

Datos Alternativos a la Justicia Algorítmica en Riesgo de Crédito: Evidencia Bibliométrica Global con Enfoque en Tecnologías Disruptivas

Gladys Elizabeth, Proaño-Altamirano *
<https://orcid.org/0000-0001-6809-7687>

Diana Andrea Rosero-Jiménez **
<https://orcid.org/0009-0002-3668-1816>

Isaac Agustín Paredes-Flor **
<https://orcid.org/0000-0003-1465-4016>

Evelyn Alexandra Gavilanes-Carranza ***
<https://orcid.org/0000-0002-1854-8125>

Instituto Superior Tecnológico España - ISTE, Ambato, Ecuador *
Pontificia Universidad Católica del Ecuador - Sede Ambato, Ambato, Ecuador **
Universidad Técnica de Ambato, Ambato, Ecuador ***

Email: gladys.proano@iste.edu.ec

Recibido: 07 de julio de 2025 / **Aprobado:** 26 de septiembre de 2025 / **Publicado:** 27 de octubre de 2025

DOI: <https://doi.org/10.24133/bm57md85>

Resumen

Este artículo tiene como propósito caracterizar, desde una perspectiva bibliométrica global, el estado y la evolución del campo “datos alternativos y justicia algorítmica en riesgo de crédito” en el contexto de tecnologías disruptivas; los objetivos son mapear productividad e influencia, describir la colaboración científica, delinear la estructura temática y la dinámica de términos emergentes, y examinar el tono discursivo. Se analizan 1.619 documentos recuperados de Scopus mediante indicadores de productividad y citación, redes de coautoría por autores y países, co-ocurrencia de palabras clave, mapa temático y curvas TF-IDF, complementados con análisis de polaridad y subjetividad; el procesamiento se realiza con pharos natural. El hallazgo más consistente indica una producción concentrada en pocos autores e instituciones con liderazgo asiático especialmente India y una red de coautoría fragmentada con puentes limitados; la estructura cognitiva evidencia cuatro clústeres diferenciados que convergen en tópicos motores vinculados con datos alternativos, explicabilidad y mitigación de sesgo; los

términos fintech, digital y data muestran trayectoria ascendente sostenida; el tono del campo es predominantemente positivo (polaridad media $\approx 0,77$; subjetividad $\approx 0,36$). Se concluye que el potencial de los datos alternativos para mejorar la evaluación del riesgo de crédito se materializa cuando se densifica la colaboración internacional, se institucionalizan auditorías reproducibles con métricas de equidad y se validan modelos en poblaciones y marcos regulatorios diversos, equilibrando desempeño predictivo y justicia algorítmica.

Palabras Clave: datos alternativos; justicia algorítmica; riesgo de crédito; bibliometría; FinTech.

Alternative Data to Algorithmic Justice in Credit Risk: Global Bibliographic Evidence with a Focus on Disruptive Technologies

Abstract

The purpose of this article is to characterize, from a global bibliometric perspective, the state and evolution of the field of “alternative data and algorithmic justice in credit risk” in the context of disruptive technologies. The objectives are to map productivity and influence, describe scientific collaboration, outline the thematic structure and dynamics of emerging terms, and examine the discursive tone. A total of 1,619 documents retrieved from Scopus are analyzed using productivity and citation indicators, co-authorship networks by authors and countries, keyword co-occurrence, thematic map, and TF-IDF curves, complemented by polarity and subjectivity analysis; processing is performed with Pharos Natural. The most consistent finding indicates that production is concentrated among a few authors and institutions with Asian leadership, especial y India, and a fragmented co-authorship network with limited bridges. The cognitive structure shows four distinct clusters that converge on key topics related to alternative data, explainability, and bias mitigation. The terms fintech, digital, and data show a sustained upward trend. The tone of the field is predominantly positive (average polarity ≈ 0.77 ; subjectivity ≈ 0.36). It is concluded that the potential of alternative data to improve credit risk assessment is realized when international collaboration is intensified, reproducible audits with equity metrics are institutionalized, and models are validated in diverse populations and regulatory frameworks, balancing predictive performance and algorithmic fairness.

