European Journal of Sustainable Development Research
2026, 10(4), em0436
e-ISSN: 2542-4742
https://www.ejosdr.com/

MODESTUM

Review Article

OPEN ACCESS

Artificial intelligence in tax compliance: A bibliometric mapping and research agenda (2015–2025)

Houda Zaim [1*], Siham Sahbani [1]

[1] Sidi Mohamed Ben Abdellah University, National School of Commerce and Management, MOROCCO
***Corresponding Author:** houda.zaim1@usmba.ac.ma

Citation: Zaim, H., & Sahbani, S. (2026). Artificial intelligence in tax compliance: A bibliometric mapping and research agenda (2015–2025). *European Journal of Sustainable Development Research, 10*(4), em0436. https://doi.org/10.29333/ejosdr/19260

ARTICLE INFO

Received: 14 Mar. 2026

Accepted: 23 Jul. 2026

ABSTRACT

While AI is integral to the tax compliance ecosystem, academic research is scattered across tax enforcement, governance, and sustainability. This bibliometric analysis of 1,803 Scopus-indexed articles from 2015 to 2025 shows that the field is not yet mature, but is undergoing rapid thematic change, with three waves, enforcement, governance, and sustainability, compressed into a decade. The central efficacy-legitimacy challenge is clear: AI improves anomaly detection and audit targeting, but algorithmic opacity can undermine taxpayer trust. The geographic inversion is also telling: Ukraine, Jordan, and Malaysia have more studies than OECD operational centers, suggesting reform urgency rather than administrative maturity. The study introduces the ATCGM, combining effectiveness, legitimacy, and sustainability, and identifies five empirically grounded research gaps. The practical implications are real: Peru's e-invoicing rollout increased declared VAT liabilities by over 5% in its first year, but the governance conditions behind such gains remain under-researched.

Keywords: artificial intelligence, tax compliance, bibliometric analysis, digital taxation, algorithmic governance, tax modernization

INTRODUCTION

Tax compliance is one of the fundamental requirements of public finance and plays a crucial role in the functioning of the State, in social redistribution and in economic stability (Alm et al., 2012). The field of studies around tax compliance in the last 50 years has moved from a classical economic standpoint (as reflected in the expected-utility tax compliance model developed by Allingham and Sandmo 1972) towards a more multi-disciplinary approach that includes elements of psychology, sociology and institutional theory (Feld & Frey, 2002; Kirchler, 2007; Murphy, 2005). Today, it is recognized that elements of trust, procedural justice and perceived fairness contribute more to voluntary compliance than one might attribute by looking at the odds of an audit and the size of the penalties.

Three developments converged and spurred the digitization of tax administration to take off at a rapid pace since 2015: the development of machine learning to analyze large-scale administrative data (Bengio et al., 2013); the OECD's (2020) Tax Administration 3.0 initiative, which requires real-time digital reporting; and the widespread introduction of electronic invoicing around the world (Bird & Zolt, 2008). Artificial Intelligence (AI) is directly relevant to tax compliance as it puts the fundamental functions of modern tax administration into practice. Machine learning algorithms identify tax filing irregularities, such as "red flags", as indicators of tax evasion; Natural Language Processing (NLP) provides structured data from unstructured tax forms and invoices; and predictive analytics target tax files for audit based on risk score, not on a random selection basis (Aggarwal, 2024; Aslett et al., 2024; McKinsey & Company, 2021). In its operations, HMRC's Connect system compares more than one billion data points each year, while Brazil's Receita Federal also uses machine learning to combat VAT fraud in millions of electronic invoices and China's Golden Tax System III combines corporate tax risk profiling with AI-powered invoice verification at the national level (OECD, 2023). AI acts not only as compliance add-on, but a core component of compliance infrastructure that enables the monitoring, enforcement, and measurement of compliance.

Meanwhile, many issues such as effectiveness, legitimacy and administrative responsibility, remain unresolved with AI. Concerns have emerged with regards to algorithmic transparency, data privacy and embedded bias (Adadi & Berrada, 2018; Veale & Binns, 2017). AI systems face particular challenges in tax administration, due to the "black box" nature of some of these systems, where decisions need to be legally explainable and contestable (Kroll et al., 2017). Additionally, the pace of AI adoption is faster than what institutions can keep up with, leading to information asymmetries between tax authorities and taxpayers (Braithwaite, 2003).

There are previous bibliometric reviews of tax compliance in general (Thaha et al., 2023), AI in finance and accounting in general, and digital taxation (Belahouaoui & Attak, 2024). None has specifically synthesized the field of AI and tax compliance behavior as a unit that combines predictive enforcement, taxpayer legitimacy and modernization towards sustainability. This represents a major blind spot: without a clear picture of what has been researched, by whom, and where, it will be hard for policy makers and researchers to develop clear evidence bases for decisions on AI deployment or the governance architectures that would enable these deployments to be deemed legitimate and equitable.

The period 2015-2025 is analytically relevant. As far as technology is concerned, 2015 was the year of industrialization of deep learning, with the release of production-grade frameworks (Abadi et al., 2015) and the shift of neural network applications to operational public administration. On the institutional level, technology-driven tax transformation was acknowledged as a policy priority in the OECD's BEPS Action 1 report (2015) and further developed in OECD Tax Administration 3.0 (OECD, 2020). In practice, since 2015, the amount of AI-taxation research indexed in Scopus is great enough to allow for meaningful bibliometric network mapping. In regulatory terms, the post-2015 period covers the GDPR framework (2016/2018) on algorithmic data processing, as well as the introduction of national AI strategies, which explicitly took fiscal governance applications into consideration. Pre-2015 output was negligible, and the 2015-2025 window covers the full scope of AI-powered tax compliance.

The general research question of this work is: ***What is the conceptualization and application of Artificial Intelligence in the global academic literature on tax compliance in the last decade (2015-2025)?*** This is realized by the following sub-questions:

RQ1: What are the most influential sources, documents and authors in AI and tax compliance research?

RQ2: What are the regional patterns and who are the main contributors in terms of institutions and countries?

RQ3: How is the literature organized into themes, and how have these themes changed over the past 10 years?

RQ4: What implications for theory building and future research can be drawn from the bibliometric evidence?

The paper makes an original contribution in three ways. It first maps and interprets a three-wave thematic organization of the field, as indicated by the bibliometric evidence. Secondly, it offers the AI Tax Compliance Governance Matrix (ATCGM), which serves as an integrative analytical matrix with effectiveness, legitimacy, and sustainability as dimensions that are not competing but interconnected. Third, it identifies five empirically based research gaps that have direct practical relevance for tax authorities, policymakers, and researchers.

METHODOLOGY

The design of this study is based on a bibliometric approach and science mapping within a positivist paradigm (Aria & Cuccurullo, 2017; Donthu et al., 2021; Zupic & Čater, 2015). Bibliometric analysis is suitable for domains that are emerging and fragmented, as it enables the systematic examination of the growth of publications, co-authorship and the keywords relation (Tranfield et al., 2003). Two software environments are deployed, one for descriptive indicators and publication trend analysis (Biblioshiny, v.4.1.3) and another one for network visualization of co-authorship and keyword co-occurrence (VOSviewer, v.1.6.20) (Van Eck & Waltman, 2010). The interpretive synthesis links the bibliometric patterns with theory on tax administration, taxpayer behavior, and tax policy modernization, following quantitative mapping.

Data Collection and Filtering

Because the governance and behavioral dimensions of AI-driven taxation are largely invisible to computer-science-only framing, this study deliberately filters for social-science scholarship rather than algorithmic optimization research. A comprehensive Boolean search was executed on Scopus in July 2025, combining AI-technology terms (machine learning, deep learning, neural networks, natural language processing, predictive analytics, anomaly detection, generative AI, chatbots, robotic process automation) with tax-compliance terms (tax evasion, tax audit, VAT fraud, compliance risk management, taxpayer behavior, digital taxation, e-filing). The Boolean search query was:

ALL FIELDS ((“artificial intelligence” or “AI” or “machine learning” or “deep learning” or “neural network*” or “expert system*” or “intelligent system*” or “knowledge-based system*” or “data mining” or “predictive analytics” or “data analytics” or “anomaly detection” or “risk scoring” or “natural language processing” or “large language model*” or “generative AI” or “chatbot*” or “robotic process automation” or “computational intelligence”) **AND** (“tax compliance” or “fiscal compliance” or “taxpayer compliance” or “tax evasion” or “tax avoidance” or “tax fraud” or “tax audit” or “risk-based audit” or “tax administration” or “VAT fraud” or “GST fraud” or “sales tax fraud” or “tax gap” or “tax crime” or “compliance risk management” or “taxpayer behavior” or “digital taxation” or “e-filing”))

The initial query returned 7,059 documents. A five-phase PRISMA-compliant filtering process (Page et al., 2021) produced the analytical corpus:

Phase 1, Temporal boundary: Restriction to 2015–2025 (7,059 documents). The lower bound aligns with the maturation of deep learning frameworks and OECD BEPS Action 1 institutional recognition of digital tax transformation.

Phase 2, Subject area filtering: Retention of Economics, Econometrics and Finance (ECON), Business, Management and Accounting (BUSI) and Social Sciences (SSCI) subject areas (3,679 documents). This social science lens is critical because the governance and behavioral dimensions of AI-driven taxation, taxpayer trust, procedural justice, and institutional legitimacy, are largely invisible to computer science framing.

