*, 74(9), 790-799.
<https://doi.org/10.1016/j.recesp.2021.06.016>
- Parra, M. A., Baez, S., Sedeño, L., Gonzalez Campo, C., Santamaría-García, H., Aprahamian, I., Bertolucci, P. H., Bustin, J., Camargos Bicalho, M. A., Cano-Gutierrez, C., Caramelli, P., Chaves, M. L. F., Cogram, P., Beber, B. C., Court, F. A., De Souza, L. C., Custodio, N., Damian, A., De La Cruz, M., ... Ibanez, A. (2021). Dementia in Latin America: Paving the way

- toward a regional action plan. *Alzheimer's & Dementia*, 17(2), 295-313. <https://doi.org/10.1002/alz.12202>
- Pereira, G. H. de A., Fusioka, A. M., Nassu, B. T., & Minetto, R. (2021). Active Fire Detection in Landsat-8 Imagery: A Large-Scale Dataset and a Deep-Learning Study. *ISPRS Journal of Photogrammetry and Remote Sensing*, 178, 171-186. <https://doi.org/10.1016/j.isprsjprs.2021.06.002>
- Potsdam Institute for Climate Impact Research. (2023). Amazon Rain Forest Is Losing Resilience. *Current Opinion in Environmental Sustainability*, 50. <https://doi.org/10.1016/j.cosust.2023.101014>
- Redkina, N. S. (2021). The Library in the Information Ecosystem of Open Science. *Scientific and Technical Information Processing*, 48(4), 239-247. <https://doi.org/10.3103/S0147688221040043>
- Rodriguez-Villamizar, L. A., Belalcázar-Ceron, L. C., Fernández-Niño, J. A., Marín-Pineda, D. M., Rojas-Sánchez, O. A., Acuña-Merchán, L. A., Ramírez-García, N., Mangones-Matos, S. C., Vargas-González, J. M., Herrera-Torres, J., Agudelo-Castañeda, D. M., Piñeros Jiménez, J. G., Rojas-Roa, N. Y., & Herrera-Galindo, V. M. (2021). Air pollution, sociodemographic and health conditions effects on COVID-19 mortality in Colombia: An ecological study. *Science of The Total Environment*, 756, 144020. <https://doi.org/10.1016/j.scitotenv.2020.144020>
- Rossini, P. (2024). Networks as interpretative frameworks: Using co-citation analysis to explore large corpora of early modern letters. *Digital Scholarship in the Humanities*, 39(1), 321-334. <https://doi.org/10.1093/llc/fqad086>
- Ruel, T. D., Acosta, E. P., Liu, J. P., Gray, K. P., George, K., Montañez, N., Popson, S., Buchanan, A. M., Bartlett, M., Dayton, D., Anthony, P., Brothers, C., Vavro, C., Singh, R., Koech, L., Vhembo, T., Hazra, R., Townley, E., & Wiznia, A. (2021). Pharmacokinetics, Safety, Tolerability and Antiviral Activity of Dolutegravir Dispersible Tablets in Infants and Children with HIV-1: Results of the IMPAACT P1093 Study, a Phase I/II Open-Label Trial. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3954096>
- Saetnan, E. R., & Kipling, R. P. (2016). Evaluating a European knowledge hub on climate change in agriculture: Are we building a better connected community? *Scientometrics*, 109(2), 1057-1074. <https://doi.org/10.1007/s11192-016-2064-5>
- San-Juan-Heras, R., Gabriel, J. L., Delgado, M. M., Alvarez, S., & Martinez, S. (2024). Scientometric analysis of cover crop management: Trends, networks, and future directions. *European Journal of Agronomy*, 161, 127355. <https://doi.org/10.1016/j.eja.2024.127355>
- Schneider, M. C., Min, K., Hamrick, P. N., Montebello, L. R., Ranieri, T. M., Mardini, L., Camara, V. M., Raggio Luiz, R., Liese, B., Vuckovic, M., Moraes, M. O., & Lima, N. T. (2021). Overview of snakebite in Brazil: Possible drivers and a tool for risk mapping. *PLOS Neglected Tropical Diseases*, 15(1), e0009044. <https://doi.org/10.1371/journal.pntd.0009044>