Keywords: alternative data; algorithmic fairness; credit risk; bibliometrics; FinTech.

Introducción

A nivel global, el uso de datos alternativos en evaluación de riesgo de crédito ha pasado de pruebas de concepto a implementaciones que combinan huellas digitales, registros transaccionales y señales de telefonía para ampliar cobertura y granularidad predictiva, aunque con tensiones sobre validez externa y sesgos diferenciales que pueden afectar poblaciones con menor inclusión digital. En el plano macro, los sistemas financieros enfrentan el dilema entre eficiencia y justicia en decisiones automatizadas; en el plano meso, las entidades deben gobernar variables, umbrales y seguimiento; en el plano micro, los solicitantes exigen explicaciones útiles y no discriminación. Este estado del arte muestra que las ganancias de exactitud no garantizan trato justo sin métricas, auditorías y trazabilidad integradas desde el diseño del modelo y su operación en producción, lo que justifica un mapa bibliométrico comparativo del campo emergente de datos alternativos y justicia algorítmica en crédito (Berg, Burg, Gombović, & Puri, 2020; Björkegren & Grissen, 2019; Barocas & Selbst, 2016; Hurley & Adebayo, 2016).

Este resumen la equidad algorítmica y tecnologías disruptivas en mercados emergentes, cuyo propósito es caracterizar la evolución intelectual y empírica del dominio, identificar clústeres temáticos y estimar vacíos de evidencia aplicable. El objetivo es construir una base de datos reproducible que permita analizar productividad, colaboración,

coocurrencias y acoplamientos, y contrastar cómo cambian los enfoques de datos alternativos y justicia según contextos regulatorios y tecnológicos. La pregunta guía es cómo y dónde se operacionaliza la justicia algorítmica cuando se incorporan datos alternativos en scoring crediticio, y qué trayectorias metodológicas dominan en la última década. La estrategia bibliométrica combinará normalización rigurosa, mapas de ciencia y reporte transparente siguiendo buenas prácticas y lineamientos de calidad para revisiones sistemáticas (Aria & Cuccurullo, 2017; van Eck & Waltman, 2010; Donthu, Kumar, Mukherjee, Pandey, & Lim, 2021; Page et al., 2021).

Metodología

Se realizó un estudio bibliométrico global de la literatura sobre datos alternativos y justicia algorítmica en riesgo de crédito, con un enfoque mixto predominantemente cuantitativo de carácter exploratorio, bibliográfico y documental. La recopilación de información se llevó a cabo mediante una búsqueda exhaustiva en la base de datos scopus.

En concreto, se empleó una ecuación de búsqueda booleana que combinó términos relacionados con riesgo de crédito, con datos alternativos y con justicia algorítmica/ética en IA, además de incluir términos sobre tecnologías disruptivas en finanzas como: artificial intelligence, machine learning, blockchain, digital platforms. Esta estrategia de búsqueda permitió recuperar un total de 11.890 documentos iniciales en Scopus.

Adicionalmente, la consulta se limitó a documentos de acceso abierto, del tipo artículo o revisión, pertenecientes a las áreas temáticas de Computer Science, Economics/Econometrics/Finance, Social Sciences y Business/Management/Accounting, para garantizar la relevancia disciplinaria y la disponibilidad de los textos completos filtrando 1619 documentos para el análisis.

De todos los registros obtenidos, se descargaron los metadatos completos (información bibliográfica, resúmenes, palabras clave, referencias, etc.) para su análisis. Posteriormente, se eliminaron los duplicados para depurar la base de datos final. El proceso de identificación y selección de la literatura se documentó siguiendo las directrices, registrando de manera transparente el flujo de inclusión/exclusión de publicaciones (identificación, cribado, elegibilidad y selección final). Tras la depuración, la muestra final de documentos fue empleada para los análisis bibliométricos y estadísticos descritos a continuación.