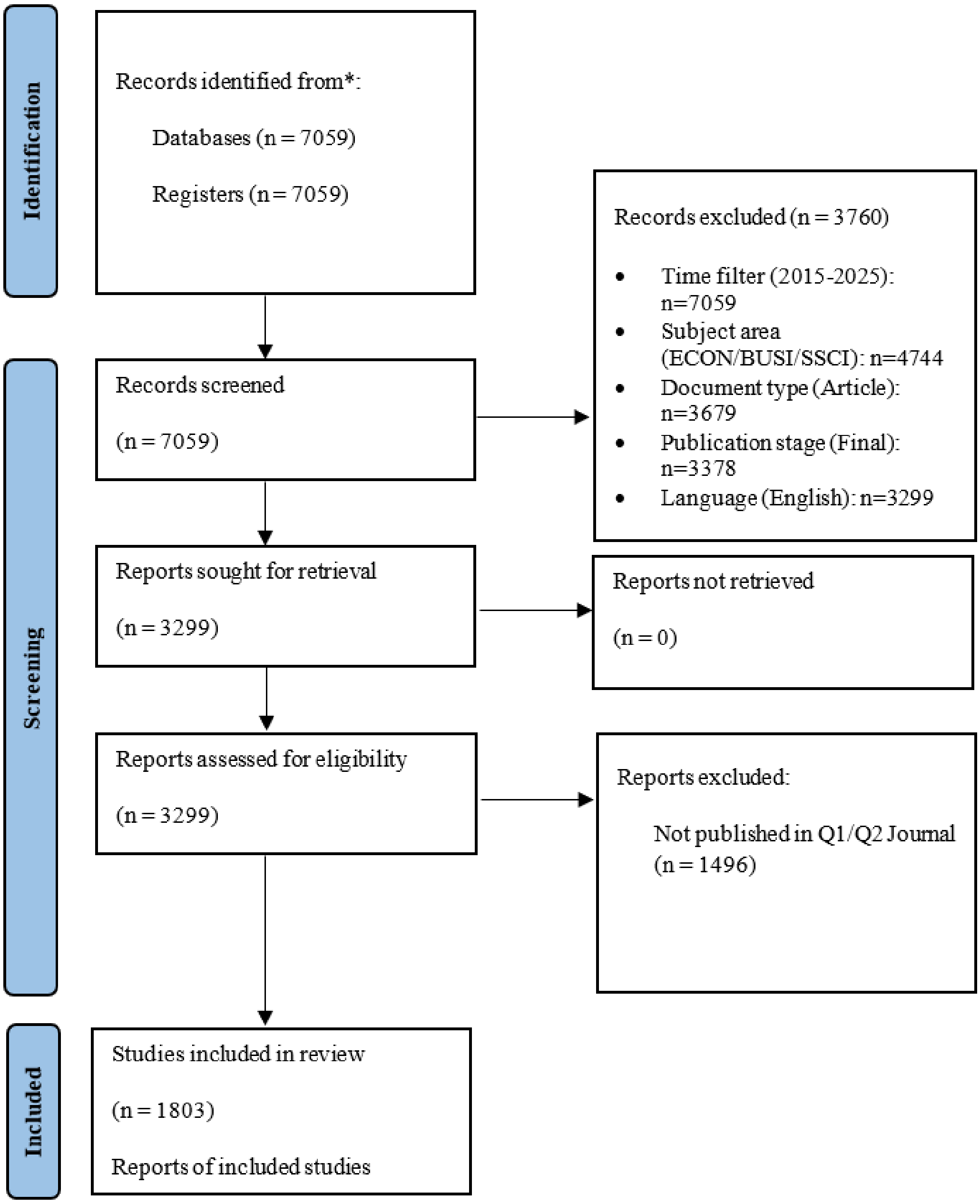

Figure 1. PRISMA 2020 flow diagram of the study selection process (Source: Authors' own elaboration)

To avoid treating algorithmic taxation as only a computer science problem, studies that did not use AI methods relevant to tax administration or compliance were omitted, and the map depicts the social science of algorithmic taxation, not the computer science of algorithmic optimization.

Phase 3, Document type: Peer-reviewed articles only (3,378 documents). To establish methodological uniformity, conference papers, book chapters and editorials were not included.

Phase 4, Language: English-language final versions (3,299 papers).

Phase 5, Quality filter: Scimago Q1–Q2 journals (1,803 articles), allowing the analytical focus to be on established scholarly contributions (Podsakoff et al., 2008).

Figure 1 presents the PRISMA 2020 flow diagram for the selection process. The funnel structure illustrates that the most significant reduction occurred at the subject-area filter (Phase 2), where over half of initial records were excluded to retain only social-science-oriented scholarship. The quality filter (Phase 5) then reduced the corpus from 3,299 to 1,803 articles, confirming that the majority of relevant English-language research on this topic is published in established Q1–Q2 outlets, a finding that itself reflects the growing institutional legitimacy of AI-tax compliance as a scholarly field.

Analytical Approach

For RQ1 and RQ2, descriptive bibliometric analysis (Biblioshiny) was performed, which looked at publication trends, country and institutional productivity, journal concentration and citation indicators. RQ3 was answered using the science mapping tool, VOSviewer, which created keyword co-occurrence networks (threshold: ⩾12 occurrences, 91 keywords) and co-authorship networks (threshold: ⩾3 publications per author). Thematic evolution was illustrated

Table 1. Main informations

Description	Results
Timespan	2015-2025
Sources (journals)	664
Documents	1803
Annual growth rate %	36.04
Document average age	2.35
Average citations per doc	16.56
References	
Keywords plus (id)	4905
Authors keyword	13588
Authors	8114
Co-authors per doc	9.13
International co-authorships %	36.77

using temporal overlay visualization during the past ten years. Research gap synthesis involved combining network structural analysis with substantive interpretation of the content of the cluster, to answer RQ4.

Limitations

Scopus-only coverage may omit relevant publications found in Web of Science or other databases. Emerging journals and scholarship in languages other than English may not be included because of the English-language and Q1–Q2 restrictions. The Q1–Q2 quality filter was applied to ensure the analytical focus rests on established scholarly contributions with rigorous peer review (Podsakoff et al., 2008); however, this criterion may exclude high-quality empirical work published in journals below Q2 ranking, particularly regional or practitioner-oriented outlets that carry operationally relevant evidence on AI deployment in developing-country tax authorities. Citation-lag bias means 2024–2025 publications are structurally under cited. AI research evolves rapidly; peer-reviewed cycles may lag recent developments such as large language model applications in tax advisory. These constraints are reframed as future research directions in Section *Conclusion*.

RESULTS

Main Information

The article corpus consists of 1,803 articles across 664 journals and is growing by 36.04% per year, with 36.77% international co-authorship (**Table 1**). These overall statistics conceal a more meaningful pattern: growth here is less an indicator of field maturity than of rapid thematic reorientation. Enforcement, governance and sustainability are three research waves discernible within a single 10 year window. The 2.35-year average document age suggests that foundational questions tend to be revisited rather than resolved as AI modalities evolve, suggesting that the field has not yet converged on a canonical theoretical framework.

Figure 2 shows a near-flat baseline from 2015–2018 (21–39 articles), followed by an inflection in 2019 and exponential growth through 2024 (456 articles). The 2024 peak captures the overlap of three thematic waves, enforcement, governance, and sustainability, compressed into a single decade rather than unfolding sequentially. The 2025 decline (398) reflects both citation-lag bias and the Q1–Q2 indexing delay noted in Section *Limitations*, confirming that the bibliometric present consistently trails operational deployment by roughly one to two years. Overall, the curve signals not field maturity, but accelerating thematic turbulence.

Most Relevant Sources

The top three journals are Finance Research Letters (66 articles), International Review of Economics and Finance (37) and Journal of Money Laundering Control (32) (**Table 2**). This “disciplinary centering consequence” is visible in the fact that AI-tax scholarship is mostly offered in finance publications rather than in public economics or public administration, where tax compliance historically anchored. This continued hierarchy reflects bibliometric inertia: an effectiveness-first framing persists despite the thematic shift toward governance and sustainability in 2024. There is a bifurcated periphery that can be identified: ethics and environmental journals mark the legitimacy–sustainability turn, while risk-analytics venues keep the financial orientation. Remarkably, there are no generalist public administration journals, confirming that AI and tax compliance remains a niche field within finance and applied ethics. This disciplinary positioning is a reproduction of fragmentation; effectiveness is overrepresented, legitimacy is scattered, and sustainability is isolated in non-communicating ecosystems.