- Shah, F. A., Meyer, N. J., Angus, D. C., Awdish, R., Azoulay, É., Calfee, C. S., Clermont, G., Gordon, A. C., Kwizera, A., Leligdowicz, A., Marshall, J. C., Mikacenic, C., Sinha, P., Venkatesh, B., Wong, H. R., Zampieri, F. G., & Yende, S. (2021). A Research Agenda for Precision Medicine in Sepsis and Acute Respiratory Distress Syndrome: An Official American Thoracic Society Research Statement. *American Journal of Respiratory and Critical Care Medicine*, 204(8), 891-901. <https://doi.org/10.1164/rccm.202108-1908ST>
- Shao, L., Cao, Y., Jones, T., Santosh, M., Silva, L. F. O., Ge, S., Da Boit, K., Feng, X., Zhang, M., & Bérubé, K. (2022). COVID-19 mortality and exposure to airborne PM2.5: A lag time correlation. *Science of The Total Environment*, 806, 151286. <https://doi.org/10.1016/j.scitotenv.2021.151286>
- Shaw, E., Walpole, S., McLean, M., Alvarez-Nieto, C., Barna, S., Bazin, K., Behrens, G., Chase, H., Duane, B., Gibbs, T. J., Groome, J., Hackett, F., Harden, J., Hothersall, J., Hourihane, M., Huss, N. M., Ikiugu, M., Leedham-Green, K., MacKenzie-Shalders, K., ... Woollard, R. (2021). *AMEE Consensus Statement: Planetary health and education for sustainable healthcare*.
- Silva, J. L. B. D., Moura, G. B. D. A., Silva, M. V. D., Oliveira-Júnior, J. F. D., Jardim, A. M. D. R. F., Refati, D. C., Lima, R. D. C. C., Carvalho, A. A. D., Ferreira, M. B., Brito, J. I. B. D., Guedes, R. V. D. S., Lopes, P. M. O., Nóbrega, R. S., Pandorfi, H., Bezerra, A. C., Batista, P. H. D., Jesus, F. L. F. D., Sanches, A. C., & Santos, R. C. (2023). Environmental degradation of vegetation cover and water bodies in the semiarid region of the Brazilian Northeast via cloud geoprocessing techniques applied to orbital data. *Journal of South American Earth Sciences*, 121, 104164. <https://doi.org/10.1016/j.jsames.2022.104164>
- Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4), 265-269. <https://doi.org/10.1002/asi.4630240406>
- Stapleton, F., Abad, J. C., Barabino, S., Burnett, A., Iyer, G., Lekhanont, K., Li, T., Liu, Y., Navas, A., Obinwanne, C. J., Qureshi, R., Roshandel, D., Sahin, A., Shih, K., Tichenor, A., & Jones, L. (2023). TFOS Lifestyle: Impact of societal challenges on the ocular surface. *The Ocular Surface*, 28, 165-199. <https://doi.org/10.1016/j.jtos.2023.04.006>
- Stoler, J., Miller, J. D., Brewis, A., Freeman, M. C., Harris, L. M., Jepson, W., Pearson, A. L., Rosinger, A. Y., Shah, S. H., Staddon, C., Workman, C., Wutich, A., Young, S. L., Adams, E., Ahmed, F., Alexander, M., Asiki, G., Balogun, M., Boivin, M. J., ... Tutu, R. (2021). Household water insecurity will complicate the ongoing COVID-19 response: Evidence from 29 sites in 23 low- and middle-income countries. *International Journal of Hygiene and Environmental Health*, 234, 113715. <https://doi.org/10.1016/j.ijheh.2021.113715>

- Urban, R. C., & Nakada, L. Y. K. (2021). COVID-19 pandemic: Solid waste and environmental impacts in Brazil. *Science of The Total Environment*, 755, 142471. <https://doi.org/10.1016/j.scitotenv.2020.142471>
- Valente, T. W. (2011). Social Networks and Health: Models, Methods, and Applications. Thomas W. Valente. *International Journal of Epidemiology*, 40(6), 1742-1743. <https://doi.org/10.1093/ije/dyr160>
- Van Eck, N. J., & Waltman, L. (2014). Visualizing Bibliometric Networks. En Y. Ding, R. Rousseau, & D. Wolfram (Eds.), *Measuring Scholarly Impact* (pp. 285-320). Springer International Publishing. https://doi.org/10.1007/978-3-319-10377-8_13
- van Eck, N. J., & Waltman, L. (2020). *VOSviewer Manual*. 53.
- Vargas-Munoz, J. E., Srivastava, S., Tuia, D., & Falcao, A. X. (2021). OpenStreetMap: Challenges and Opportunities in Machine Learning and Remote Sensing. *IEEE Geoscience and Remote Sensing Magazine*, 9(1), 184-199. <https://doi.org/10.1109/MGRS.2020.2994107>
- Viljur, M., Abella, S. R., Adámek, M., Alencar, J. B. R., Barber, N. A., Beudert, B., Burkle, L. A., Cagnolo, L., Campos, B. R., Chao, A., Chergui, B., Choi, C., Cleary, D. F. R., Davis, T. S., Dechnik-Vázquez, Y. A., Downing, W. M., Fuentes-Ramirez, A., Gandhi, K. J. K., Gehring, C., ... Thorn, S. (2022). The effect of natural disturbances on forest biodiversity: An ecological synthesis. *Biological Reviews*, 97(5), 1930-1947. <https://doi.org/10.1111/brv.12876>
- Wallace, B., Posnik, Z., Hurley, B., DiMatteo, A., Bandimere, A., Rodriguez, I., Maxwell, S., Meyer, L., Brenner, H., Jensen, M., LaCasella, E., Shamblin, B., Abreu Abreu-Grobois, F., Stewart, K., Dutton, P., Barrios-Garrido, H., Dalleau, M., Dell'amico, F., Eckert, K., ... Mast, R. (2023). Marine turtle regional management units 2.0: An updated framework for conservation and research of wide-ranging megafauna species. *Endangered Species Research*, 52, 209-223. <https://doi.org/10.3354/esr01243>
- Yuan, T., Da Rocha Neto, W., Rothenberg, C. E., Obraczka, K., Barakat, C., & Turletti, T. (2022). Machine learning for next-generation intelligent transportation systems: A survey. *Transactions on Emerging Telecommunications Technologies*, 33(4), e4427. <https://doi.org/10.1002/ett.4427>
- Zaneti, L. A. L., Arias, N. B., De Almeida, M. C., & Rider, M. J. (2022). Sustainable charging schedule of electric buses in a University Campus: A rolling horizon approach. *Renewable and Sustainable Energy Reviews*, 161, 112276. <https://doi.org/10.1016/j.rser.2022.112276>
- Zhang, M. (2010). Social Network Analysis: History, Concepts, and Research. En B. Furht (Ed.), *Handbook of Social Network Technologies and Applications* (pp. 3-21). Springer US. https://doi.org/10.1007/978-1-4419-7142-5_1
- Zhao, Q., Guo, Y., Ye, T., Gasparrini, A., Tong, S., Overcenco, A., Urban, A., Schneider, A., Entezari, A., Vicedo-Cabrera, A. M., Zanobetti, A., Analitis, A., Zeka, A., Tobias, A., Nunes, B., Alahmad, B., Armstrong, B., Forsberg,