El diseño del estudio adoptó un enfoque metodológico mixto con predominio del análisis bibliométrico cuantitativo. Se optó por utilizar exclusivamente la base de datos Scopus como fuente de información, decisión que se justificó con base en comparativas de cobertura y calidad reportadas en la literatura. Estudios previos han demostrado que Scopus ofrece una cobertura amplia y adecuada de revistas en las disciplinas analizadas, siendo comparable o superior a otras bases de datos internacionales (Mongeon & Paul-Hus, 2016;

Archambault et al., 2009; Martín-Martín et al., 2018). Asimismo, se tomaron en cuenta lineamientos metodológicos reconocidos que respaldan esta elección y fomentan la exhaustividad en la búsqueda de información (Page et al., 2021). De este modo, la recopilación de datos a través de Scopus aseguró la integridad y calidad del corpus bibliográfico analizado.

Para el análisis de los datos, se aplicaron diversas técnicas bibliométricas y herramientas de cienciometría, complementadas con análisis estadísticos inferenciales. En particular, se llevaron a cabo las siguientes actividades y métricas bibliométricas principales:

En cuanto a las herramientas empleadas, el procesamiento y análisis de los datos se llevaron a cabo con soporte de software especializado, se utilizó el lenguaje R- Studio y pharos natural para la realización de los cálculos bibliométricos y la extracción de indicadores cuantitativos (Aria & Cuccurullo, 2017), así como también, la construcción y visualización de mapas de redes (coautoría, cocitación, palabras clave, etc.). En particular, la detección de comunidades y clusters en las redes se basó en el algoritmo de agrupación modular optimizado propuesto por Waltman & van Eck (2013), lo que garantizó una identificación consistente de estructuras temáticas.

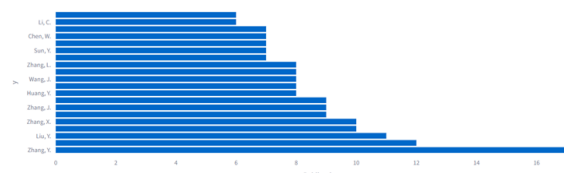
En síntesis, la metodología aplicada integró técnicas bibliométricas establecidas y análisis estadísticos complementarios para ofrecer una

visión integral de la producción científica en la intersección de datos alternativos, riesgo de crédito y justicia algorítmica. Se siguieron enfoques validados en la literatura especializada (Aria & Cuccurullo, 2017; van Eck & Waltman, 2010; Waltman & van Eck, 2013; Cobo et al., 2011), con el fin de garantizar rigor, reproducibilidad y validez en los resultados obtenidos.

Resultados

Figura 1

Top autores más productivos



Nota. Elaboración propia con datos de Scopus procesados en pharos natural (2025).

La figura muestra una concentración marcada de productividad con Zhang, Y como líder destacado, seguido a distancia por autores del segundo estrato como Zhang, K, Zhang, X, Huang, K, Wang, A, Zhang, A, Tan, S, Chen, W y Li, C.C.; este patrón, coherente con la ley de Lotka, sugiere que unos pocos nodos sostienen la cadencia de publicación y podrían estar configurando estándares metodológicos en datos alternativos y justicia algorítmica para riesgo de crédito.

Figura 2

Afiliaciones de países más frecuentes

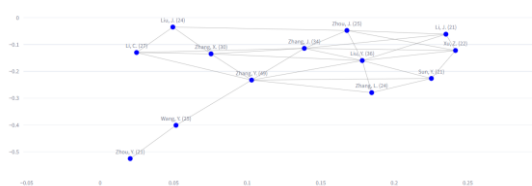


Nota. Elaboración propia con datos de Scopus procesados en pharos natural (2025)

La figura evidencia una concentración institucional marcada en Asia, con predominio de India Amity University Noida, Chandigarh University, Christ University y el Department of Management Studies de Graphic Era Deemed to be University y con presencia secundaria de China y Hong Kong, además de nodos puntuales en Singapur, Reino Unido, Jordania y Brasil. En términos comparativos, el peso relativo de India sugiere ecosistemas universitarios con capacidad instalada en ciencia de datos y crédito digital, lo que acelera la producción sobre datos alternativos y justicia algorítmica.