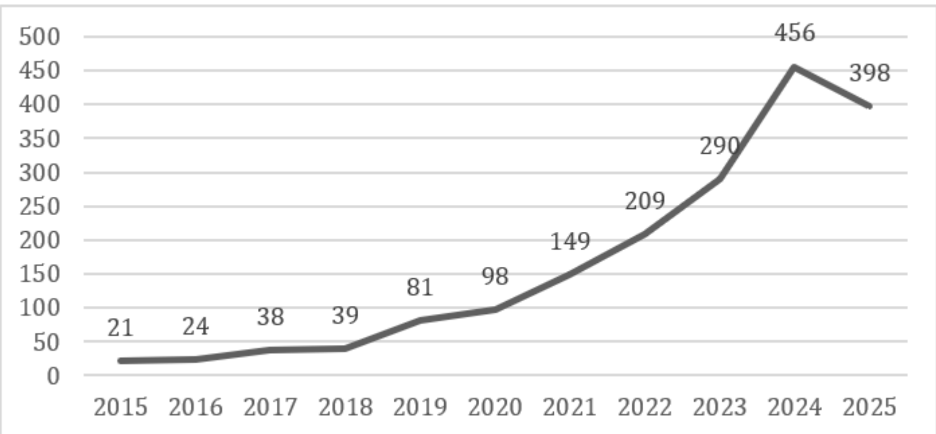

Figure 2. Publication trend from Scopus (2015–2025) (Source: Authors’ own elaboration)

Table 2. Most relevant sources from Scopus

Sources	Volume	Publisher	Index	SJR
Finance Research Letters	66	Elsevier Ltd	Scopus	Q1
International Review of Economics and Finance	37	Elsevier Inc.	Scopus	Q1
Journal of Money Laundering Control	32	Emerald Publishing	Scopus	Q1
Journal of Risk and Financial Management	31	MDPI	Scopus	Q2
Research in International Business and Finance	27	Elsevier Ltd	Scopus	Q1
Journal of Business Ethics	18	Springer Science and Business Media B.V.	Scopus	Q1
Energy Economics	17	Elsevier B.V.	Scopus	Q1
Accounting Review	17	American Accounting Association	Scopus	Q1
Problems and Perspectives in Management	17	LLC CPC Business Perspectives	Scopus	Q2
Journal of Cleaner Production	16	Elsevier Ltd	Scopus	Q1
Contemporary Accounting Research	16	John Wiley and Sons Inc	Scopus	Q1
International Journal of Accounting Information Systems	16	Elsevier Inc.	Scopus	Q1
Intertax	16	Kluwer Law International	Scopus	Q2
British Accounting Review	15	Academic Press	Scopus	Q1
Environment, Development and Sustainability	15	Springer Science and Business Media B.V.	Scopus	Q1
Business Strategy and the Environment	15	John Wiley and Sons Ltd	Scopus	Q1
Managerial and Decision Economics	15	John Wiley and Sons Ltd	Scopus	Q2
International Journal of Data and Network Science	14	Growing Science	Scopus	Q2
Corporate Social Responsibility and Environmental Management	13	John Wiley and Sons Ltd	Scopus	Q1
Transforming Government: People, Process and Policy	12	Emerald Publishing	Scopus	Q2

Table 3. Most relevant affiliations in AI and tax compliance research (2015–2025)

Affiliations	Country	# of articles
Sumy State University, Sumy, Ukraine	Ukraine	33
Universiti Teknologi MARA, Shah Alam, Malaysia	Malaysia	30
King Faisal University, Al-Ahsa, Saudi Arabia	Saudi Arabia	21
The Hong Kong Polytechnic University, Hong Kong, Hong Kong	Hong Kong	20
Universitatea Babeş-Bolyai, Cluj Napoca, Romania	Romania	19
Applied Science Private University, Amman, Jordan	Jordan	16
Central University of Finance and Economics, Beijing, China	China	16
Shanghai University of Finance and Economics, Shanghai, China	China	16
Southwestern University of Finance and Economics, Chengdu, China	China	16
The University of Jordan, Amman, Jordan	Jordan	16
Wuhan University, Wuhan, China	China	15
Azerbaijan State University of Economics (UNEC), Baku, Azerbaijan	Azerbaijan	15
Middle East University Jordan, Amman, Jordan	Jordan	14
Xi'an Jiaotong University, Xi'an, China	China	13
Universiti Malaya, Kuala Lumpur, Malaysia	Malaysia	12
Universiti Sains Malaysia, Gelugor, Malaysia	Malaysia	12
Universiti Kebangsaan Malaysia, Bangi, Malaysia	Malaysia	12
Zhongnan University of Economics and Law, Wuhan, China	China	12
Monash University, Melbourne, Australia	Australia	11
Xiamen University, Xiamen, China	China	11

Most Relevant Affiliations and Regional Context

The productivity map (**Table 3**) shows a remarkable geographic inversion: Sumy State University (Ukraine, 33 articles) and Universiti Teknologi MARA (Malaysia, 30) are the top two, and the top-20 affiliations are completely lacking OECD operational centers such as IRS, HMRC, and BZSt. This pattern confirms the “leapfrogging hypothesis” that there is a relationship between research intensity and institutional reform urgency, rather than administrative maturity.

There are three regional modalities that can be identified. Ukraine is damage-responsive: post-Maidan governance reforms targeted limited institutional resources toward algorithmic audit as an institutional reconstruction process, which resulted in a research intensity that is related to fragility rather than strength. This is also the case with Romania (19 articles) and Azerbaijan (15). Jordan (46 articles from three institutions) is an example of opportunity seizing: anticipatory positioning as an Islamic fintech hub, linking with its knowledge partners, Saudi Arabia and Malaysia, in a horizontal South–South diffusion rather than the vertical OECD-to-developing-world technology transfer assumed by conventional models of technology adoption. Malaysia plays a pivotal role in supporting digital governance testbeds for ASEAN through the national e-government agenda. Although China is operationally leading through the Golden Tax System III, its overall presence in the middle tier (around 16 articles per institution) demonstrates state-scholarship decoupling: advanced deployments are administratively driven and not necessarily reflected in English-language Scopus journals.

Co-authorship Network

The co-authorship network (threshold: ⩾3 publications) shows that there are three epistemic communities that are

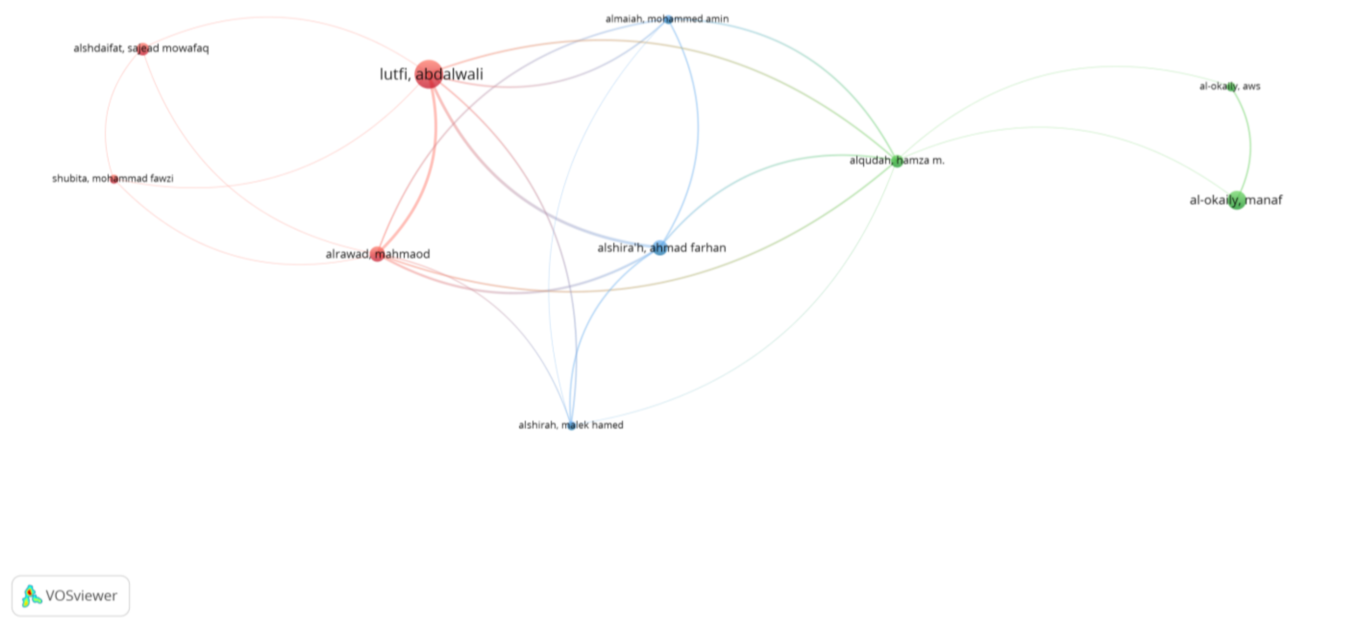

Figure 3. Co-authorship network (Source: Authors' own elaboration, using VOSviewer)

relatively closed and with little interaction between them (**Figure 3**). The Red Cluster (Lutfi, Alrawad, Alshdaifat, Shubita) is the foundation for micro-level corporate compliance, digitalization and fintech. The Blue Cluster (Almaiah, Alshira'h, Alshirah) is dedicated to public-sector technology adoption, AI integration and e-government services. The Green Cluster (Alqudah, Al-Okaily) also encompasses fintech infrastructure, cloud-based accounting systems and AI-driven governance in the Middle East.

Three structural features are important to note. Firstly, the absence of cross-cluster co-authorship indicates that the dimensions of effectiveness, legitimacy and sustainability are not theorized as interdependent, although they are all operationally simultaneous in deployed AI systems. Second, Anglo-American research centers are absent entirely, meaning that OECD jurisdictions generate tax AI systems, while emerging-economy scholars generate the academic literature on those systems. Third, there is no evidence of brokerage nodes, which means there is no integrative hub connecting the three clusters. The Jordan–Saudi Arabia–Malaysia corridor is an epistemic network characterized by horizontal transfer of knowledge across emerging economies, in contrast to the vertical OECD diffusion models.