Referencias:

B., Pan, S.-C., ... Li, S. (2021). Global, regional, and national burden of mortality associated with non-optimal ambient temperatures from 2000 to 2019: A three-stage modelling study. *The Lancet Planetary Health*, 5(7), e415-e425. [https://doi.org/10.1016/S2542-5196\(21\)00081-4](https://doi.org/10.1016/S2542-5196(21)00081-4)

RESUMEN

El presente libro investiga la evolución del acceso abierto a la información científica relacionada con el medio ambiente en los países miembros de la Organización del Tratado de Cooperación Amazónica entre 2021 y 2023. Como objetivo tiene identificar patrones de producción científica en acceso abierto, así como las barreras y oportunidades que presenta este modelo para fomentar la investigación colaborativa y apoyar la conservación de la región amazónica. Mediante un análisis bibliométrico basado en el protocolo PRISMA, se analizaron más de mil publicaciones científicas disponibles en acceso abierto. Los resultados indican que Brasil es el principal productor de investigación en la región, seguido de Colombia y Ecuador, aunque aún persisten importantes desafíos como la brecha digital y la falta de infraestructura en otras zonas amazónicas. La discusión resalta cómo el acceso abierto ha facilitado la cooperación internacional, aunque se advierte que las disparidades en la disponibilidad de recursos limitan el impacto del conocimiento generado. En conclusión, el estudio subraya que el acceso abierto es esencial para promover la investigación y la toma de decisiones informadas en la Amazonía, pero se requiere un esfuerzo conjunto para superar las barreras estructurales que afectan la difusión equitativa del conocimiento científico en la región.

Palabras Clave: biodiversidad amazónica, ciencia abierta, medio ambiente, métricas de colaboración científica, tendencias científicas.

Abstract

This book investigates the evolution of open access to scientific information related to the environment in the member countries of the Amazon Cooperation Treaty Organization between 2021 and 2023. It aims to identify patterns of scientific production in open access, as well as the barriers and opportunities presented by this model to foster collaborative research and support conservation in the Amazon region. Through a bibliometric analysis based on the PRISMA protocol, more than one thousand scientific publications available in open access were analyzed. The results indicate that Brazil is the main producer of research in the region, followed by Colombia and Ecuador, although there are still important challenges such as the digital divide and the lack of infrastructure in other Amazonian areas. The discussion highlights how open access has facilitated international cooperation, although it warns that disparities in the availability of resources limit the impact of the knowledge generated. In conclusion, the study underlines that open access is essential to promote research and informed decision making in the Amazon, but a concerted effort is required to overcome the structural barriers that affect the equitable dissemination of scientific knowledge in the region.

Keywords: amazon biodiversity, open science, environment, scientific collaboration metrics, scientific trends.



<http://www.editorialgrupo-aea.com>



[Editorial Grupo AeA](#)



[editorialgrupoea](#)



[Editorial Grupo AEA](#)

ISBN: 978-9942-651-63-1



9 789942 651631