Figura 3

Red de Coautoría



Nota. Elaboración propia con datos de Scopus procesados en pharos natural (2025)

La red de coautoría exhibe una estructura dispersa con pocos puentes efectivos entre clústeres, donde autores como Zhang, Y, Li, J, Liu, Y, Zhou, L, Sun, W, Wang, Y y Luo, A concentran los vínculos más visibles pero sin configurar un núcleo

denso; este patrón sugiere baja centralidad intermedia y una intermediación limitada, lo que favorece la formación de silos temáticos y reduce la difusión transversal de métodos sobre datos alternativos y justicia algorítmica en riesgo de crédito.

Figura 4

Mapa 2D de documentos según tópicos



Nota. Elaboración propia con datos de Scopus procesados en pharos natural (2025)

El mapa 2D sugiere cuatro conglomerados temáticos diferenciados sobre dimensiones latentes de co-ocurrencia, con dos grupos dominantes que ocupan la mayor densidad de puntos y exhiben mayor dispersión horizontal, indicio de variedad metodológica y aplicaciones amplias en el uso de datos alternativos para riesgo de crédito.

En contraste, los clústeres restantes concentran menos documentos y aparecen parcialmente solapados, lo que apunta a líneas más especializadas o emergentes por ejemplo, enfoques de equidad algorítmica y explicabilidad cuya proximidad a los núcleos dominantes anticipa trayectorias de convergencia. En términos comparativos, la separación entre centroides sugiere diferenciación temática suficiente para sostener agendas propias, mientras que el solapamiento moderado permite transferencia de técnicas y métricas, condición favorable

para madurar estándares de auditoría y mitigación de sesgo en el dominio analizado.

Figura 5

Mapa de calor – promedio de polaridad por años y tipo



Nota. Elaboración propia con datos de Scopus procesados en pharos natural (2025)

El mapa de calor muestra una predominancia sostenida de polaridad positiva a lo largo del periodo, con intensidades altas y estables, mientras la negativa disminuye y la neutral permanece cerca de cero.

Comparativamente, el gradiente sugiere un desplazamiento de la literatura hacia valoraciones favorables conforme maduran marcos metodológicos y métricas de equidad. Persisten, sin embargo, focos negativos que apuntan a controversias sobre sesgo, gobernanza y transferibilidad, lo que demanda contraste con análisis cualitativos para corroborar la interpretación.

Figura 6

Boxplot de polaridad por tipo de documento

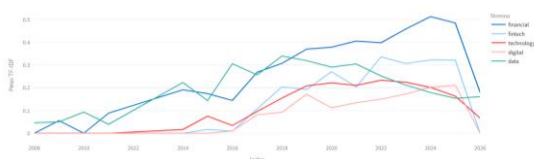


Nota. Elaboración propia con datos de Scopus procesados en pharos natural (2025)

El boxplot evidencia un sesgo general hacia polaridades positivas, con medianas altas y consistentes en conference papers y artículos. Libros y capítulos muestran dispersión moderada, mientras revisiones y capítulos exhiben colas negativas más pronunciadas, señal de heterogeneidad crítica en sesgo, gobernanza y transferibilidad. Comparativamente, la menor variabilidad de artículos y ponencias sugiere metodologías más consolidadas, frente a debates aún abiertos en las síntesis y desarrollos teóricos.

Figura 7

Curvas de crecimiento TF-IDF de términos emergentes



Nota. Elaboración propia con datos de Scopus procesados en pharos natural (2025)

Las curvas TF-IDF muestran una trayectoria ascendente y sostenida de la saliencia para fintech, digital y data desde 2015, con picos entre 2023 y 2025, mientras technology despegue más tarde y alcanza su máximo ligeramente después, lo que sugiere un desplazamiento del foco desde infraestructuras tecnológicas generales hacia aplicaciones financieras y datos alternativos específicos del crédito. Comparativamente, fintech domina la cota superior y digital mantiene niveles altos y estables, en tanto financial y technology exhiben aceleraciones más tardías. La contracción al final del periodo es compatible con efectos de ventana

truncada y no necesariamente con pérdida de relevancia, por lo que la lectura sustantiva es de consolidación temática alrededor de datos alternativos y digitalización del riesgo, con creciente conexión a debates de justicia algorítmica.