Keyword Co-occurrence Network

Keyword co-occurrence analysis (threshold: ⩾12 occurrences, 91 keywords) yields five epistemic communities whose spatial arrangement encodes the field's power geometry, which concepts function as gravitational centres, which bridge clusters, and which remain structurally isolated (**Figure 4**).

Cluster 1 (Red; 25 items): Institutional digitalization and behavioral compliance

This cluster is located in the heart of the biggest community and focuses on tax compliance, tax administration, e-government, digitalization and trust. This stream indicates that perceived usefulness is connected to the technology acceptance model and institutional keywords, implying that the stream is not focused on infrastructure, but rather on the behavioral acceptance of AI technology. Shadow economy, tax evasion and corruption are closely located, but they are in Cluster 5, which generates a tension zone where enforcement logic clashes with modernization logic. COVID-19 also shows up here, making it clear that COVID-19 is indeed a driver of e-government compliance research in emerging economies.

Cluster 2 (Green; 22 items): Corporate financial governance and risk architecture

This cluster brings together tax research and mainstream finance in the fields of corporate governance, earnings management, financial performance and stock price crash risk. Digital finance and information asymmetry are the two themes that indicate scholarship on how AI is changing firm-level disclosure and tax avoidance, but not tax administration. Jordanian research does not seem to be grouped under the developing-country set in Cluster 1, but rather under "Jordan" as a country keyword, confirming that Jordanian research is focused on GCC corporate governance, not on public-sector tax administration.

Cluster 3 (Blue; 19 items): Algorithmic infrastructure and computational taxonomy

This cluster is the methodological center of the field at the visual and structural center, providing technical jargon to all other clusters, and comprising artificial intelligence, machine learning, big data, data mining, data analytics, and decision making. The self-reflexive presence of literature review, bibliometric analysis and information systems point to the ongoing paradigmatic self-definition of the field, as it maps its methodological limits. The reason that the United States is not included in regional-policy clusters is that the contributions made by the United States have been mostly methodological through algorithmic tools, and not context-specific governance studies.

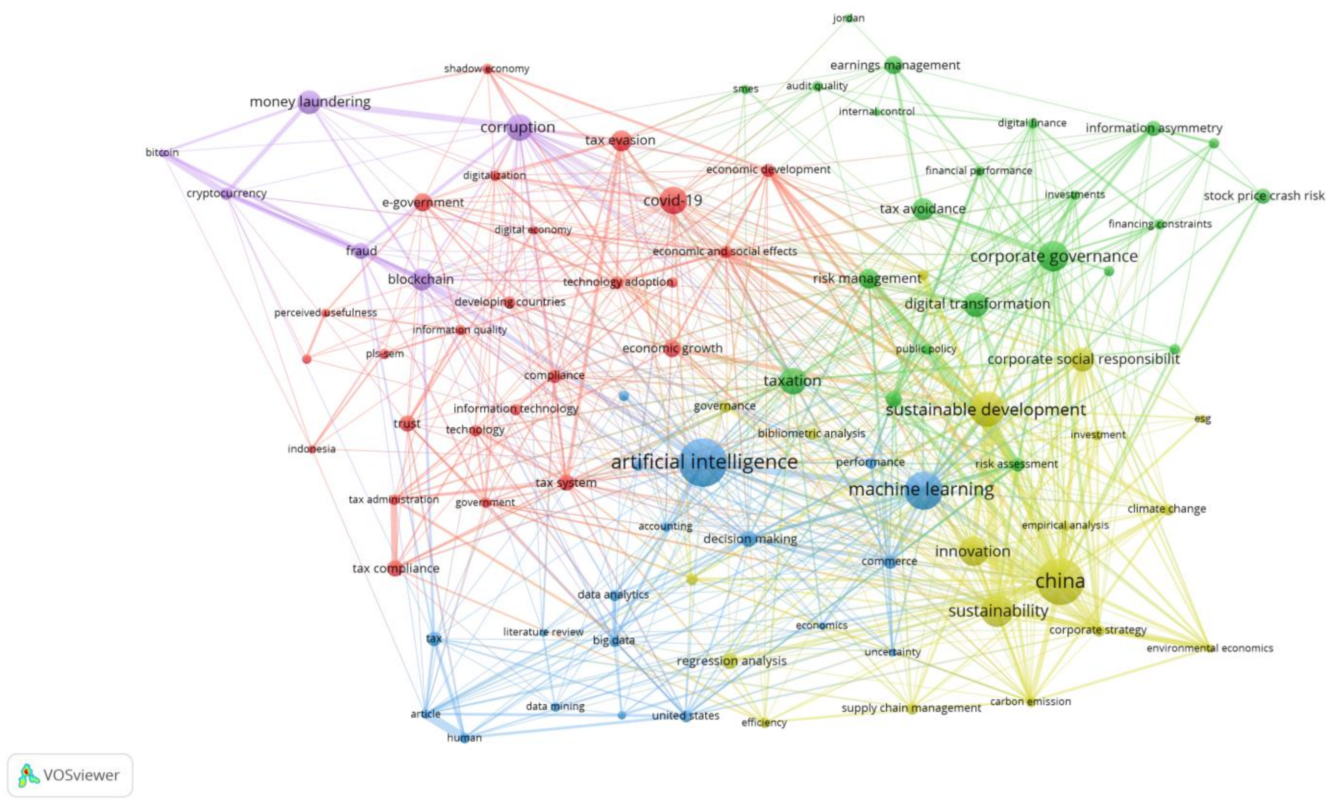

Figure 4. Keyword co-occurrence map (Source: Authors' own elaboration, using VOSviewer)

Cluster 4 (Yellow; 19 items): Sustainability governance and environmental fiscal policy

This cluster is the newest expansion of the field around sustainability, ESG, climate change, carbon emission, and environmental economics, and is densely connected to Cluster 2 through governance, innovation, and investment, but is located on the lower right periphery. It is noteworthy that China is not part of Cluster 1's developing-country group, as this suggests that China has become an arena for environmental tax experimentation and ESG disclosure regulations, but not a general compliance reform. Alongside empirical analysis, panel data and regression analysis are also emerging; this shows the speed with which econometric methods are being adopted in the field of research on sustainability-tax.

Cluster 5 (Purple; 6 items): Crypto-assets and illicit financial flows

This cluster is located on the extreme left-hand side, having the highest internal density and the lowest linkages from the outside, and includes bitcoin, blockchain, cryptocurrency, fraud, money laundering, and corruption. Theoretically, it has an important spatial separation from the central AI node, as blockchain analytics and ML enforcement share very few common references in the literature. The link to the wider network is mainly through fraud and corruption as boundary concepts of Cluster 1. The fact that this research is isolated from social science-oriented research suggests that the study of crypto-tax research is a methodological bubble within computer sciences and forensic tax accounting, rather than embedded in the social science of algorithmic tax governance.

Structural bridges and implications

The network structure is in line with the tripartite structure of the ATCGM. The left side (Clusters 1 and 5) represents the tension between efficacy and legitimacy in enforcement: enforcement keywords are spatially close but conceptually distinct from trust and procedural justice. In the center (Cluster 3), the algorithmic infrastructure that enables all three functions is present. The right side (Clusters 2 and 4) contains the sustainability paradox, as corporate governance and ESG are closely intertwined, and sustainability compliance is theorized through corporate governance and not administrative enforcement. Three boundary objects cross several clusters: digital transformation links Clusters 1 and 2, governance links Clusters 2 and 4, and taxation links Clusters 2 and 3. A structural gap with the central core is observed in the fine lines between Cluster 5 and the mainstream AI governance frameworks, where there is a lack of integrated language that connects blockchain-based illicit flows with governance concepts.

Keyword Overlay Visualization

The temporal overlay (**Figure 5**) shows a directionality of scholarly focus (dark blue = 2021; yellow = 2024), establishing a scholarly focus changing over time along the trajectory of enforcement, governance and sustainability.

Phase I (2021–2022): Enforcement, crypto-surveillance, and behavioral digitalization

The left periphery is in deep blue. In addition to the institutional-behavioral keywords of Cluster 1, such as shadow economy, tax evasion, corruption, and developing countries, Cluster 5 uses keywords concerning bitcoin, blockchain,