Conclusión

La evidencia presentada responde a la pregunta de investigación al mostrar que el campo de los datos alternativos y la justicia algorítmica en riesgo de crédito se encuentra en consolidación, con una estructura productiva concentrada en pocos autores líderes y con fuerte gravitación institucional en Asia, especialmente India. La red de coautoría se presenta fragmentada y con puentes limitados, por lo que la circulación transversal de métodos y métricas depende de alianzas puntuales más que de un núcleo cohesionado. Los mapas temáticos y el 2D confirman la existencia de cuatro clústeres diferenciados con zonas de solapamiento que facilitan transferencia metodológica, mientras las curvas TF-IDF muestran una trayectoria ascendente de fintech, digital y data que desplaza el foco desde infraestructuras generales hacia aplicaciones específicas del crédito.

La caracterización del tono científico indica predominio de valoraciones positivas, con polaridad alta y estable en artículos y ponencias, pero con heterogeneidad crítica en revisiones y capítulos, donde emergen debates sobre sesgo, gobernanza y explicabilidad. Esta combinación sugiere que la comunidad

reconoce beneficios técnicos y de eficiencia de los datos alternativos, pero mantiene reservas sobre su justicia y transferibilidad entre jurisdicciones y segmentos de población. En este marco, la influencia real no se explica solo por productividad, sino por la posición estructural de autores e instituciones como nodos puente capaces de articular estándares y buenas prácticas.

Con base en estos hallazgos, la conclusión final establece que el desarrollo responsable del campo exige, en el presente, tres desplazamientos concretos: densificar la colaboración internacional para superar la fragmentación observada, institucionalizar marcos de auditoría reproducible que unifiquen métricas de equidad y criterios de explicabilidad, y promover evaluaciones externas que prueben la validez de los modelos en contextos regulatorios y poblacionales diversos. Bajo estas condiciones, el potencial de los datos alternativos para mejorar la evaluación del riesgo de crédito se materializa sin sacrificar principios de justicia algorítmica, y el área avanza desde una consolidación liderada por polos regionales hacia una agenda global con estándares compartidos.

Referencias bibliográficas

- Aria, M., & Cuccurullo, C. (2017). *bibliometrix: An R-tool for comprehensive science mapping analysis*. *Journal of Informetrics*, 11(4), 959–975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Archambault, É., Campbell, D., Gingras, Y., & Larivière, V. (2009). Comparing bibliometric statistics obtained from the Web of Science and Scopus. *Journal of the American Society for Information Science and Technology*, 60(7), 1320–1326. <https://doi.org/10.1002/asi.21062>
- Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, 104(3), 671–732. <https://doi.org/10.15779/Z38ZK5D>
- Berg, T., Burg, V., Gombović, A., & Puri, M. (2020). On the rise of fintechs: Credit scoring using digital footprints. *The Review of Financial Studies*, 33(7), 2845–2897. <https://doi.org/10.1093/rfs/hhz099>
- Björkegren, D., & Grissen, D. (2019). Behavior revealed in mobile phone usage predicts credit repayment. *The World Bank Economic Review*, 33(2), 367–393. <https://doi.org/10.1093/wber/lhz018>
- Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology*, 62(7), 1382–1402. <https://doi.org/10.1002/asi.21525>

- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285–296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
- Hurley, M., & Adebayo, J. (2016). Credit scoring in the era of big data. *Yale Journal of Law & Technology*, 18, 148–216. <https://yjolt.org>
- Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & Delgado López-Cózar, E. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison. *Scientometrics*, 116, 2175–2188. <https://doi.org/10.1007/s11192-018-2711-2>
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106, 213–228. <https://doi.org/10.1007/s11192-015-1765-5>
- 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., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- Waltman, L., & van Eck, N. J. (2013). A smart local moving algorithm for large-scale modularity-based community detection. *The European Physical Journal B*, 86, 471. <https://doi.org/10.1140/epjb/e2013-40829-0>