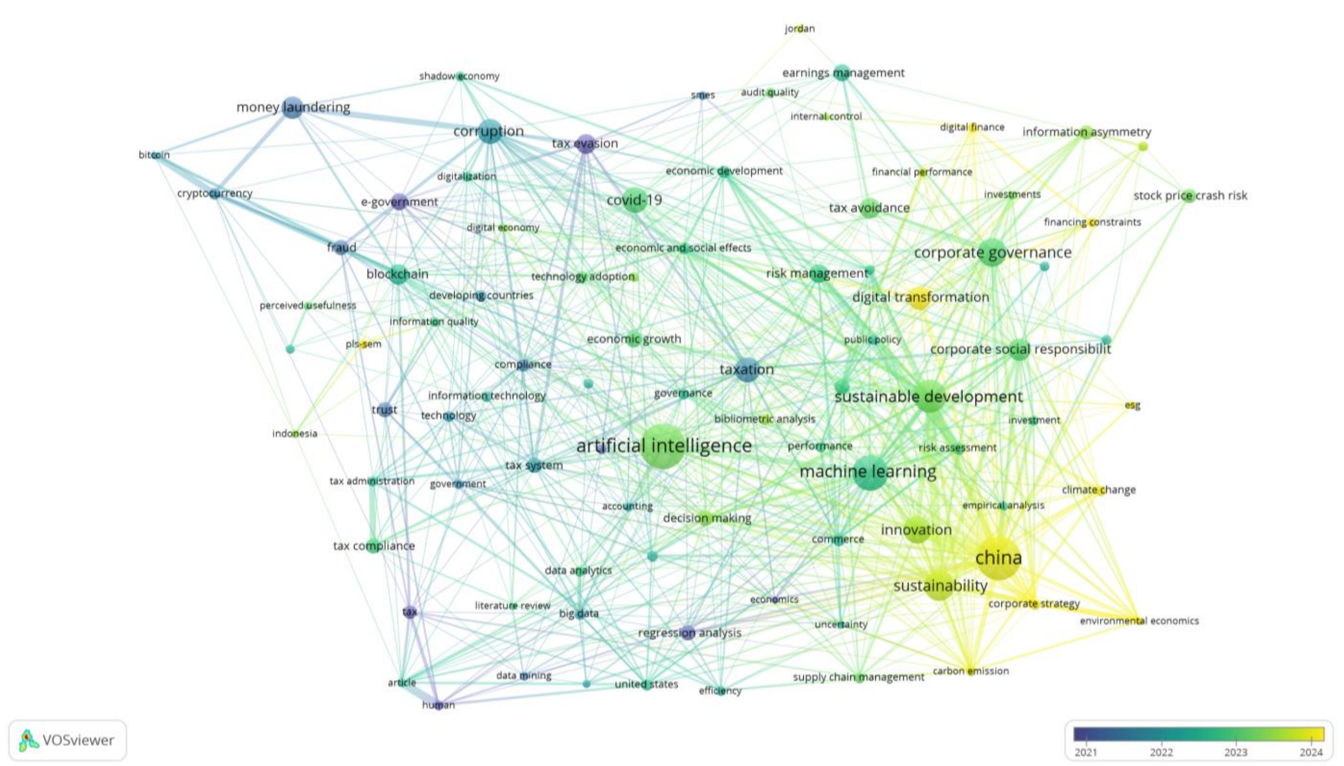

Figure 5. Keyword overlay visualization (Source: Authors' own elaboration, using VOSviewer)

cryptocurrency, money laundering and fraud in the 2021 band. The founding phase was centered on two fears, the taxing of crypto-assets and the digitization of administration for low-capacity environments. The co-temporal presence of technology acceptance model and perceived usefulness with e-government and tax compliance showed early research in a deterrence-cooperation binary: how can technology be used to expand the reach of compliance enforcement, or to encourage voluntary compliance, but not both at once?

Phase II (2022–2023): Algorithmic consolidation and methodological self-definition

The network center turns green and yellow green. This window is the perfect time for Cluster 3 keywords (AI, Machine Learning, Big Data, Data Analytics, Data Mining), with AI becoming an object of scholarly attention rather than just a tool. The emergence of accounting, information systems and economics in the same band is an attempt to legitimate the disciplines, given that business and social science journals took over computational techniques. At the boundary between Phase I and Phase II, COVID-19 acts as a temporal hinge between administrative urgency and methodological refinement. This range is also shaded by literature review and bibliometric analysis, which show that technical maturation took place at the same time as the self-reflexive mapping of the field.

Phase III (2023–2024): Sustainability turn and governance expansion

Bright yellow is illuminated in the right hemisphere. The keywords of Cluster 4 (sustainability, ESG, climate change, carbon emission, corporate social responsibility) are most prevalent in 2024, and Cluster 2 keywords (corporate governance, digital transformation, risk management) concentrate in 2023. It is a meaningful sequencing theoretically, because the establishment of governance frameworks was necessary first before environmental tax applications could be theorized. The 2024 arrival of China and supply chain management reinforces the geographical shift (Section *Most Relevant Affiliations and Regional Context*), with Chinese scholarship on carbon-border adjustments becoming part of the global body of knowledge. Sustainable development and governance approach seem to be located at the border of the 2023-2024 period, indicating that scholars now explicitly theorize the dual use of algorithmic infrastructure in terms of governance approaches.

Synthesis: From Bifurcation to Integration. The temporal gradient is not linear but directional. The cognitive core moved from the enforcement-oriented left in 2021 toward the sustainability-oriented right in 2024, whereas the technical core (Cluster 3) continues to be seen as persistent infrastructure over 2022-2023. The lack of yellow, however, on the far left, means Cluster 5 has stayed entirely blue, indicating that crypto-enforcement research has either stagnated or moved away from the mainstream, missing the governance and sustainability turns. On the other hand, saturation in Clusters 2 and 4, in the form of dense yellow, suggests that the upcoming cycle will be increasingly shaped by ESG, climate taxation and corporate governance. The three-phase pathway is in line with the idea of rapid thematic restructuring rather than linear cumulation. The 2021-2022 literature focused on AI's role in detecting tax fraud, the 2022-2023 literature focused on AI's role in tax administration, and the 2023-2024 literature focused on AI's role in ensuring sustainable taxation. In all phases, the interactions between these themes were renegotiated, and developed the concurrent paradoxes that are evident in the ATCGM, including efficacy vs. legitimacy, global asymmetry, and the sustainability paradox.

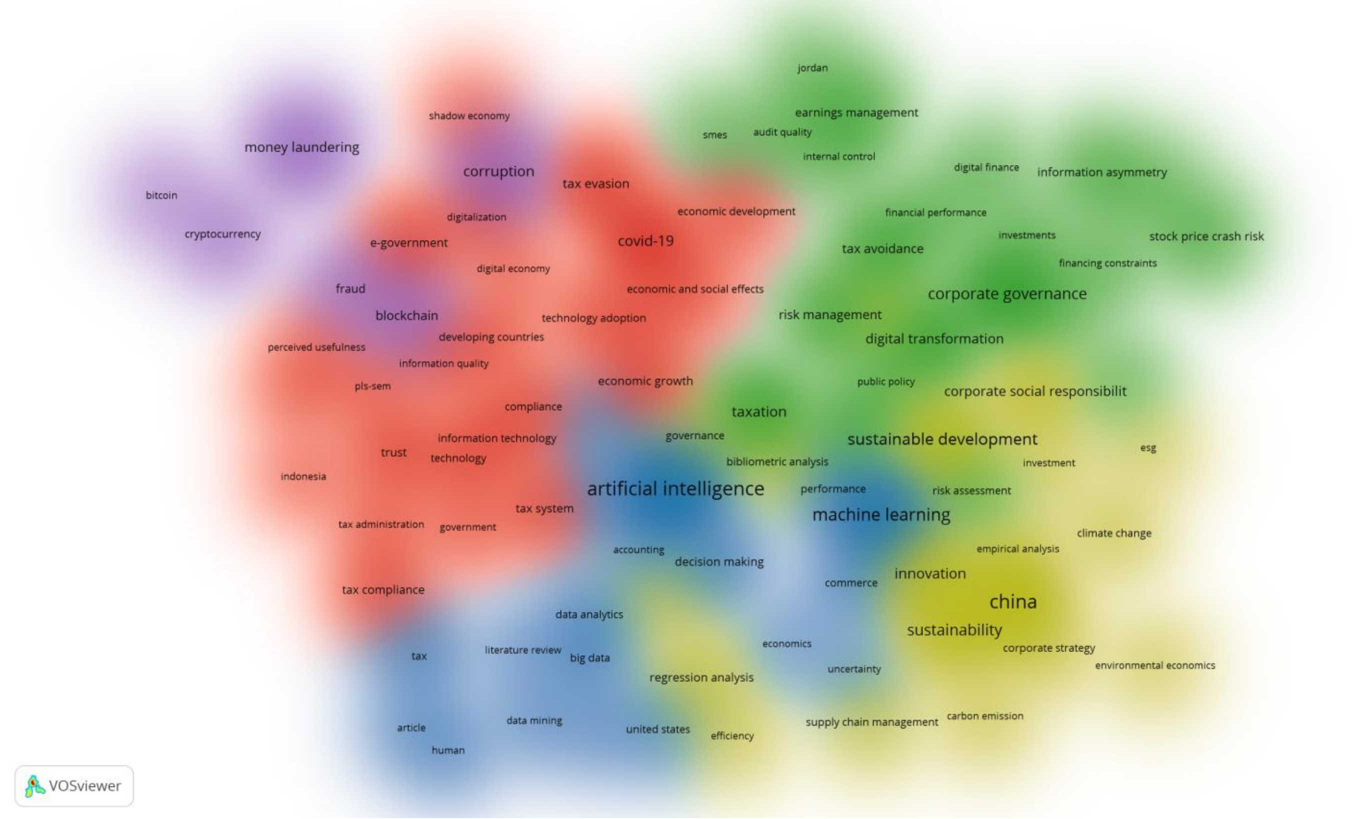

Figure 6. Thematic clusters (density visualization) (Source: Authors' own elaboration, using VOSviewer)

Thematic Density Map

The density visualization (**Figure 6**) also supports the four interconnected thematic domains that converge around the main theme of tax transformation via artificial intelligence, which are: (1) compliance/enforcement, (2) governance/accountability, (3) technology/innovation, and (4) policy/decision support. As discussed above in the operationalization deficit, the clusters of governance and policy are more dispersed, whereas the clusters of enforcement and technology are more dense and internally coherent in the visual density distribution. This architecture also validates the three-dimensional nature of the ATCGM with effectiveness (Cluster 1 & Cluster 3), legitimacy (Cluster 2 & Cluster 5) and sustainability (Cluster 4) being separate research spaces yet interdependent.

DISCUSSION

From Descriptive Mapping to Theory Building: The ATCGM

Given the above bibliometric evidence, this paper proposes an integrative analytical framework, the AI Tax Compliance Governance Matrix (ATCGM), that is composed of three interdependent dimensions: (1) Effectiveness, which covers technical performance, detection accuracy and audit efficiency, (2) Legitimacy, which covers algorithmic transparency, procedural justice and taxpayer trust, and (3) Sustainability, which covers ESG integration, environmental fiscal policy and long-run compliance culture. While there are existing bibliometric reviews that treat these dimensions separately, the ATCGM theorizes them as co-constituted by any use of AI in tax administration. The ATCGM captures a real-time simultaneity: when underpayment is detected by HMRC's Connect System, Effectiveness (revenue collected) is optimized, Legitimacy is potentially compromised if targeting patterns are not clearly known or if they are biased, and Sustainability is compromised if enforcement disadvantages compliant firms pursuing ESG objectives. In practice, a revenue agency deploying a machine-learning model for VAT audit selection applies the matrix through three simultaneous checks prior to rollout. Under the Effectiveness pillar, the agency reviews the model against random audit baselines and establishes a minimum detection threshold. Under the Legitimacy pillar, audit selection decisions must be explainable through SHAP values, with an appeal pathway for flagged taxpayers. Under the Sustainability pillar, the agency examines whether the risk model disproportionately burdens SMEs that comply with environmental tax incentives and adjusts the weightings accordingly. The three checks, which precede their rollout, are a practical ATCGM audit cycle. The model is moderated by four conditions: data quality, administrative capacity, legal accountability, and taxpayer AI literacy, or the understanding that taxpayers have about how algorithmic audit decisions are made, how they interpret risk assessments, and the ability to make decisions regarding their data in automated systems for compliance. The full **ATCGM structure** is detailed in **Table 4**.

It is important to recognize that there is a clear imbalance in the existing literature, as shown in **Table 4**: While there is a vast amount of empirical research on the model's accuracy, and on audit targeting, much less research has been conducted on Legitimacy and Sustainability. The four items in the

Table 4. AI tax compliance governance matrix (ATCGM): Pillars, risks, and research agenda

Pillar	AI-Tax Function	Main Risk	Future Research Priority	ATCGM Dimension
Predictive Effectiveness	Risk scoring, anomaly detection, audit selection	False positives, selective-label bias, revenue-only optimization	Longitudinal evaluation of audit yield, deterrence, and voluntary compliance	Effectiveness
Institutional Legitimacy	Explainable decisions, appealability, procedural fairness	Opacity, bias, weak contestability, reduced taxpayer trust	Fairness audits, SHAP/LIME explainability, taxpayer perception studies	Legitimacy
Sustainability Orientation	ESG monitoring, carbon-tax verification, supply-chain analytics	Greenwashing detection; compliance burden on SMEs	Integrated environmental tax data and proportional governance safeguards	Sustainability
Moderating Capabilities	Data quality, administrative capacity, legal accountability, AI literacy	AI dependency, poor data governance, unequal taxpayer understanding	Comparative case studies across mature, reforming, and emerging administrations	All three

Moderating Capabilities row are often cited in the literature as separate issues, but no study in the corpus looks at them simultaneously. This indicates that the main value of the ATCGM is not just descriptive but diagnostic, showing the gaps in structure, not just a list of gaps.

How AI Is Actually Deployed: From Technical Tools to Administrative Systems

Four main tax administration modalities of AI have been recorded in the literature. First, supervised machine learning for audit targeting: Random Forest classifiers (Breiman, 2001) and gradient boosting machines (Friedman, 2001) are trained to score the risk of taxpayers from historical audit results, with a reported accuracy that reaches 93% in VAT fraud detection contexts (Belahouaoui & Alm, 2025). Second, anomaly detection for real-time monitoring: isolation forest algorithms (Liu et al., 2008) can detect statistical outliers in transaction data streams so that auditing can be done continuously instead of periodically (Alles et al., 2006). Third, automation of correspondence using NLP: NLP systems categorize taxpayer correspondence, provide specific replies, and flag high-risk correspondence. Fourth, graph neural networks for money-laundering and illicit financial flow analysis: mapping shell-company ownership structures and cryptocurrency transaction graphs to uncover coordinated evasion schemes.

There are different governance issues associated with each modality. The supervised learning systems incorporate historical enforcement patterns, which can continue to reinforce income-based targeting bias. Anomaly detection is based on statistical thresholds whose legality is in dispute (Kroll et al., 2017). NLP correspondence systems may systematically misclassify queries from low-literacy taxpayers. The analysis of cryptocurrency flows requires graph analysis and cross-border data sharing, which exceed existing legal frameworks. The governance challenges need to be addressed in the analytical space of the ATCGM's Legitimacy dimension.

The Efficacy–Legitimacy Paradox

This efficiency-equity paradox also reflects a governance failure, with fiscal consequences, as long as the notion of perceived equity is lowered, the chances of voluntary compliance also drop, to varying degrees, offsetting gains in enforcement (Kirchler et al., 2008; Murphy, 2005). This study was able to build upon the observations that were previously made: Belahouaoui and Attak (2024) focused on the technical use of AI in tax administration, and Thaha et al. (2023) explored drivers of compliance among SMEs, but both were not in a position to demonstrate that legitimacy and enforcement keywords cluster in tension rather than in concert, a structural pattern visible through network mapping.

If mechanisms for building trust among taxpayers are not in place, taxpayers who are willing to provide information voluntarily will adopt enforced compliance in response to an erosion of trust, resulting in higher administrative expenses and a lower intrinsic motivation. This is complicated by the impact of AI: the algorithmic auditing process may feel like a surveillance operation, rather than a service, and can further erode the compliance culture that it should be safeguarding. Given the use of AI for tax decision-making, tax authorities should therefore be co-developing explainability infrastructure, as well as tools like SHAP values (Lundberg & Lee, 2017), LIME (Ribeiro et al., 2016) and procedural justice communication, which would make the algorithmic decisions understandable and contestable.

Geographic Knowledge Asymmetry and Its Consequences

The geographic inversion mentioned in Section *Most Relevant Affiliations and Regional Context* does not only have bibliometric interest. The asymmetry is that the literature comes from emerging economies and the systems are implemented by OECD economies, and that the knowledge most useful in operational decision making in this sector comes from those who have the least access to operational data. This presents a paradox for developing-economy tax officials trying to find the most usable evidence from other countries, because the most applicable evidence comes from peers with limited institutional capacity, while the most scalable systems are mainly documented in administrative reports that are not readily available in academic circles. This gap is also highlighted by research on VAT transitions in developing economies, which reveals that indirect tax instruments are effective only if there is adequate administrative infrastructure (Adandohoin, 2021). The contradiction is exacerbated by the fact that developing country tax authorities are experiencing simultaneous pressure to modernize the way revenue is collected, and the need to reduce compliance gaps, while not having the institutional anchors that OECD tax authorities enjoy (Belahouaoui & Attak, 2024).

Evidence from Peru shows that the rollout of e-invoicing led to an initial increase in firm-declared VAT liabilities by more than 5% in the first year and network spillovers decreased false purchase invoices by 2.3% (Bellon et al., 2022, 2023). A second line of evidence shows that beyond tax compliance, e-invoicing implementation leads to wider digitalization spillovers, such as the increase of data governance within firms and the decrease in information asymmetries between firms and tax authorities (Bojanc et al., 2024). These studies are however focused on e-invoicing as a precursor to AI-enabled analytics, and not on machine learning systems themselves, which highlights the causal vacuum at the core of the field. This gap between digitalization precursors and full AI deployment is especially important for multinational enterprises, whose cross-border structuring opportunities expand as developing-country authorities build up their digital capacity (Chen & Lei, 2025). Prior reviews were not in a position to map this knowledge asymmetry at scale: the present study's co-authorship network analysis, covering 8,114 authors across 664 journals, is the first to reveal that the Jordan–Saudi Arabia–Malaysia corridor constitutes a distinct horizontal knowledge-transfer channel that bypasses OECD diffusion models entirely, a finding with direct implications for how international technical assistance programs should be designed.

The ESG–Tax Nexus: Integration, Not Diffusion

The time frame evidence indicates real, albeit early, integration between AI tax monitoring and sustainability governance. The sequence from Section *Keyword Overlay Visualization*, where governance frameworks appear before environmental applications, suggests that sustainability applications build upon existing AI governance frameworks instead of just tacking the word "environment" onto existing tax papers. The concrete integration mechanism is dual-use algorithmic monitoring, the same monitoring systems that can detect tax underreporting can also be used for monitoring ESG disclosure compliance, reporting on carbon emissions, and the manipulation of transfer prices by MNCs. Empirical evidence demonstrates that digitalization of tax administration not only curtails the opacity of information but, at the same time, enhances accountability channels and, consequently, corporate ESG performance (Hai et al., 2024). The clearest example is China's carbon-border adjustment experimentation, which appears in the 2024 cohort. In the eyes of the tax authorities, this convergence also indicates that investments in AI-enabled enforcement systems can have co-benefits in environmental tax monitoring, as long as these systems are conceived from the ground up with the potential for dual-use. As illustrated in Wang et al. (2025) and Yanto et al. (2025), there is a bidirectional relationship between tax behavior and sustainability commitments: tax-aggressive firms have systematically lower ESG engagement, while firms with high ESG engagement have higher voluntary compliance. The bidirectional nature of this phenomenon further underscores the importance of integrated monitoring architectures, as AI tools that identify tax underreporting can also highlight discrepancies between reported ESG pledges and actual financial actions. In addition, environmental protection taxes have been found to have the potential to drive corporate digital transformation, with implications for the technological readiness needed to facilitate AI-driven compliance efforts (Zhang et al., 2025). The convergence is already being realized in a number of ways through audit technologies: for instance, through satellite-image verification of GHG emissions disclosures, which offers a tangible example of dual-use algorithmic audit in which tax and sustainability data are tracked via a common analytical path (Gu et al., 2023).

Practical Implications for Tax Authorities

The results apply to five operational priorities. First, introduce AI in the areas where there is plenty of data (VAT, e-invoicing, customs) and then expand it to other areas, such as income-tax applications. The findings from IRS audit data provide empirical support that big-data-based selection procedures can enhance compliance rates, but that taxpayer perceptions of fairness and procedural legitimacy mediate the impact of such selection procedures on compliance (Grana et al., 2025; Neuman & Sheu, 2022). Second, any audit system that relies on AI should have SHAP-value reporting and algorithmic impact assessments built into it as part of its basics, before being scaled up for enforcement use. The ethical challenges of using AI systems that are not transparent are significant: public acceptance of algorithmic decision making relies on perceived transparency, explainability and contestability (Kieslich et al., 2022; Nannini et al., 2023). Thirdly, put in place governance policies that define data sources, model-update cycles, appeal processes, and safeguards to prevent discriminatory targeting. In terms of tax reform, stakeholder evidence suggests that weak trust in opaque procedures in the tax system is one of the most frequent causes of tax non-compliance (Hasan et al., 2024). Fourth, take a more taxpayer-service approach to the use of AI, preventing inadvertent mistakes, rather than ramping up enforcement. Studies have consistently shown that the relationship between trust in government, perceived procedural fairness and tax morale is independent of deterrence effects (Anjarwi et al., 2025; Supriyati et al., 2024). Fifth, build environmental and ESG-related tax incentive monitoring into the design of incentives, up front, to avoid greenwashing and to safeguard public revenues. Beyond compliance, the adoption of AI in tax administration has been linked to externalities such as an improved investor perception of risk and a more stable supply of market signals, suggesting that implementing AI with a focus on transparency can have wider repercussions beyond direct compliance benefits (Tu et al., 2025).

In times of conflict and/or post-reform situations (Ukraine, post-conflict economies): define transparency protocols before predictive risk scoring so as to avoid losing trust in fragile institutional environments. One institutional design that can build verifiable transparency by design is the blockchain-based tax system (Kabir, 2021). In advanced economies (HMRC, IRS, BZSt): integrate explainable AI into current systems without undermining the credibility and trust of the processes already established over many years. Findings from audit digitalization evidence drawn from China show that digital monitoring by the government has a positive impact on the internal control weaknesses of large firms, while the effects are heterogeneous across types of enterprises and need to be calibrated at the institutional level (Shi & Zhang, 2025). For international institutions (OECD FTA, IMF FAD): develop differentiated international rules for transparency in tax administration, based on institutional capacity, and not on uniform technical requirements. When the technical implementation is paired with institutional readiness, big-data-driven tax collection systems have been proven to boost the quality of corporate accounting information, highlighting the need for careful sequencing between technology and institutions (Zhou & Hao, 2025).

Research Agenda: Five Empirically Grounded Gaps

The bibliometric evidence shows five missing points in the structure which should inspire further investigation.

The Efficacy–Equity Paradox. Ongoing and future research will be required to create fairness-aware targeting algorithms, longitudinal studies of voluntary compliance over 3–5 year time horizons, and multi-jurisdictional bias audits. Pre-deployment algorithmic impact assessments should be published by tax authorities.

Geographic Knowledge Asymmetry. The volume of literature is generated in emerging economies, and OECD centers put in place the systems, resulting in a disconnect between volume and impact. Cross-jurisdictional comparative research across different institutional capacities is a research priority, while translating the operational evidence of systems like China's Golden Tax System III into a readily accessible scholarly form is also a research priority.

The Longitudinal–Behavioral Void. Despite 1,803 articles, none of the studies document how taxpayers act before and after the deployment of AI systems. There is a need for partnerships with revenue administrations to gain access to administrative microdata and mixed methods designs that involve panel data as well as taxpayer and inspector experience.

Crypto-Enforcement Integration. The temporal isolation of Cluster 5, "frozen at the 2021 band", is a warning sign of the disconnect between blockchain analytics and the wider application of AI governance. With the expansion of digital asset markets, the integration of crypto-asset transaction monitoring with explainable audit frameworks is becoming a high priority.

The Operationalization Deficit. There are myriad measurements of model accuracy in the literature, but little guidance for the tax inspector's daily use of AI for audits. Future work should yield: (1) explainability protocols tailored to the tax appeal legal context; (2) human-AI collaboration protocols that define conditions for overrides; and (3) governance protocols that allow non-technical officials to assess AI proposals.

CONCLUSION

This study traces the intellectual landscape of tax compliance research with AI over 1,803 peer-reviewed articles (2015–2025) by making three contributions. First, it maps three discernible thematic waves, enforcement, governance, and sustainability, that are compressed into the last ten years, suggesting rapid reorientation rather than linear accumulation. Second, it presents the AI Tax Compliance Governance Matrix (ATCGM), where effectiveness, legitimacy, and sustainability are shown to be interdependent as aspects of AI used in tax administration. Third, it outlines five research gaps from a network structural analysis that provides a priority program for the field.

As generative AI enters the field of taxpayer service and policymaking, the balance between efficiency and equity in algorithmic tax governance will become even more dynamic. The geographical distribution of research shows an over-representation of developing economies, particularly in Eastern Europe, the Middle East and Southeast Asia, in the production of knowledge, thereby supporting a leapfrogging trend driven by a need for reform and not administrative maturity. This is an indication that global models of capacity development designed for OECD countries may not be applicable in emerging economies.

There is significant potential for using AI to modernize tax systems, with potential benefits for better enforcement, evidence-based policymaking, and governance. The evidence in the bibliometric analysis is self-evident: effectiveness that is gained through AI is only lasting if legitimacy is established from the beginning and sustainability aspects are considered from the design stage and not after. For tax authorities, the takeaway is that with no transparency processes, appeals, or bias audits, AI adoption can negatively impact voluntary compliance. As AI technology advances in the coming decade, the real challenge is not whether AI can detect noncompliance, but whether it can be governed in a way that maintains taxpayer confidence and fiscal equity.

The main limitations of our study are that it includes only journals indexed in Scopus, only English-language publications, only journals with Q1–Q2 ratings, and citation lag effects. There is a need for further research utilizing bibliometrics with interviews, case studies, administrative microdata and the longitudinal evaluation of actual AI deployments to determine not only whether AI can detect noncompliance, but also whether AI can create more effective, legitimate, and sustainable tax systems.

Author contributions: **HZ:** conceptualization, methodology, data collection, writing – original draft, visualization; **SS:** validation, formal analysis, writing – review & editing, supervision. Both authors agreed with the results and conclusions.

Funding: The first author benefits from the PhD Associate Scholarship provided by the National Centre for Scientific and Technical Research (N* = 135USMBA2024).

Acknowledgements: The first author would like to thank the National Centre for Scientific and Technical Research (CNRST) for its support.

Ethical statement: The authors stated that the bibliometric study used publicly accessible data retrieved from the Scopus database. Ethical approval was not required because it involved no human participants, personal information, or experimental interventions. The data were accessed through authorized academic means and analyzed according to established research ethics. All sources were appropriately cited and acknowledged.

AI statement: The authors stated that they have not used any generative AI or AI-assisted technologies in the preparation of this manuscript.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from corresponding author.

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M. ... Zheng, X. (2015). TensorFlow: Large-scale machine learning on *heterogeneous distributed systems [Preprint]. arXiv*. https://doi.org/10.48550/arXiv.1603.04467

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138-52160. https://doi.org/10.1109/ACCESS.2018.2870052

Adandohoin, K. (2021). Tax transition in developing countries: Do value added tax and excises really work? *International Economics and Economic Policy, 18*(2), 379-424. https://doi.org/10.1007/s10368-021-00492-8

Aggarwal, S. (2024). The role of artificial intelligence in tax administration and compliance: A new era of digital taxation. *Educational Administration: Theory and Practice, 30*(1), 3831-3837. https://doi.org/10.53555/kuey.v30i1.7581

Alles, M., Brennan, G., Kogan, A., & Vasarhelyi, M. A. (2006). Continuous monitoring of business process controls. *International Journal of Accounting Information Systems, 7*(2), 137-161. https://doi.org/10.1016/j.accinf.2005.10.004

Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics, 1*(3-4), 323-338. https://doi.org/10.1016/0047-2727(72)90010-2

Alm, J., Kirchler, E., & Muehlbacher, S. (2012). Combining psychology and economics in the analysis of compliance. *Economic Analysis and Policy, 42*(2), 133-151. https://doi.org/10.1016/S0313-5926(12)50016-0

Anjarwi, A. W., T, S., Baridwan, Z., & Iqbal, S. (2025). The deterrent role of social norms on tax evasion in Indonesia. *Deviant Behavior, 46*(7), 927-946. https://doi.org/10.1080/01639625.2024.2375017

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

Aslett, J., Hamilton, S., Gonzalez, I., Hadwick, D., & Hardy, M. A. (2024). *Understanding artificial intelligence in tax and customs administration* (Technical Notes and Manuals No. 2024/006). International Monetary Fund. https://doi.org/10.5089/9798400290435.005

Belahouaoui, R., & Alm, J. (2025). Tax fraud detection using artificial intelligence-based technologies: Trends and implications. *Journal of Risk and Financial Management, 18*(9), Article 502. https://doi.org/10.3390/jrfm18090502

Belahouaoui, R., & Attak, E. H. (2024). Digital taxation, artificial intelligence and Tax Administration 3.0: Improving tax compliance behavior, a systematic literature review using textometry (2016–2023). *Accounting Research Journal, 37*(2), 172-191. https://doi.org/10.1108/ARJ-12-2023-0372

Bellon, M., Dabla-Norris, E., Khalid, S., & Lima, F. (2022). Digitalization to improve tax compliance: Evidence from VAT e-invoicing in Peru. *Journal of Public Economics, 210*, Article 104661. https://doi.org/10.1016/j.jpubeco.2022.104661

Bellon, M., Dabla-Norris, E., & Khalid, S. (2023). Technology and tax compliance spillovers: Evidence from a VAT e-invoicing reform in Peru. *Journal of Economic Behavior & Organization, 212*, 756-777. https://doi.org/10.1016/j.jebo.2023.06.004

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 35*(8), 1798-1828. https://doi.org/10.1109/TPAMI.2013.50

Bird, R. M., & Zolt, E. M. (2008). Technology and taxation in developing countries: From hand to mouse. *National Tax Journal, 61*(4), 791-821.

Bojanc, R., Pucihar, A., & Lenart, G. (2024). E-invoicing: A catalyst for digitalization and sustainability. *Organizacija, 57*(1), 3-19. https://doi.org/10.2478/orga-2024-0001

Braithwaite, V. (Ed.). (2003). *Taxing democracy: Understanding tax avoidance and evasion*. Ashgate.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Chen, Y., & Lei, S. (2025). Tax avoidance opportunity for multinational enterprises: Effects of digitalized tax administration in China. *Journal of International Financial Markets, Institutions and Money, 102*, Article 102177. https://doi.org/10.1016/j.intfin.2025.102177

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

Feld, L. P., & Frey, B. S. (2002). Trust breeds trust: How taxpayers are treated. *Economics of Governance, 3*(2), 87-99. https://doi.org/10.1007/s101010100032

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5), 1189-1232. https://doi.org/10.1214/aos/1013203451

Grana, J., Sarzosa, M., Plumley, A., & Rodriguez, D. (2025). *The comprehensive general indirect effect of auditing individual income tax returns* (IRS SOI Working Paper). U.S. Department of the Treasury, Internal Revenue Service.

Gu, Y., Dai, J., & Vasarhelyi, M. A. (2023). Audit 4.0-based ESG assurance: An example of using satellite images on GHG emissions. *International Journal of Accounting Information Systems, 50*, Article 100625. https://doi.org/10.1016/j.accinf.2023.100625

Hai, B., Shi, H., Piao, T., & Dou, Z. (2024). Digitalization of tax administration and its impact on corporate ESG performance. *Finance Research Letters, 68*, Article 105791. https://doi.org/10.1016/j.frl.2024.105791

Hasan, A., Sheikh, N., & Farooq, M. B. (2024). Exploring stakeholder perceptions of tax reform failures and their proposed solutions: A developing country perspective. *Meditari Accountancy Research, 32*(3), 721-755. https://doi.org/10.1108/MEDAR-03-2023-1961

Kabir, M. R. (2021). Behavioral intention to adopt blockchain for a transparent and effective taxing system. *Journal of Global Operations and Strategic Sourcing, 14*(1), 170-201. https://doi.org/10.1108/JGOSS-08-2020-0050

Kieslich, K., Keller, B., & Starke, C. (2022). Artificial intelligence ethics by design: Evaluating public perception on the importance of ethical design principles of artificial intelligence. *Big Data & Society, 9*(1). https://doi.org/10.1177/20539517221092956

Kirchler, E. (2007). *The economic psychology of tax behaviour*. Cambridge University Press. https://doi.org/10.1017/CBO9780511628238

Kirchler, E., Hoelzl, E., & Wahl, I. (2008). Enforced versus voluntary tax compliance: The "slippery slope" framework. *Journal of Economic Psychology, 29*(2), 210-225. https://doi.org/10.1016/j.joep.2007.05.004

Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Yu, H. (2017). Accountable algorithms. *University of Pennsylvania Law Review, 165*(3), 633-705.

Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. In *2008 Eighth IEEE International Conference on Data Mining* (pp. 413-422). IEEE. https://doi.org/10.1109/ICDM.2008.17

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in neural information processing systems* (Vol. 30, pp. 4765-4774).

McKinsey & Company. (2021). *The future of tax: How AI and automation are transforming tax functions*.

Murphy, K. (2005). Regulating more effectively: The relationship between procedural justice, legitimacy, and tax non-compliance. *Journal of Law and Society, 32*(4), 562-589. https://doi.org/10.1111/j.1467-6478.2005.00338.x

Nannini, L., Balayn, A., & Smith, A. L. (2023). Explainability in AI policies: A critical review of communications, reports, regulations, and standards in the EU, US, and UK. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1198-1212). Association for Computing Machinery. https://doi.org/10.1145/3593013.3594074

Neuman, E. L., & Sheu, R. J. (2022). Big data analytics in IRS audit procedures and its effects on tax compliance: A moderated mediation analysis. *Journal of the American Taxation Association, 44*(2), 97-113. https://doi.org/10.2308/JATA-2020-038

OECD. (2015). *Addressing the tax challenges of the digital economy, Action 1: 2015 final report* (OECD/G20 Base Erosion and Profit Shifting Project). OECD Publishing. https://doi.org/10.1787/9789264241046-en

OECD. (2020). *Tax Administration 3.0: The digital transformation of tax administration*. OECD Publishing. https://doi.org/10.1787/ca274cc5-en

OECD. (2023). *Tax Administration 2023: Comparative information on OECD and other advanced and emerging economies*. OECD Publishing. https://doi.org/10.1787/900b6382-en

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*, Article n71. https://doi.org/10.1136/bmj.n71

Podsakoff, P. M., MacKenzie, S. B., Podsakoff, N. P., & Bachrach, D. G. (2008). Scholarly influence in the field of management: A bibliometric analysis. *Journal of Management, 34*(4), 641-720. https://doi.org/10.1177/0149206308319533

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135-1144). Association for Computing Machinery. https://doi.org/10.1145/2939672.2939778

Shi, W., & Zhang, Z. (2025). Exploring the role of digital governance: The effect of audit digitalization on firms' internal control weaknesses in China. *International Journal of Accounting Information Systems, 56*, Article 100756. https://doi.org/10.1016/j.accinf.2025.100756

Supriyati, S., Hapsari, I., & Nahumury, J. (2024). The influence of trust in the government, perceived fairness, and tax morale on taxpayer compliance: Implications for budget formation. *Public and Municipal Finance, 13*(2), 129-139. https://doi.org/10.21511/pmf.13(2).2024.11

Thaha, A. R., Antoro, A. F. S., & Muhtarom. (2023). Analytical review of tax compliance studies in the SMEs sector: A bibliometric approach. *Journal of Tax Reform, 9*(3), 398-412. https://doi.org/10.15826/jtr.2023.9.3.149

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management, 14*(3), 207-222. https://doi.org/10.1111/1467-8551.00375

Tu, W., Du, A. M., & Saddler, S. B. (2025). Digital tax administration, investor risk perception, and stock return volatility. *International Review of Economics & Finance, 103*, Article 104434. https://doi.org/10.1016/j.iref.2025.104434

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

Veale, M., & Binns, R. (2017). Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data. *Big Data & Society, 4*(2). https://doi.org/10.1177/2053951717743530

Wang, X., Chan, K. F., Chang, M., Shan, Y. G., & Yang, J. W. (2025). Does tax-aggressive behavior motivate more CSR engagement? *Economics & Politics, 37*(2), 669-684. https://doi.org/10.1111/ecpo.12333

Yanto, H., Hajawiyah, A., & Baroroh, N. (2025). The reciprocal effect of environmental, social, and governance (ESG) practices and tax aggressiveness in Indonesian and Malaysian companies. *Problems and Perspectives in Management, 23*(1), 339-351. https://doi.org/10.21511/ppm.23(1).2025.25

Zhang, X., Liu, Y., Yu, S., Lin, O., & Meng, L. (2025). Impact of environmental protection tax on enterprise digital transformation: Evidence from Chinese listed firms. *International Review of Economics & Finance, 97*, Article 103743. https://doi.org/10.1016/j.iref.2024.103743

Zhou, Y., & Hao, Y. (2025). Big data tax collection and administration and quality of corporate accounting information. *Finance Research Letters, 85*, Article 108041. https://doi.org/10.1016/j.frl.2025.108041

Zupic, I., & Čater, T. (2015). Bibliometric methods in management and organization. *Organizational Research Methods, 18*(3), 429-472. https://doi.org/10.1177/1094428114562629